

비디오 콘텐츠를 위한 색인 온톨로지 저장소

황우연[†], 양정진^{**}

요 약

나날이 방대해지는 디지털 콘텐츠들의 홍수 속에서 원하는 정보를 찾아내는 데 필요한 정밀한 색인 기술의 필요성은 날로 증대되어 가고 있다. 이러한 요구 수용을 위해서는 지능형 개체가 정보의 검색 주체가 되어야 하며, 지능형 개체간의 상호 운용성이 뒷받침되어야 한다. 본 논문에서는 Snoek, Worring들이 제안한 다중 형상 색인을 위한 통합 기반구조를 분석하고 콘텐츠 기반의 자동화된 색인 기법에서 색인 정보의 신뢰성을 높이기 위한 방법을 살펴본다. 그리고 이를 통해 시맨틱 웹 기술의 온톨로지 개념을 적용하여 추상화된 색인 정보의 생성과 제어를 지원하고 사람과 기계간, 기계와 기계간의 의미적 상호 운용성이 보장되는 기반 모델을 제시하고자 한다. 온톨로지의 메모리 상주 모델 처리 방식은 방대한 색인 정보를 수용하기에 부적절하기 때문에 온톨로지 저장소의 사용을 필요로 하며, 논리적으로 표현된 지식의 일관성과 추론을 위해 추론엔진의 사용이 요구된다. 본 논문에서는 온톨로지 저장소의 기능 및 성능적인 요구사항을 모두 만족시키는 Minerva 온톨로지 저장소를 이용하여 비디오 콘텐츠를 표현한 설계 지식을 저장하고 회수하는 실험을 하였다. 또한 마지막으로 다른 연구과제들과 연계한 효율적인 색인 가능성에 대해 고찰해 본다.

Index Ontology Repository for Video Contents

WooYeon Hwang[†], Jung-Jin Yang^{**}

ABSTRACT

With the abundance of digital contents, the necessity of precise indexing technology is consistently required. To meet these requirements, the intelligent software entity needs to be the subject of information retrieval and the interoperability among intelligent entities including human must be supported. In this paper, we analyze the unifying framework for multi-modality indexing that Snoek and Worring proposed. Our work investigates the method of improving the authenticity of indexing information in contents-based automated indexing techniques. It supports the creation and control of abstracted high-level indexing information through ontological concepts of Semantic Web skills. Moreover, it attempts to present the fundamental model that allows interoperability between human and machine and between machine and machine. The memory-residence model of processing ontology is inappropriate in order to take-in an enormous amount of indexing information. The use of ontology repository and inference engine is required for consistent retrieval and reasoning of logically expressed knowledge. Our work presents an experiment for storing and retrieving the designed knowledge by using the Minerva ontology repository, which demonstrates satisfied techniques and efficient requirements. At last, the efficient indexing possibility with related research is also considered.

Key words: Video Contents(비디오 콘텐츠), Indexing(색인), Ontology Repository(온톨로지 저장소)

※ 교신저자(Corresponding Author) : 양정진, 주소 : 경기도 부천시 원미구 역곡2동 산43-1(420-102), 전화 : 02)2164-4377, FAX : 02)2164-4777, E-mail : jungjin@catholic.ac.kr
접수일 : 2009년 1월 2일, 수정일 : 2009년 4월 2일
완료일 : 2009년 7월 14일

[†] 준회원, 가톨릭대학교 컴퓨터전자공학부
(E-mail : hwangwooyeon@hanmail.net)

^{**} 정회원, 가톨릭대학교 컴퓨터전자공학부

※ 본 연구는 2007년도 가톨릭대학교 교비연구비지원으로 수행되었음

1. 서 론

색인(Index)은 전자 도서관의 검색 기술에서부터 사용자의 특징에 따라 상황에 적합한 맞춤형 콘텐츠 제공기술의 핵심적인 역할을 수행한다. 정보의 검색 주체가 사람인 전자 도서관의 검색 기술보다 정보의 검색 주체가 지능형 개체인 맞춤형 콘텐츠 제공 기술은 풍부하고 완전한 색인 정보 제공을 요구하고 지능형 개체간의 의미적 상호 운용성을 바탕으로 협력적으로 제공된다.

기존에는 주로 사람에 의하여 색인 정보가 제공되었지만 근래에는 콘텐츠 기반의 자동화된 색인 기법에 관한 연구가 활발히 진행되고 있다[1]. 색인 정보의 신뢰성을 높이기 위해서는 사람과 지능형 개체간의 상호 합의와 상호 협력을 요구한다. 예를 들어, 사람은 제공된 콘텐츠의 평점을 색인 정보에 포함할 수 있고 지능형 개체는 콘텐츠가 제공된 시간과 지역을 색인 정보에 포함 할 수 있다. 혹은 지능형 개체가 생성한 색인 정보를 사람이 이해할 수 있어야 한다.

