

시맨틱 웹을 이용한

분산 시각미디어 검색 프레임워크 아키텍처

손영수^o 양명미 김정선
한양대학교

{ysson^o, mmyang, jskim}@cse.hanyang.ac.kr

Distributed Visual Media Retrieval Framework Architecture Using Semantic Web

YoungSu Son^o MyoungMi Yang Jungsun Kim
Dept. of Computer Engineering, Hanyang University

요 약

대부분의 Ontology를 이용한 이미지 검색 시스템들은 중앙 집중적(Centralization)으로 구성되어 있어, 단일, 소수의 이미지 제공자의 특화된 Semantic기반의 이미지 검색을 제공한다. 이에 반해 간단한 구조를 가진 메타데이터 기반의 분산 이미지 검색 시스템들은 이미지 제공자들의 자율성을 보장하기 용이하지만, Semantic 기반의 이미지 검색을 제공하지 못한다. 본 논문은 위 시스템들의 단점들을 극복하기 위해 다양한 이미지 제공자들의 자율성을 보장하면서, Semantic 기반의 이미지 검색을 지원하는 프레임워크의 아키텍처를 제안한다.

1. 서론

본 논문은 과학 재단의 특정 기초 과제인 “시맨틱 웹을 이용한 지식 기반 분산 시각미디어 검색 프레임워크”를 위해 제안된 아키텍처에 관한 논문이다.

이미지를 검색하는 방법은 메타 데이터 기반과 콘텐츠 기반으로 나뉘지만, 콘텐츠 기반으로 이미지 검색에는 많은 한계가 있어 메타 데이터 기반의 이미지 검색에 관한 많은 연구가 이루어졌다. 하지만 Photo-Book[1], QBIC[2]같은 메타데이터 시스템들은 이미지 제공자들의 자율성을 보장하지 못하는 단점들이 있었다.

이를 보완하기 위해 Chariot System[3]은 다양한 이미지 제공자들을 지원하는 구조를 제안했지만, 메타 데이터 기반의 구조이기 때문에 본 프로젝트의 요구에 부합되는 Semantic 기반의 검색을 지원하지 못하였다.

2000년도 이후 Semantic Web의 ontology를 이용한 이미지 검색시스템[4][5]에 대한 많은 연구가 진행되어, Semantic 기반의 검색이 가능해졌다. 하지만 Ontology기반의 시스템은 단일 또는 소수의 이미지 제공자들만 지원하여 제한된 도메인의 이미지 검색만 가능하다.

여, 다양한 이미지 제공자들의 자율성을 보장하면서 Semantic 기반의 이미지 검색을 지원하는 KVRF (Knowledge-based Distributed visual Media Retrieval Framework)의 아키텍처를 제안한다.

2. 기반기술

2.1 Web Service 기반 아키텍처

다양한 이미지 제공자들을 등록 및 관리하기 위해서 플랫폼 독립적인 Web Services[6]기반으로 구성한다.

2.2 지식 기반의 Image Search 플랫폼

Semantic 기반의 이미지 검색을 위해서 Metadata가 아닌 Semantic Web의 Ontology[7]를 이용한다.

2.3 워크플로우 기반 아키텍처

Client의 Query에 따라 컴포넌트의 호출순서가 다양하게 변경되어야 하므로, Ontology기반의 Workflow 생성모델[8]을 이용하여 Query에 따라 유동적으로 호출순서가 생성되도록 한다.

3. Architecture

우리가 제안한 KVRF의 전체적인 Architecture는 [그림2]와 같다.

3.1 KVRF

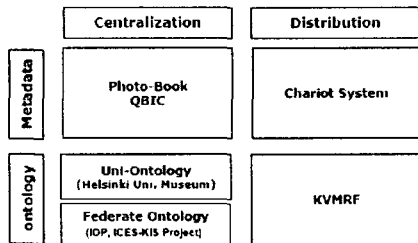
Image Service Broker로서 클라이언트로부터 Query를 입력받아 이미지 제공자로부터 적합한 이미지를 검색해 주는 Broker의 역할을 한다.

