

온톨로지 기반의 의미 연관성을 이용한 효과적인 방송 정보 검색 방법

*정병희, **김항규, ***박창섭, *오연희, *박성준, **이윤준

*한국방송공사 방송기술연구소, **KAIST 전자전산학과, ***수원대학교 정보공학과

{*{bhjung, yeonche, sparko}@kbs.co.kr, **{hkkim, yjlee}@dbserver.kaist.ac.kr,

***park@suwom.ac.kr

An Effective Broadcast Information Retrieval Method Using Ontology-Based Semantic Relevance

*Byunghee Jung, **Hang-kyu Kim, ***Chang-sup Park, *Yeonhee Oh, *Sungchoon Park, and

**Yoonjoon Lee

*Broadcast Technical Research Institute, KBS, **Dept. of EECS, Korea Advanced Institute of Science and Technology, ***Dept. of Internet Information Engineering, The University of Suwon

요약

방송된 프로그램의 VOD 서비스는 웹 및 다양한 매체를 통해 중요한 정보 서비스가 되고 있다. 기존의 멀티미디어 검색은 파일명이나 메타데이터를 바탕으로 직접 검색하는 정적인 검색 방법이었다. 본 논문에서는 이런 정적인 검색 방법에서 벗어나 의미적으로 연관된 비디오를 동적으로 검색하여 풍부한 검색 결과를 보여주는 방법을 제안하고, 이를 수행하는 프로토타입 시스템을 설계, 구현함으로써 제안한 방법의 타당성을 보인다.

제안하는 방법은 온톨로지를 이용하여 특정 도메인상에서 사용자가 관심을 가지는 정보의 연관성을 표시하고, 이를 검색 과정에 적용한다. 검색 결과는 의미적으로 연관된 정보를 보여줄 뿐 아니라, 연관 카테고리를 바탕으로 관련된 정보를 정리하여 보여줌으로써 사용자는 풍부한 검색 결과 가운데에서 원하는 정보를 쉽고 빠르게 접근할 수 있다. 예를 들면, '고혈압'이라는 검색어를 사용자가 입력하면, 고혈압에 관련된 방송 프로그램 뿐 아니라, '고혈압'에 좋은 음식이나 그 음식의 효능, 효과, 요리법 등 방송을 통해 볼 수 있는 관련 정보들을 한번에 검색할 수 있다.

1. 서론

방송된 프로그램의 VOD 서비스는 웹 및 다양한 매체를 통해 중요한 정보 서비스가 되고 있다. VOD 서비스에 있어서 원하는 비디오를 쉽게 찾을 수 있는 멀티미디어 콘텐츠 검색에 대한 요구는 날이 증가하고 있다. 이러한 요구를 반영하며 사용자 생성 콘텐츠 기반의 동영상 전문 포털인 판도라, 엠군 등이 등장하게 되었고, 네이버, 다음, 네이트 등 유명 포털 사이트들도 멀티미디어 검색에 큰 관심을 가지고 동영상 검색 서비스를 시작하게 되었다.

기존의 멀티미디어 검색은 파일명이나 메타데이터를 바탕으로 직접 검색하는 정적인 검색 방법이었다. 파일명 기반 검색 방법은 검색 엔진이 멀티미디어 콘텐츠의 파일명에 대해 키워드 검색을 하는 것으로, 가장 초보적인 멀티미디어 콘텐츠 검색 방법이다. 멀티미디어 콘텐츠가 가지고 있는 내용을 바탕으로 검색하지 않기 때문에 실제 내용과 파일명의 불일치로 인한 잘못된 검색이 나타날 뿐만 아니라, 파일명으로 콘텐츠의 모든 정보를 담을 수 없기 때문에 풍부한 검색 방법을 제공하기 어렵다.