본 연구에서는 사람과 지능형 개체간 혹은 지능형 개체들 간의 상호 합의를 위하여 시맨틱 웹의 기반 기술인 온톨로지를 기반으로 색인 정보를 위한 메타데이터를 생성한다. 방대한 양의 온톨로지 기반의 색인 정보를 저장, 회수하기 위하여 온톨로지 저장소를 도입하고 시맨틱 웹, 웹 서비스, 웹 2.0에서 의미 정보 회수를 위한 프로토콜을 이용하여 색인 정보를 회수하기 위한 기반 구조를 제시한다.

2. 관련 연구

2.1 다중형상 비디오 색인(Multimodal Video Indexing)을 위한 프레임워크

색인 기법에는 '무엇', '어떻게', '어떤 것'을 지칭할 것인가의 세 가지 이슈가 있다. 색인 대상의 범위를 영상매체의 전 지역으로 할 것인지, 혹은 프레임마다 할 것인지 결정하는 것이 '무엇'에 해당한다. 색인 기준으로 영상매체의 어떤 특징을 선별하고 분석할 것인지 결정하는 것이 '어떻게'에 해당하며, 색인하고자 하는 정보에 해당하는 것을 선별하는 것이 '어떤 것'에 해당한다.

Snoek, Worring[2]들은 영상 매체의 색인 정보의 '어떻게'를 해결하기 위한 기법들이 영상, 오디오, 텍

스처의 단일 형태로 대상연구 되고 있음을 지적하고 '무엇', '어떻게', '어떤 것'에 해당하는 세 가지 이슈를 모두 수용하는 자동화된 다중형상 비디오 색인을 위한 기반구조를 제안하였다.

다중형상(Multimodality)은 시스템의 도메인으로 부터 관찰되며 Nigay, Coutaz들에 의하여 "다른 형식의 전달 매체를 통하여 자동적으로 의미를 전달하고 추출하기 위해 유저와 상호작용하는 시스템 포용력"으로 정의되었다[3]. Snoek, Worring[2]들은 영상 매체 색인 기법의 '어떻게'와 '무엇'을 결정하기 위해 이 정의를 저작자의 관점을 반영하여, "특정 내용(Content)과 함께 구획(Layout)을 조합하여 미리 정의된 의미적 아이디어를 표현하기 위해 다음에 나열된 두 개 이상의 정보 채널을 이용한 영상 매체의 저작자의 포용력"으로 확장하였다.

- 시각적 형상: mise-en-scene을 포함한, 영상매체를 통해 볼 수 있는 자연적이거나 인공적으로 만들어진 모든 것
- 청각적 형상: 화법, 음악, 환경적인 음향을 포함한 영상매체에서 들리는 모든 것
- 문자 형상: 영상 매체의 내용을 묘사하는 문자 자원들

내용의 관점은 색인 기법의 '어떤 것'과 관련이 있으며 구획과 저작자가 영상매체를 생성하기 위해 사용한 요소에 관련이 있다. 배경(Setting)은 줄거리가 가지는 장소상의 시간과 공간이며 분위기 혹은 심리역시 강조 할 수 있다. 사물(Object)은 영상 매체상의 두드러진 정적, 혹은 동적인 개체들이며 사람(People)은 영상 매체 상에 출현하는 사람이다.

내용의 요소들은 특정 형상 양식의 요소들을 이용한다. 시각적인 형상은 저작자가 특정 색, 빛, 카메라 각도, 카메라 거리, 카메라의 이동을 이용할 수 있으며 청각적인 양식 요소들은 소란스러움, 리듬, 음악의 특성이 있고 문자 형상 출현은 필적과 문체에 의하여 결정된다.

구획은 색인 기법의 '무엇'과 관련이 있으며 각 형상의 문법 구조이며 다음 요소들로 구성된다. 기본 단위(Fundamental Units)들은 다른 형상을 구별하는 주요 요인이 되고, 시각적 형상의 기본 단위는 하나의 영상 프레임들이며, 청각적인 형상의 기본 단위는 오디오 샘플들이다. 센서 샷(Sensor shots)들은

기본 단위들의 집단이며 다음 샷들로 구분된다. 카메라 샷(Camera shots)은 카메라의 연속된 기록의 결과이고, 마이크로폰 샷(Microphone shots)은 마이크의 연속된 기록의 결과이며, 텍스트 샷(Text shots)은 연속된 텍스트 표현이다. 변환 편집(Transition edits)은 다른 센서 샷들을 통일성 있는 문서 구조로 연결하는 것이다. 시각적인 변환은 와이프기법, 페이드아웃, 오버랩 등이 있으며, 청각적인 변환은 음악을 이용한 부드러운 변환, 정적을 이용한 갑작스런 변화 등이 있고, 문자 변환은 자막이나 색변환 등이 있다. 특별한 효과(Special effects)는 형상의 영향을 강하게 하거나 의미를 추가 할 때 사용된다. 문자 효과를 예로 들면 생산 시간, 크기나 공간적인 위치로 표현될 수 있다.