3.1.1 Ontology 관련 컴포넌트

- Common MetaData Ontology - 여러 가지 Domain에 공통적으로 속하는 MetaData들을 가지고 있다.
- Service Ontology - Service Provider들이 제공하는 Domain에 대한 정보가 저장되어 있으면 Semantic Web Service에서 사용되는 Matchmaking방법을 사용하고 있다.[9]

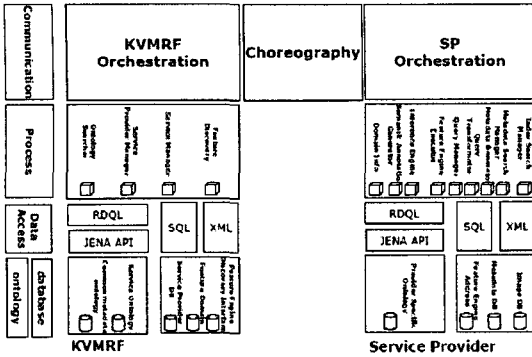
3.1.2 Database 관련 컴포넌트

- Feature Engine Discovery Interface - Feature Engine들의 위치 및 실행시키기 위한정보들을 저장한다.
- Service Provider DB - Service Provider들이 어떤 Domain을 제공하는



[그림 1] 이미지 검색시스템의 분류

본 논문은 위에 열거된 시스템[1],[2],[3],[4],[5]들의 단점을 보완하



[그림 2] KVMRF의 전체 구성도
 이에 대한 정보와 Service Provider들의 실제 위치정보를 저장한다.

3.1.3 Process관련 컴포넌트

- Feature Discovery - Domain에 따라 추출해야 할 Feature 정보들을 가져온다.
- Ontology Searcher - Common MetaData Ontology를 검색한다.
- Service Finder - Client로부터 입력받은 Query를 이용해 Service Provider DB에서 적합한 Provider를 검색한다
- Service Provider Manager - Domain과 Service Provider의 주소집 입력받아 Service Provider DB에 저장한다.

3.2 Service Provider

3.2.1 Ontology 관련 컴포넌트

- Provider Specific Ontology - Service Provider들이 가지고 있는 Image들을 Semantic 기반으로 검색하기 위하여 Domain Specific한 Knowledge들을 가지고 있다.

3.2.2 Database 관련 컴포넌트

- Image DB - Service Provider들이 가지고 있는 Image들을 저장한다.
- MetaData DB - Service Provider들이 가지고 있는 Image들에 대한 MetaData 정보를 저장한다.

3.2.3 Process관련 컴포넌트

- Domain Info - Service Provider에서 제공하는 Image Domain 정보란 필요한 Component에게 전달한다.
- Feature Engine - Image들의 특정 Feature들을 추출한다.
- Feature Engine Executor - Feature Engine들을 실행한다.
- Inference Engine - Query RDQL을 입력받아 Ontology를 기반으로 여러 가지 가능한 다른 Query들을 추론한다.
- Index Service Manager - Image Reference를 입력받아 각 Reference에 해당하는 Image들을 Image DB에서 검색한다.
- MetaData Generator - Image DB에서 Image들을 가져와 각 Image들에 대한 Feature들을 추출하고, 결과를 MetaDataDB에 저장한다.
- MetaData Search Manager - RDQL을 입력받아 Inference Engine을 실행시켜 여러 가지 가능한 RDQL을 가져온 후, 이를 Query Transformer에게 전달하고 실행한다
- Query Manager - Feature Value와 기존의 RDQL을 입력받아 Feature Value가 포함되어 있는 Modify된 RDQL을 생성한다.
- Query Transformer - RDQL은 DB에 직접 Query로 쓸수 없으므로, RDQL을 XML로 다시 SQL로 변환 시킨 후, 원하는 DB를 검색한다.
- Semantic Annotation Generator - Common MetaData Ontology를 이용하여 Service Provider가 제공하는 Image들에 대해서 Provider Specific한 Ontology를 생성한다