내용 기반 검색은 멀티미디어 콘텐츠가 가지고 있는 내용을 바탕으로 검색하는 방법으로, 콘텐츠의 내용을 텍스트로 기술한 메타데이

터 기반 검색과 콘텐츠가 가지고 있는 이미지 또는 비디오 기반 검색이 있다. 메타데이터 기반 검색은 파일명을 포함한 멀티미디어 콘텐츠에 대한 정보를 텍스트로 기술하고, 기술되어 있는 메타데이터 정보를 바탕으로 콘텐츠를 검색하는 방법이다. 현재 MPEG을 포함한 많은 동영상 표준들에서 사용하고 있는 검색 방식이다. 그러나 콘텐츠와 별도로 존재하는 메타데이터에 대한 검색이라는 점에서 한계를 가진다. 반면 콘텐츠 내의 이미지나 비디오에 대해 검색을 하는 방법은, 이미지나 비디오의 프레임이 가지고 있는 색상, 이미지의 구도 등에 대한 정보를 바탕으로 콘텐츠를 검색하기 때문에, 검색을 위한 별도의 메타데이터가 필요하지 않다. 그러나 검색 비용이 높으며, 질의 방식이 용이하지 않다는 단점을 가지고 있다.

의미 기반 검색 방법[1-4]은 멀티미디어 콘텐츠가 가지는 의미 정보를 기술하는 방식으로 MPEG-7의 시맨틱 기술 방식이 이에 속한다. 검색에 유용한 의미를 기술해 놓는다는 점에서 유익한 정보를 주는 검색이 가능하나, 기술을 위한 비용이 높다는 단점이 있다.

기존의 검색 방식들은 검색어와 관련된 멀티미디어 콘텐츠를 검색해 주고, 콘텐츠의 검색 경로, 콘텐츠의 크기, 형태(비디오, 오디오, 이미지, 텍스트 등)와 같은 특성 등과 같이 이미 정해져 있는 카테고리에 따라 검색 결과를 보여준다. 즉, 검색된 콘텐츠들 간의 의미적 관계

를 나타내 주지 못하고 있다. 본 연구에서는 '연관 검색'이라 정의한 검색된 멀티미디어 콘텐츠들 사이의 의미적 관계에 따라 결과를 보여주는 의미 연관성을 이용한 방송 정보 검색 방법과 이를 구현한 시스템을 제안하고자 한다. 제안한 방법은 온톨로지를 이용하여 의미 정보를 기술하고, 이에 근간한 유의한 정보들을 동적으로 보여주기에 때문에 풍부하고 다양한 검색이 가능하다.

다음의 2 장에서는 본 연구에서 제안하고 있는 의미 연관성을 이용한 방송 정보 검색에 대해 간단한 예를 통해 설명을 하고자 한다. 그리고 설계된 검색 시스템 구조를 3 장에서 제시한다. 설계한 시스템 구조를 바탕으로 구현한 내용을 4 장에서 보여주도록 하고, 마지막 5 장에서 본 연구에 대한 결론을 도출한다.

2. 의미 연관성을 이용한 방송 정보 검색

멀티미디어 콘텐츠 검색 포털 사이트의 검색어로 '사과'를 입력해 보자. '사과'와 관련된 콘텐츠들이 최근 자주 접근된 정도에 따라 나열되어 있다. 각 동영상에서의 '사과'가 가지는 의미를 검색 결과에서 전혀 반영을 하지 못하고 있다는 것을 알 수 있다. 만일 '사과가 효능을 나타낼 수 있는 질병'에 대해 사용자가 알고 싶다면, '사과'를 입력한 다음 모든 콘텐츠를 직접 살펴보든지, 아니면 '사과 질병', '사과 효능' 등과 같은 검색어 집합을 사용자가 직접 조합해 보면서 찾아보아야 하는 번거로움이 발생한다.