Snoek, Worring[2]들은 색인 정보의 '어떤 것'과 '무엇'을 결정하기 위하여 다음에 나열된 색인 정보의 종류를 제시하였고 다중 형상을 생성하기 위하여 형상을 추출하는 기법들을 함께 소개하였다.

- 목적(Purpose): 유사한 의도를 공유하는 영상매체의 집합. (예: 오락, 정보, 의사소통)
- 장르(Genre): 유사한 양식을 공유하는 영상매체의 집합. (예: 장편영화, 뉴스방송, 광고)
- 하부장르(Sub-genre): 유사한 내용을 공유하는 영상매체 장르의 부분 집합.
- 논리적 단위들(Logical units): '지정된 사건들', 다른 '논리적 단위들'이 합쳐져 의미를 가지는 집합으로 구성된 영상매체의 연속적인 부분의 내용. (예: 드라마 내에서의 대화, 축구경기의 전반전, 뉴스 방송에서의 일기예보)
- 지정된 사건들(Named events): 일정 시간 동안 변하지 않는 의미가 배정된 짧은 구획.

Snoek, Worring[2]들이 제안한 다중 형상 색인을 위한 통합 기반구조는 그림 1에 표현되었으며, 본문에서는 시맨틱 색인 부분을 온톨로지화 및 기술하여 의미적 상호 운용성을 보장하고자 한다.

2.2 온톨로지 저장소

W3C에서는 데이터를 위한 의미를 기술하기 위한 표준으로 XML(eXtensible Markup Language)을 권고한다[4]. XML은 기계가 읽을 수 있는 구조로 정의되었으며, XML 스키마와 더불어 성분과 속성으로

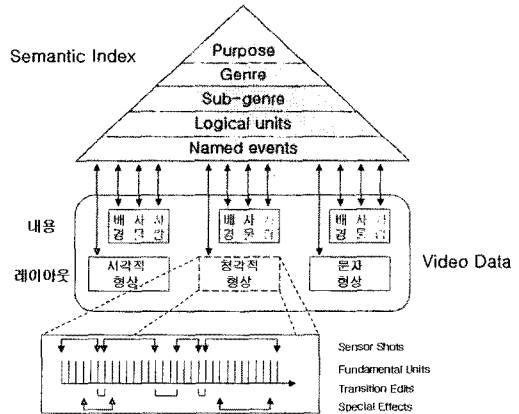


그림 1. 다중 형상 색인을 위한 통합 프레임워크

메타데이터 어휘를 기술하고 개체 수 제약을 이용하여 구문을 기술한다. RDF(Resource Description Framework)는 XML을 기반으로 유, 무형의 자원을 주어·술어·목적어 기반의 트리플 형태로 선언하며 URI(Uniform Resource Identifier)와 더불어 유일한 메타데이터를 기술하기 위한 기반구조로 사용된다 [4]. RDF 스키마는 RDF의 용어와 역할의 개념을 정의하는 경량의 온톨로지를 정의하기 위한 메타언어로 사용되며, 응용계층은 각 개념들이 선언된 온톨로지를 명시하여 의미적 상호 운용성을 보장받는다. OWL(Web Ontology Language)은 W3C에서 권고하는 온톨로지 언어이며 서술논리(Description logic)를 기술하여 논리적인 표현력과 합리적인 시스템을 위한 자료구조로 사용될 수 있다[4-6].

그림 2는 W3C에서 제시하는 시맨틱 웹 스택이다. 최하층에서는 유일한 자원 식별자 (URD)와 공통의 언어 셋을 이용하여 범용 어플리케이션에서의 데이터 표현의 상호 운용성을 보장한다. 이를 기반으로 확장 가능한 마크업 언어(XML)와 네임스페이스를

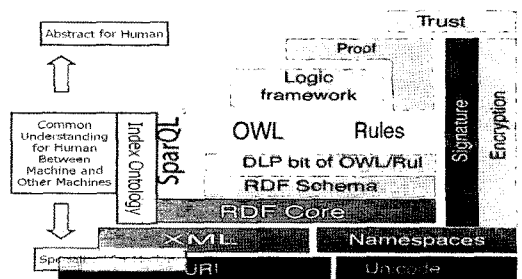


그림 2. 시맨틱 웹 스택의 시맨틱 색인 온톨로지 경계

이용하여 메타데이터의 구문과 어휘 구조를 생성하며, 메타언어의 특징으로 인해서 기계가 일관된 방법으로 XML기반의 새로운 언어를 읽고 쓸 수 있게 되었다.