3.3 Communication

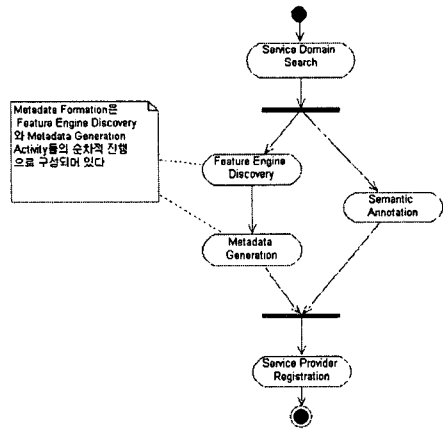
2.3의 내용처럼 워크플로우 기반의 모델을 지원하기 위한 컴포넌트들로 Orchestration과 Choreography Engine [10]을 사용한다.

4. 프로그램 흐름도

프로그램의 전체 흐름은 이미지 제공자를 KVMRF(Image Service Broker)에 등록하는 시나리오와 Client의 Query에 따라 적합한 이미지를 검색하는 두 가지의 시나리오로 나뉘어 진다.

4.1 Registration Scenario

이미지 서버를 KVMRF(Image Service Broker)에 등록하기 위해서 [그림3]과 같은 순서대로 세부 Activity로 나눌 수 있다.



[그림 3] Registration Activity Diagram

4.1.1 Service Domain Search

이미지 제공자에 적합한 Domain 정보를 KVMRF(Image Service Broker)로부터 검색하고 선택함으로써 이미지 제공자가 가지고 있는 Domain정보를 선택한다.

4.1.2 Feature Engine Discovery

Domain에 해당되는 이미지의 Feature Engine들을 검색해서 Feature Engine들의 물리적 주소를 얻는다.

4.1.3 Metadata Generation

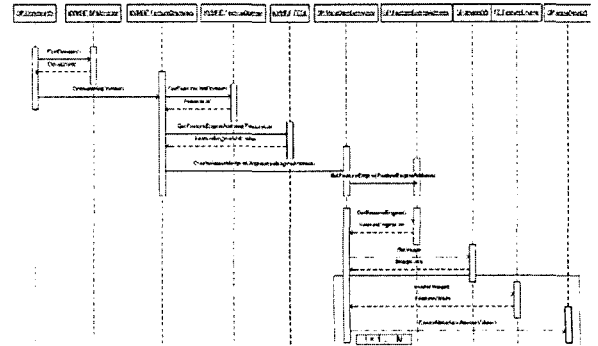
Feature Engine Discovery로부터 얻어온 Feature Engine을 이용해 이미지의 Feature 값(Metadata)들을 생성한다.

4.1.4 Semantic Annotation

KVMRF(Image Service Broker)에 있는 Common Metadata Ontology의 있는 데이터베이스를 이용해 이미지 제공자에 관련된 ontology를 만든다.

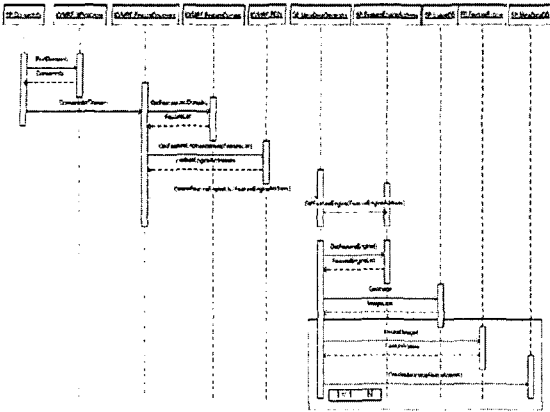
4.1.5 Service Provider Registration

위 절차가 마지면 서비스를 하기 위해서, KVMRF에 등록한다.



[그림 4] Metadata 생성 시나리오

[그림 4]의 시나리오는 4.1.1부터 4.1.3에 시나리오를 나타낸 것이다.



[그림 5] Semantic Annotation과 등록 시나리오

[그림 5]의 시나리오는 4.1.4부터 4.1.5에 시나리오를 나타낸 것이다.