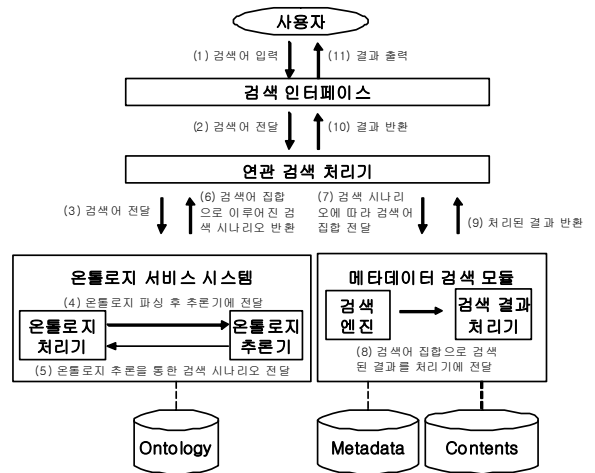
본 연구에서 정의한 연관 검색은, 콘텐츠들에서 나타나는 '사과'의 의미에 따라 콘텐츠들을 분리함으로써 멀티미디어 콘텐츠들 사이의 연관 정보도 함께 제공해 주는 검색 방식을 뜻한다. 검색어로 '사과'가 입력이 된다면, 과일 의미로서의 '사과'가 소개되는 멀티미디어 콘텐츠들이 한 카테고리 안에 나열이 되고, '사과'의 구성 물질 등을 알려주는 콘텐츠들이 다른 카테고리에 나열되고, 어떤 질병에 이로운지에 대한 정보를 제공하는 콘텐츠들이 또 다른 카테고리에서 결과로 출력된다. 만일 질병에 대한 효능이 '사과'와 유사한 과일로 '귤'이 나타난다면, 이와 관련된 동영상 콘텐츠도 다른 카테고리로 출력될 수 있다. <그림 1>은 본 연구를 통해 구현된 연관 검색 결과 화면이다. 연관 검색을 통해 검색어 위주의 정보 검색 뿐 아니라, 검색어와 연관된 의미 있는 정보들까지도 함께 검색된다.



<그림 1> 연관 검색 예

이러한 연관 검색을 위해서는 특정 도메인에서 검색어로 쓰일 만한 키워드들을 위주로 이들의 의미 및 관계를 기술할 필요가 있다. MPEG-7 표준에서는 실 세계를 기술하기 위한 시맨틱 기술자를 정의하고 있다. MPEG-7의 메타데이터에서 기술되어 있는 객체들간의 관계를 시맨틱 기술자에서 이미 정의되어 있는 관계들을 이용해서 연관 정보를 기술할 수 있게 되어 있다. 이미 정의되어 있는 기술자와 관계들을 바탕으로 의미 관계 정보를 기술할 수 있기 때문에 표현력에 한계를 가질 수 밖에 없다. 본 연구에서는 표현력이 풍부할 뿐만 아니라, 메타데이터 형식과 멀티미디어 콘텐츠와 독립적인, 온톨로지를 통해 실 세계 정보를 기술하고자 한다. 온톨로지는 유사 객체들을 묶어서 분류하고, 다른 각 객체들 사이의 관계를 기술하는 능력을 가지고 있다. 이러한 온톨로지를 활용해서 연관 검색을 할 수 있게 된다.

3. 검색 시스템 구조



<그림 2> 제안하는 검색 시스템 구조

본 절에서는 온톨로지를 이용한 연관 검색 시스템의 구조를 제안하고자 한다. <그림 2>는 온톨로지를 기반으로 하는 연관 검색 시스템 구조를 보여주고 있다. 연관 검색 시스템은 멀티미디어에 대한 메타데이터를 검색하는 내용 기반 검색에 온톨로지 기술된 시맨틱 정보를 이용해서 검색 결과를 멀티미디어 콘텐츠가 가지는 의미들간의 관계에 따라 분류해서 보여준다. 이러한 과정은 <그림 2>와 같이 11 단계의 처리 과정을 거쳐 이루어진다. (1) 사용자가 검색 인터페이스를 통해 검색어를 입력한다. (2) 검색 인터페이스를 통해 입력된 검색어를 연관 검색 처리기에 전달한다. (3) 연관 검색 처리기는 입력 받은 검색어를 온톨로지 서비스 시스템에 전달한다. (4) 온톨로지 처리기가 해당 온톨로지를 파싱 처리해서 온톨로지 추론기에게 전달한다. (5) 온톨로지 추론기는 온톨로지 처리기로부터 받은 온톨로지를 바탕으로 연관 검색 처리기로부터 받은 검색어를 연관 정보를 가진 검색어들의 집합으로 변경해서 온톨로지 처리기에게 전달한다. (6) 온톨로지 처리기는 온톨로지 추론기로부터 받은 결과를 바탕으로 검색 시나리오를 구성해서 연관 검색 처리기에게 반환한다. (7) 연관 검색 처리기는 검색 시나리오에 따라 검색어 집합을 메타데이터 검색 모듈에게 전달한다. (8) 검색 엔진은 검색어 집합으로 검색된 메타데이터 파일 정보를 검색 결과 처리기에게 전달한다. (9) 검색 결과 처리기는 검색 엔진으로부터 받은 메타데이터 파일 정보를 바탕으로 연관 검색 처리기가 필요로 하