본 연구에서는 Snoek, Worring[2]들이 제안한 기반 구조를 바탕으로 그림 2의 RDF Core와 OWL의 표현 범위를 가지는 색인 온톨로지를 기술한다. 온톨로지의 메모리 상주 모델 처리 방식은 방대한 색인 정보를 수용하기에 부적절하기 때문에 온톨로지 저장소의 사용이 필요하며, 논리적으로 표현된 지식의 일관성 유지와 추론을 위해 추론엔진의 사용이 요구된다. 온톨로지 저장소의 기능적 요구사항과 성능적인 요구사항을 충족시키기 위해 KAON, DLDB등이 연구되었으며, 본고에서는 두 요구사항을 모두 수용하는 Minerva 온톨로지 저장소를 이용하여 표현된 설계지식을 저장하고 회수한다[7,8].

2.3 비디오 색인과 추출을 위한 온톨로지 구축

현존하는 비디오 이벤트에 대한 추출방식은 크게 두 가지 방식으로 분류된다. 첫째는 비디오의 전체적인 콘셉트를 분석해 비디오 이벤트를 추출하는 형식이다. 도메인에 상관없이 서로 다른 종류의 비디오를 분류하며 하나의 비디오를 TV프로그램의 광고, 뉴스, 날씨 비디오 등으로 분류할 수 있다. 또한 비디오의 물리적 계층인 프레임, 그룹 별로 콘셉트를 추출해 이벤트를 분류한다[9].

둘째는 특정 도메인에서 객체의 모션 및 각종 저수준 시각특징들을 이용해 프레임 단위에서 세밀한 이벤트를 추출하는 방식이다. 이는 감시카메라 및 스포츠 동영상에서 객체의 모션을 분석해 특정 이벤트를 얻는 방식이다. 예로 교통신호 위반 상황이 담긴 이벤트 검색 및 축구 동영상의 골 장면 등이 있다[9].

내용기반 연구는 검색의 정확도를 높이고, 속도를 빠르게 하기 위하여 사전에 색상과 같은 저수준의 정보를 자동적으로 추출하여 메타데이터로 생성해 놓는 연구방법이다[10]. 추출된 메타데이터 정보는 멀티미디어 데이터를 관리하고 검색하는데 이용된다. 하지만 초기의 내용기반 연구들은 메타데이터의 정보표현 및 구조에 대한 표준이 되어있지 않아 일반화된 검색을 지원하는데 어려움이 있었다. 이러한 표준화에 대한 요구와 멀티미디어 검색분야에 관한 지속적인 관심으로 MPEG-7표준이 제정되면서 검색

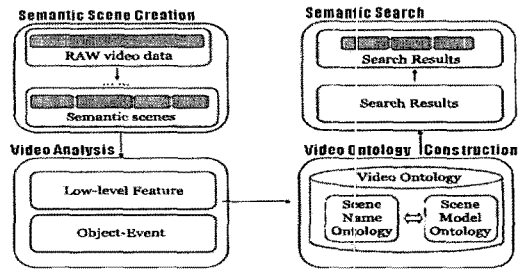


그림 3. 의미 기반 검색을 위한 온톨로지 시스템

을 위한 정보표현 및 저장방법이 통일되었다[11].

이처럼 초기의 내용기반 연구는 저수준과 고수준 정보의 관련성을 정의하지 않아 멀티미디어 데이터의 의미적 검색에는 어려움이 있었다. 이러한 문제들을 해결하기 위해서는 저수준 정보와 고수준 특징 정보 사이의 의미적 차이(semantic gap)를 극복하기 위한 연구가 필요하게 되었다. 이를 위해 온톨로지나 관련어집에 의해 구조화된 배경지식을 이용하는 연구가 시도되고 있다.

온톨로지를 이용해 장면검색을 하는 연구로는 Hoogs, Stein, Hollink들이 제안한 저수준의 특징정보와 고수준의 특징정보를 확장된 WordNet이라는 단어사전을 이용해 기술한 연구가 있다[12-14]. 하지만 이 연구들은 기존의 온톨로지인 WordNet을 이용했기 때문에 구조가 너무 방대하고 복잡하여 실제 동영상 검색을 위한 구조로 적합하지 못하다는 문제가 있다. 그리고 색인 구조가 키프레임 영상에 대해서만 기술하기 때문에 비디오 검색이 아니라 이미지 검색과 같아지는 문제가 있다.