4.2 Image Search Scenario

사용자의 입력 Query별로, 이미지를 검색하는 시나리오는 여러 가지가 있지만, 대표적인 경우 하나를 예로 들겠다.

Query를 기반으로 이미지를 검색하는 시나리오는 4.2.1부터 4.2.3의 순으로 실행되며 전체 Sequence Diagram은 [그림 6]과 같다.

4.2.1 Service Provider Search

Query에 해당되는 가장 적합한 이미지 제공자를 검색하는 방법으로 Semantic Web Service의 MatchMaking[9] 알고리즘을 적용하여 서비스를 검색한다.

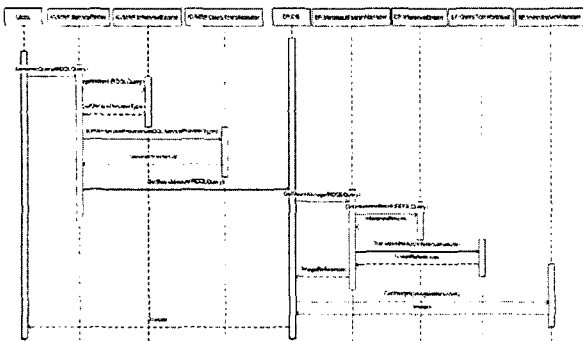
4.2.2 Metadata Search

KVMRF로부터 전송받은 RDQL을 추론, DB에 맞게 형태를 변환함으로써, 적합한 Metadata를 반환한다.

4.2.3 IR2Image

4.2.2의 결과로 받아온 Database의 Index의 Reference에 해당하는 Image를 Client로 반환한다.

이미지를 검색하는 전체 시나리오는 아래와 같다.



[그림 6] 이미지 검색 시나리오

5. 결론 및 향후 과제

본 논문에서는 이미지 제공자의 자율성을 보장하면서, Semantic 기반

의 검색이 가능한 분산 시각미디어를 효율적으로 검색할 수 있는 Framework의 Architecture를 제안하였다. 하지만 더 효율적인 검색을 위하여, 다양한 이미지 검색 질의를 지원하는 Ontology를 기반의 워크플로우 생성에 관한 연구와 이미지 프로세싱에 사용되는 Rule들을 어떻게 Ontology에 적용할 것인지에 대한 세부 Structure들과 Rule에 관한 연구가 필요하다.

참고 문헌

[1]A. Pentland et al. Photobook: tools for content-based manipulation of image databases. In Storage and Retrieval for Image and Video Databases, SPIE, pages 34-47, 1994.

[2]M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. Query by image and video content: The QBIC system. Computer, 28(9):23-32, September 1995.

[3]R. Weber, J. Bolliger, T. Gross, and H.-J. Schek. In Proc. of the ACM Conference on Information and Knowledge Management (CIKM '99), pages 430-441, Nov. 1999.

[4]Eero Hyvonen, Avril Styman and Sampo Saarela. Ontology-Based Image Retrieval. University of Helsinki, Department of Computer Science, Helsinki Institute for Information Technology (HITS)

[5]Laura Hollink, Guus. Schreiber, Jan Wielemaker and Bob. Wielinga. Semantic Annotation of Image Collections. KCAP'03, Florida, October 2003.

[6]Web Services Activity, <http://www.w3c.org/2002/ws/>

[7]Semantic Web, <http://www.w3.org/2001/sw/>

[8] R. Cicchetti, A. Hameurlain, R. Traummüller, Domain Knowledge-Based Automatic Workflow Generation, 13th International Conference, DEXA 2002, Proceeding, pages 81 September 2-6, 2002.

[9] 최원중, 양재영, 최종민, 조현규, 조현성, 김경일, 온톨로지를 계층관계를 이용한 서비스 발견 알고리즘, Proceedings of The 30th KISS Fall Conference, 2003 가을 학술 발표 논문집 1, pages 28-30

[10] Mark Turner et al, "Turing Software into a Service".IEEE Computer Society, Oct, 2003