는 검색 결과를 추출해서 반환한다. (10) 연관 검색 처리기는 전달 받은 결과를 검색 인터페이스로 반환한다. (11) 사용자는 검색 인터페이스를 통해 결과를 출력 받게 된다. 이와 같이 동작하는 연관 검색 시스템의 각 모듈들을 살펴보도록 하겠다.

가. 검색 인터페이스

검색 인터페이스는 사용자가 연관 검색 시스템을 접근하기 위한 인터페이스이다. 웹 포털과 같이 웹 페이지가 될 수도 있으며, 응용 프로그램 클라이언트로 구현될 수도 있다. 검색 인터페이스는 검색어를 입력 받을 수 있는 입력 부분과 검색의 결과를 출력할 수 있는 출력 부분으로 구성된다. 연관 검색 시스템은 결과를 온톨로지에 따라 동적으로 반환하기 때문에 검색 인터페이스는 가능하면 연관 검색 처리기의 하위 레벨과는 독립적으로 구성될 수 있어야 한다.

나. 연관 검색 처리기

연관 검색 처리기는 온톨로지 서비스 시스템과 메타데이터 검색 모듈 사이를 연결해주면서 상위의 검색 인터페이스와 검색어 입력, 결과 출력을 하는 중앙 처리기 역할을 한다. 연관 검색 시스템의 각 모듈들 간의 통신은 연관 검색 처리기를 통해 이루어진다. 검색 인터페이스는 연관 검색 처리기를 통해 온톨로지 서비스 시스템, 메타데이터 검색 모듈을 이용하여 연관 검색 결과를 받는다. 온톨로지 서비스 시스템은 연관 검색 처리기를 통해 검색어 시나리오를 메타데이터 검색 모듈에게 전달하고, 메타데이터 검색 모듈은 검색 결과 처리기가 반환하는 결과를 연관 검색 처리기를 통해 검색 인터페이스에게 전달하게 된다.

다. 온톨로지 서비스 시스템

온톨로지 서비스 시스템은 사용자가 검색 인터페이스를 통해 입력한 검색어를 연관 정보를 바탕으로 연관 검색어 집합으로 변환해 준다. 연관 검색어 집합은 온톨로지서 기술되어 있는 정보를 바탕으로 연관된 검색어들로 이루어진 집합을 이르며, 여러 검색어 집합은 다시 검색어 시나리오를 이룬다. 온톨로지는 파일 시스템 상에 저장되어 있는 파일일 수도 있으며, 데이터 베이스 안에 존재하는 데이터 테이블, 또는 웹 상의 URL에 위치하고 있는 문서일 수도 있다. 온톨로지 처리기는 이러한 형태로 존재하고 있는 온톨로지를 파싱해서 메모리 상에 적재한 다음에 온톨로지 추론기에게 전달해서 메모리 상에 지식 베이스를 구성할 수 있도록 한다. 온톨로지 추론기는 메모리 상에 구성된 지식 베이스를 통해 검색어를 검색어 집합으로 변환할 수 있도록 연관 정보를 제공해 준다.