기존의 선행연구에서 제시된 비디오 온톨로지 시스템은 위의 그림 3과 같은 구조를 가진다[15]. 의미적 장면 생성과정은 비디오를 장면단위로 모델링하고 검색하기 위한 전처리 과정이다. 비디오 분석과정에서는 MPEG-7표준에서 제공하는 멀티미디어 특징 정보 기술자를 이용해 동영상의 특징 정보를 추출해 낸다. 비디오 분석 과정을 통해 생성된 정보는 다음 단계인 비디오 온톨로지를 구축하는데 이용된다.

비디오 온톨로지 구축은 의미적 장면이 포함하는 내용의 키워드 트리가 저장되어 있는 장면 이름 온톨로지와, 장면이 가지는 멀티미디어 특징 정보의 관련성을 저장하는 장면 모델 온톨로지의 두 가지 온톨로지 로 이루어진다. 마지막으로 동영상 검색 부분은 구축된 비디오 온톨로지를 통해 의미적 기반 검색결과

를 보여준다.

3. 다중형상 비디오를 위한 색인 온톨로지

3.1 다중형상 분석

Snoek, Worring[2]들은 분석과 변환, 통합 과정을 수행하여 색인 정보를 추출하는 방법론을 제시하였다. 본 연구에서는 색인 추출 과정이 된다는 가정 하에 수작업으로 온톨로지에 색인 정보를 입력하였다. 아래의 그림 4는 영상 데이터가 분할, 변환, 통합되어 '문자적 내용(Textual content)'인 색인 정보가 추출 되는 과정을 보여준다.

분석을 위해서는 시각적이고 청각적인 형상 요소들을 텍스트로 변환하는 것이 가장 적절하다. 예를 들면 비디오 프레임들에서의 텍스트 탐지를 위한 비디오 OCR(Optical Character Recognition) 기법은 컴포넌트 기반으로 나누어질 수 있다. 비디오 OCR 기법의 사용에 의해 시각적으로 중첩된 텍스트 물체들은 문자적 형태로 변환이 가능해진다. 다양한 비디오 OCR 결과의 품질은 사용된 글자들의 종류, 색 등과 그 비디오 자체의 화질에 의존적이다.

또한 청각적 형상으로부터 일반적으로 발음되는 음성을 문자화해야 할 경우도 있다. 유효한 음성인식 시스템들은 한 명이 말하는 경우나 제한된 어휘 내에서는 잘 작동하는 것으로 알려져 있으나, 실제 환경에서 사용될 경우는 성능이 떨어지게 된다[16]. 위에서 언급한 변환 기법들을 통하여 텍스트화된 데이터들은, 통합 과정을 거쳐 그림 1에서 소개한 다섯가지 색인 정보들로 분류된다.

3.2 시맨틱 색인 용어간의 연관성 정의

관련 연구에서 발견된 인스턴스들로 구성된 시맨

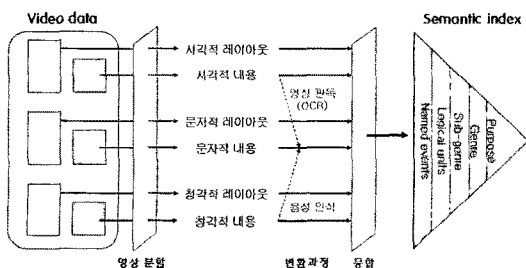


그림 4. 다중형상 분석과정

틱 색인 계층구조를 그림 5에서 나타내었다. 아래의 그림은 Snoek, Worring[2]들이 제안한 기본 구조에서 '어떤 것', '무엇'의 색인 정보를 표현한 것이며, 상위로부터 하위 인스턴스들은 목적, 장르, 하부장르, 논리적 단위, 지정된 사건 순이다. 가운데 부분의 점선은 유사한 노드들을 그룹화하는데 사용하였다.

Snoek, Worring[2]들은 시맨틱 색인을 표현함에 있어 색인 정보의 용어를 정의하였지만 용어간의 정확한 연관은 정의하지 않았다. 본 연구에서는 Snoek, Worring[2]의 연구를 기본 토대로 하여 시맨틱 색인 용어 이외에 용어간의 연관성 등을 온톨로지 언어로 확장 정의하였다. 또한 영상매체 색인 용어들을 클래스화하여 색인 온톨로지로서 구성된 내용이 사람이나 지능형 개체에 의해 저장, 회수될 수 있도록 하였다.