라. 메타데이터 검색 모듈

메타데이터 검색 모듈은 멀티미디어 콘텐츠를 설명하고 있는 메타데이터 파일들을 검색하는 모듈이다. 메타데이터는 콘텐츠에 대한 내용 기반 설명뿐만 아니라 콘텐츠를 물리적으로 접근할 수 있는 정보도 함께 가지고 있다. 따라서 해당 메타데이터를 검색하면 이에 대응되는 멀티미디어 콘텐츠도 바로 검색이 가능하다. 연관 검색 처리기로부터 검색 시나리오에 따라 검색어 집합을 전달받으면, 검색 엔진은 검색어 집합을 가지고 있는 메타데이터 파일을 검색하게 된다. 검색된 메타데이터 파일은 사용자가 원치 않는 데이터들도 포함하고 있을 수 있기

때문에 별도의 결과 처리가 필요한데, 이를 위해 검색 결과 처리기가 있어야 한다. 검색 결과 처리기는 사용자의 출력 화면에 전달되어야 하는 결과만을 추출해서 연관 검색 처리기에게 전달하도록 한다.

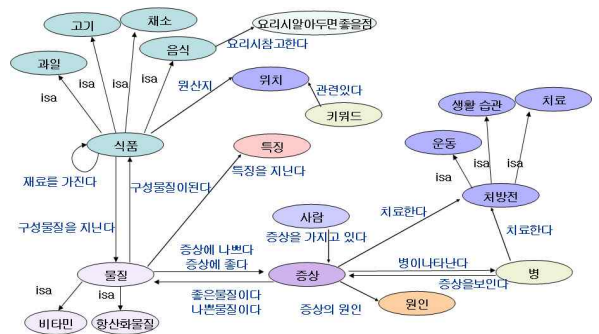
4. 프로토타입 구현

본 연구에서 제안하고 있는 연관 검색 기법과 설계한 연관 검색 시스템 구조가 활용도를 가질 수 있는지 확인하기 위해 시스템 프로토타입을 구현하였다. 본 장에서는 프로토타입 시스템 환경뿐만 아니라 온톨로지나 메타데이터, 멀티미디어 콘텐츠등과 같은 데이터 환경도 포함된 데모 환경을 설명하고, 구현된 시스템 모듈, 그리고 데모 시스템이 동작하는 모습을 보여주도록 하겠다.

가. 시스템 환경

프로토타입 시스템이 구현된 환경은, Dual Xeon 2.8c, 3 GB RAM, Red Hat Linux 3.3.3-7이다. 사용 프로그래밍 언어는 자바로, jdk 1.5.0_09를 사용했다. 사용자 인터페이스는 웹 인터페이스로 구현했는데, 이를 위해 웹 서버로는 apache tomcat 5.5.20을 사용했다.

프로토타입 시스템 데모를 위해 건강과 관련된 내용에 특화된 온톨로지를 설계했다. 온톨로지의 모습은 아래의 그림과 같다. 건강 온톨로지는 식품, 물질, 증상, 병과 같이 크게 4 개의 클래스로 구성이 된다. 식품 클래스는 고기, 과일, 채소, 음식 등과 같은 하위 클래스를 가지며, 식품 객체의 구성 물질을 나타낼 수 있도록 물질 클래스와 구성 물질 관계를 가진다. 증상 클래스는 어떠한 질병에 대해 나타나는 증상을 나타내는 클래스로, 각 증상 객체는 병 클래스와 증상 관계를 가지고, 물질 클래스와 증상에 대한 이해 관계를 가진다. 그리고 증상에 대한 치료 방법을 기술할 수 있도록 처방전 관련 부분도 건강 온톨로지에 포함되어 있다. 본 건강 온톨로지는 OWL(Web Ontology Language)[5]로 기술해 데모 시스템에서 이용하였다.



<그림 3> 건강 온톨로지

멀티미디어 콘텐츠는 한국 방송 공사에서 제작한 '비타민', '스폰지', '생로병사의 비밀', '무엇이든 물어보세요' 등과 같은 건강 관련 방송 콘텐츠를 이용했다. 콘텐츠에 대한 메타데이터는 디지털 방송 콘텐츠 표준 중의 하나인 TV-Anytime 메타데이터 스키마[6]를 바탕으로 생성했다. TV-Anytime 표준은 메타데이터 스키마에 대한 내용뿐만 아니라, 다양한 디지털 방송 서비스 플랫폼에 대한 내용까지 포괄하고 있다.

나. 시스템 모듈 구현

사용자의 검색 질의어를 받기 위한 검색 인터페이스는 웹 기반으로 구현되었다. 웹서버로 아파치 톱캣을 이용했고, 검색 인터페이스는 하부의 연관 검색 처리기를 이용하는 JSP 페이지가 된다.