Snoek, Worring[2]들의 시맨틱 색인 용어 이외에 용어간의 연관성을 정의한 것은 그림 6에 표현된 것과 같다. 온톨로지는 Index 용어를 클래스로 기술하고 있으며 가장 일반화된 클래스이다. hasPrecision 술어는 생성된 색인 정보의 정확도를 기술하는 유형화된 리터럴을 값으로 가지는 술어이다. hasPrecision은 Index 클래스를 도메인으로 가지는 술어이기 때문에 Index 클래스의 모든 인스턴스는 hasPrecision을 술어로 가질 수 있다. Index의 모든 하부 클래스의 인스턴스는 Index 클래스의 인스턴스이기 때문에(i.e. IS-A relation) hasPrecision을 술어로 가질 수 있다. hasIndex 술어와 has_Index_For_Multimodal_Video 술어는 IS-A 계층 구조를 가지며 상위 계층은 일반적인, 하위 계층은 구체적인 연관을 기술한다.

Snoek, Worring[2]들은 시맨틱 색인을 정의할 때 크게 두 가지로 색인을 분류하였다. 영상매체 전체에 관련된 색인(목적, 장르, 하부장르)과 내용에 관련된

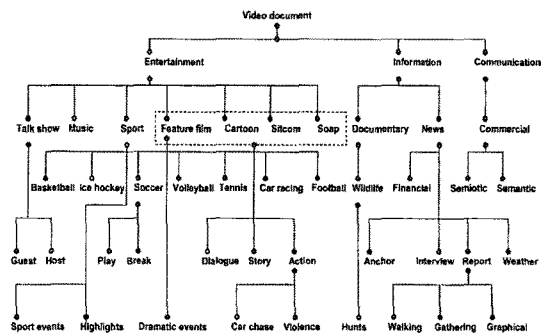


그림 5. 시맨틱 색인 정보

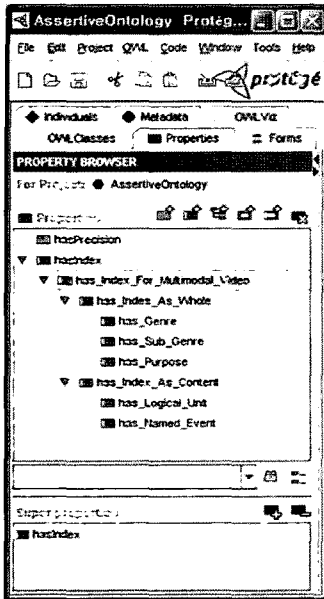


그림 6. 연관 용어 정의

색인(논리적 단위들, 지정된 사건들)으로 분류하였고, 색인들을 정의한 클래스와 색인간의 관계를 기술하기 위한 술어 역시 이 분류에 따랐으며 그림 6에 표현되어 있다.

그림 7은 영상매체 전체에 관련된 색인 용어를 클래스로 기술한 표현이다. 그림 7의 좌측 창 오른쪽 하단의 Car Racing 클래스는 Sub-Genre와 Sport의 하위 클래스로 기술 하였으며 Sub-Genre는 Genre의 하위 클래스로 기술하였다.