연관 검색 처리기는 자바로 구현된 라이브러리 형태로 제공이 되며, 검색 인터페이스의 JSP 페이지에서 연관 검색 처리기 라이브러리를 이용하게 된다. 연관 검색 처리기에서 동작하는 과정은 세 단계로 나누어진다. 첫 단계는 검색 인터페이스를 통해 전달 받은 검색어를 온톨로지 서비스 시스템에게 전달하고 결과로 카테고리들의 목록과 각 카테고리를 내의 세부 내용들을 반환 받는 과정이다. 그 다음으로는 온톨로지 서비스 시스템에서 받은 결과를 바탕으로 메타데이터를 검색한 후, 검색 인터페이스에서 출력하고자 한 내용을 추출하는 과정이다. 마지막으로 검색 인터페이스에서 설계된 페이지 형식에 맞추어 추출된 내용을 넣음으로써 새로운 페이지를 생성, 출력한다.

온톨로지 서비스 시스템은 기능을 바탕으로 크게 두 부분으로 나누어 볼 수 있다. 온톨로지에 대한 추론 기능을 가지고 있는 온톨로지 추론기와 그 외의 온톨로지 파싱 또는 질의 및 결과 처리 기능을 담당하는 온톨로지 처리기로 나누어 볼 수 있다. 본 데모 시스템에서는 추론기뿐만 아니라 처리기까지도 포괄하고 있는 Jena를 이용했다. 온톨로지 추론기는 건강 온톨로지에 대한 OWL 파일을 파일 시스템으로부터 읽어 들인 후 추론된 결과에 대한 지식베이스를 메모리 상에 구성해 놓는다. 이러한 로딩 과정은 시간 비용도 클 뿐만 아니라, 웹 환경에서 이용하고 있는 본 데모 시스템에서는 각 사용자 요구마다 하나씩의 지식베이스를 생성하면 메모리 공간 비용도 심각하게 커진다. 이를 해결하기 위해 온톨로지 서비스 시스템을 데몬 형태로 별도의 프로세스로 동작 시키고, 데몬을 실행 시킬 때 한번의 온톨로지 로딩 과정만을 거침으로써 시간 비용을 절감하고, 메모리 상 하나의 지식베이스를 공유함으로써 메모리 공간 비용도 절약할 수 있었다. 온톨로지 서비스 시스템에 요청을 하는 연관 검색 처리기는 온톨로지 서비스 시스템의 클라이언트 부분에 대한 라이브러리를 통해 온톨로지 서비스 시스템의 데몬에 질의하게 된다.

본 시스템에서는 TV-Anytime 표준에서 정의하고 있는 스키마를 바탕으로 메타데이터를 사용하고 있다. 이러한 메타데이터는 XML 형식으로 파일 시스템에 저장되고, 이를 검색하기 위한 모듈이 메타데이터 검색 모듈이다. 메타데이터 검색 모듈은 Apache에서 제공하고 있는 루씬(Lucene)을 검색 엔진으로 이용하고 있다. 루씬은 문서들에 존재하는 단어들에 대한 인덱스를 생성하고, 해당 검색어를 포함하고 있는 문서들을 검색해 준다. 본 데모 시스템에서는 루씬의 검색 엔진에 메타데이터 XML 파일들에 대한 처리와 한글의 조사 처리 부분을 추가 구현하여 사용했다.

메타데이터 검색 모듈에는 검색 엔진 외에 검색 결과 처리기라는 구성 요소가 있다. 검색 엔진은 메타데이터를 파일 단위로 검색해 주기 때문에 연관 검색 처리기에서 실제로 필요로 하는 내용들만을 XML 형식으로 전달해 주기 위한 구성 요소가 필요하다. 또한 검색 엔진은 단순히 검색어가 문서에 포함되어 있는지만을 확인할 뿐, 어떠한 의미로 포함되어 있는지는 확인해 주지 못한다. 이러한 부분 또한 검색 결과 처리기에서 담당하고 있다. 검색 엔진은 검색어 집합을 포함한 메타데이터 파일들을 검색하고, 검색 결과 처리기에서 이 중 xpath로 지정된 값들만으로 구성된 XML 메시지로 연관 검색 처리기에게 반환하게 된다. 'abcd'이라는 식별자를 가진 방송 프로그램을 검색하는 예를 들