그림 7의 좌측 창에서 짙은 색 타원형으로 표현된 Cartoon 클래스는 필요충분 조건으로서 Story,

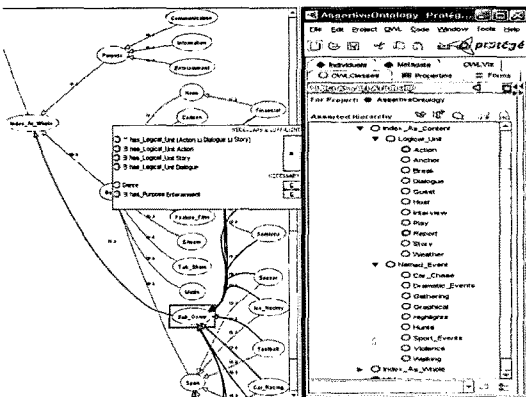


그림 7. 색인 정보를 위한 온톨로지

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns="http://idis.catholic.ac.kr/&O.owl#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:dami="http://www.dami.org/2001/03/dami+owl#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xml:base="http://idis.catholic.ac.kr/&O.owl">
  <owl:Ontology rdf:about="" />
  <owl:Class rdf:ID="Soap">
  <rdfs:subClassOf>
  <owl:Class rdf:ID="Genre"/>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
  <owl:Restriction>
  <owl:onProperty>
  <owl:ObjectProperty rdf:ID="has_Logical_Unit"/>
  </owl:onProperty>
  <owl:allValuesFrom>
  <owl:Class>
  <owl:unionOf rdf:parseType="Collection">
  <owl:Class rdf:ID="Action"/>
  <owl:Class rdf:ID="Dialogue"/>
  <owl:Class rdf:ID="Story"/>
  </owl:unionOf>
  </owl:Class>
  </owl:allValuesFrom>
  </owl:Restriction>
  </rdfs:subClassOf>
```

그림 8. XML/RDF 기반의 색인 온톨로지 일부

Action, Dialogue를 범위로 가지며, 클래스를 has_Logical_Units 술어에 의해 Closure Axiom으로 기술하고 있다. 즉, 특정 영상매체가 Story, Action, Dialogue 색인들만 has_Logical_Units 술어 관계를 가지면 Cartoon으로 클래스의 인스턴스가 된다는 서술논리 기반의 표현이다. Snoek, Worring[2]들이 정의한 시맨틱 색인 용어를 기반으로 서술논리 표현을 응용하면 추론엔진을 통하여 색인 정보의 완성된 확장이 가능하다.

그림 7의 우측 창은 영상매체의 내용에 관련된 색인 정보를 온톨로지로 기술한 모습이다.

색인 온톨로지는 그림 8과 같이 기계가 읽고 쓸 수 있는 형태의 XML/RDF 구문으로 재 표현될 수 있으며, 웹에 게시되거나 온톨로지 저장소에 적재되어 의미 정보를 회수하는 프로토콜을 통하여 색인 정보가 회수될 수 있다.

4. 모의 실험 및 결과

4.1 설계 지식의 저장과 회수

색인 온톨로지로 구성된 내용의 저장과 회수를 구현하기 위해 온톨로지 저장소의 기능 및 성능적인 요구사항을 모두 만족시키는 Minerva 온톨로지 저장소[8]를 이용하여 색인 온톨로지를 RDB(Relational Database)에 저장하였다. Minerva는 자체적으

로 사용되는 것이 아니라 기존 DB와 연동해서 사용되는 것이며, 본 구현에서는 IBM의 DB2 Express를 RDB로 사용하였다.

온톨로지 저장소를 이용하여 RDB에 저장한 XML/RDF 포맷의 색인 온톨로지는 OWL파서를 통하여 객체로 변환된 후, OWL 정보를 저장할 수 있는 스키마를 가진 데이터베이스에 저장된다. 아래의 그림 9는 색인 온톨로지에 담긴 AssertiveOntology가 RDB에 저장된 형식이다.

저장된 정보를 바탕으로 다음 쿼리를 통하여 MBC_2006_6_26_am04로 식별되는 자원의 하부장르 정보를 추출해 보았다. 아래의 그림 10은 온톨로지 저장소로부터 정보 추출을 위한 의미 기반 질의어이며, W3C에서 제공하는 RDF 스키마 형식 명세를 prefix로 참조한다.

위에서 수행한 질의에 대한 온톨로지 저장소로부터의 응답은 그림 11과 같다. MBC_2006_6_26_am04

로 식별되는 자원의 하부장르 정보가 Soccer로 식별되는 URI 형태의 결과로 회수되는 것을 확인할 수 있다.

아래의 그림 12와 같은 쿼리를 통하여 MBC_2006_6_26_am04로 식별되는 자원의 해당 장르정보를 추출하여 보았다. 다음은 정보 추출을 위한 의미 기반 질의어이다.

다음은 색인 온톨로지가 저장된 저장소로부터의 응답이다. 아래의 그림 13을 통하여 Sport로 식별되는 URI 형태의 결과가 회수됨을 알 수 있다.

4.2 실험 결과

본 구현을 통하여 Snoek, Worring[2]들이 제시한 시맨틱 색인과 같은 정보의 결과를 회수할 수 있었다. 위에서 살펴본 바와 같이 다중형상 비디오 콘텐츠의 분류와 색인에 관한 연구가 다양하게 이루어지고 있으며, 비디오 데이터의 도메인에 따라 해당 계층 구조가 다양해질 수 있음을 알 수 있다. 아래의 표 1을 통해 기존의 관련 연구와 본 연구를 데이터 모델과 시맨틱 색인 구조 및 저자의 관점에서 비교 설명하였다.

현재까지 대부분의 관련 연구에서는 비디오 콘텐츠를 계층 구조적인 객체로 모델링하고 있으나, 논문에서와 같이 각 계층 구조간의 관련성을 정의한 의미적 상호 운용성을 보장할 수 있는 방안을 제시하

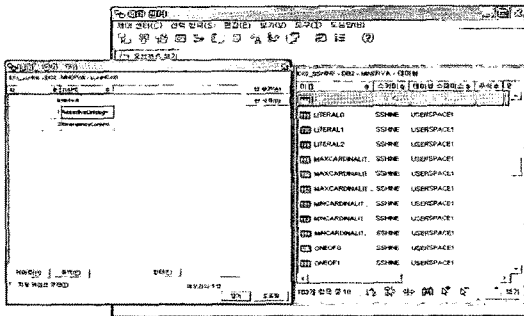


그림 9. AssertiveOntology