어보자. 검색 엔진에서는 'abcd'를 파일 상에서 포함하고 있는 모든 메타데이터 파일들을 검색한다. 검색 결과 처리기는 'abcd'가 참조가 아닌 식별자로 사용된 문서인지 확인하고 연관 검색 처리기로부터 받은 xpath를 결과로 출력하는 XQuery를 생성하고, 검색 엔진으로부터 검색된 파일들에 대해 생성된 XQuery를 질의한다. 그리고 생성된 결과는 XML 메시지 형태로 연관 검색 처리기에게 반환된다.

4. 결론

웹 기술이 발달하고 네트워크의 성능이 향상됨에 따라 단순한 문서 또는 이미지에 대한 검색뿐만 아니라, 멀티미디어 검색에 대한 요구가 크게 증가하고 있다. 그러나 이러한 요구에 비해 멀티미디어 검색 분야에서는 파일명이나 파일의 용량, 또는 인코딩 방식 등과 같은 동영상 파일에 대한 정보나 동영상에 대한 메타데이터를 바탕으로 직접 검색하는 정적인 검색 방식이 사용되고 있다. 이러한 검색 방식들은 멀티미디어의 내용에 대한 검색이 어려울 뿐만 아니라, 각 콘텐츠들이 가지고 있는 내용들을 바탕으로 콘텐츠들 간의 연관된 지식을 찾을 수 없게 된다. 본 논문에서는 이러한 기존의 검색 방식에서 벗어나고자, 온톨로지를 바탕으로 동영상의 메타데이터를 검색하는 시스템 설계를 제안하고, 데모 프로토타입 시스템을 구현함으로써 제안한 시스템의 타당성을 보여주었다. 온톨로지서 기술해 주고 있는 연관 정보를 바탕으로 멀티미디어를 검색해 주는 본 논문의 검색 방법은 검색어에 연관된 콘텐츠를 검색해 줄뿐만 아니라, 검색된 콘텐츠들 간의 관계를 온톨로지의 연관정보를 통해 보여주고 있다. 그리고, 검색어가 포함되지 않지만, 연관 정보를 가지는 콘텐츠 또한 결과로 제공한다.

현재 구현한 데모 시스템은 프로토타입 수준으로 아직 개선할 부분이 많다. 우선 문제를 간단히 하기 위해 검색어를 한 개의 단어로 한정하고 있는데, 두 개 이상의 검색어에 대한 검색 방법도 고려해 보아야 할 것이다. 멀티미디어에 대한 메타데이터 또한 대용량을 생성하지 못한 상태이다. 이 부분 또한 메타데이터를 추가적으로 생성하고 온톨로지도 확장함에 따라 발생할 수 있는 문제점들을 생각해 볼 필요가 있다.

5. 참고 문헌

- [1] D. Vallet, M. Fernández, and P. Castells, "An Ontology-Based Information Retrieval Model," In Proc. of 2nd European Semantic Web Conference, 2005.
- [2] M. Bertini, A.D. Bimbo, and C. Tornial, "Enhanced Ontologies for Video Annotation and Retrieval," In Proc. of MIR, 2005.
- [3] L. Hollink, M. Worring, and A.Th. Schreiber, "Building a Visual Ontology for Video Retrieval," In Proc. of MM, 2005.
- [4] A.Kiryakov, B.Popov, D.Ognyanoff, D.Manov, A.Kirilov and M.Goranov, "Semantic Annotation, Indexing, and Retrieval," In Proc. of Int'l Semantic Web Conference, 2003.
- [5] World Wide Web Consortium, "Web ontology language (OWL)," W3CSemantic Web Activity, <http://www.w3.org/2004/OWL/>, 2004.
- [6] European Telecommunications Standards Institute, "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems("TV-Anytime") Part 3: Metadata," ETSI TS 102 822-3-1, ETSI, 2006.