```

"PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> "+
"PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> "+
"PREFIX owl: <http://www.w3.org/2002/07/owl#> "+
"PREFIX : <http://idis.catholic.ac.kr/AO.owl#> "+
"SELECT * "+
"WHERE "+
"(:MBC_2006_6_26_am04 :has_Sub_Genre ?output)";
    
```

그림 10. 정보 추출을 위한 의미 기반 질의어

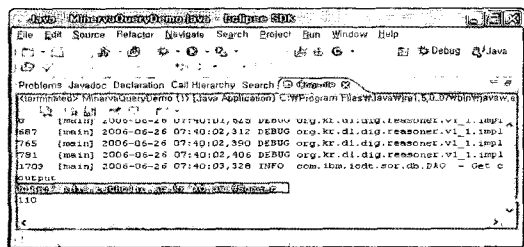


그림 11. 온톨로지 저장소로부터의 응답(Soccer)

```

"PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> "+
"PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> "+
"PREFIX owl: <http://www.w3.org/2002/07/owl#> "+
"PREFIX : <http://idis.catholic.ac.kr/AO.owl#> "+
"SELECT ?output "+
"WHERE "+
"(:MBC_2006_6_26_am04 :has_Sub_Genre ?path)+
" (?path rdf:type ?output) "+
" (?output rdfs:subClassOf :Genre)";
    
```

그림 12. 정보 추출을 위한 의미 기반 질의어

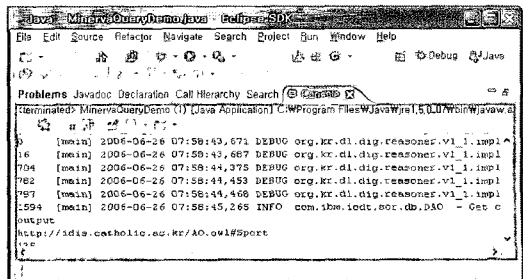


그림 13. 온톨로지 저장소로부터의 응답(Sport)

표 1. 기존 연구와의 비교

구 분	기존 연구	본 논문의 연구
데이터 모델	• 비디오 데이터 자체에 대한 논리적인 계층 구조 정보 모델링[2,9,15]	• 여러 비디오 데이터들간의 관련성을 정의
시맨틱 색인 구조	• 색인 정보의 용어를 정의하였지만, 용어간의 명확한 연관은 정의하지 않음[2,10,11,13,14]	• 시맨틱 색인 용어간의 연관성을 온톨로지 언어로 확장 정의 • 색인 온톨로지로 구성된 내용이 사람이나 지능형 개체에 의해 저장, 회수될 수 있도록 고려
저자의 관점	• 비디오 데이터를 어떻게 분류할지에 대한 방법론적인 관점 중심으로 서술[2,9,15]	• 기존 연구에서 분류 가능하도록 분석된 데이터들을 온톨로지화하여 데이터의 연관 관계까지 고려한 질의 결과를 얻을 수 있도록 함

지 못하고 있는 실정이다. 또한 시맨틱 색인 구조 측면을 살펴보면 기존의 관련 연구에서는 색인 정보의 용어까지만을 정의하고 있을 뿐, 각 용어간의 명확한 연관은 정의하지 않고 있기 때문에 실제 검색에서의 활용 범위에 제한이 있다. 그러나 본 논문에서는 시맨틱 색인 용어간의 연관성을 온톨로지 언어로 확장 정의하고, 색인 온톨로지로 구성된 내용이 사람이나 지능형 개체에 의해 저장, 회수될 수 있음을 보이는 실험을 진행하였다.

비디오 콘텐츠의 색인 기법에 관한 대부분의 연구에서는 비디오 데이터 분류를 위한 방법론들만을 제안하고 있으나, 본 연구에서는 기존 연구 결과들에 온톨로지의 장점을 결합함으로써 다양한 형태의 질의와 색인정보의 활용을 지원할 수 있는 기반 모델을 제시한다는 차이점이 있다.

5. 결 론

색인 기법은 오래 전부터 데이터 검색의 중요한 기술로 널리 사용되었으며, 지능화된 사회에 접어들면서 복잡한 콘텐츠의 자동 색인 기법에 관한 많은 연구가 진행되고 있다. 방대한 양의 영상매체에 잠재된 가치를 회수하기 위해서는 자동 색인 기능이 필수적으로 요구되며, 사람과 지능화된 기계간의 의미적 합이가 보장된 형태의 색인이 요구된다.

본 연구에서는 Snoek, Worrying[2]들의 연구를 토대로 시맨틱 색인 용어 이외에 용어간의 연관성 등을 시맨틱 웹 기술의 온톨로지를 이용하여 확장하고, 추상화된 색인 정보의 생성과 제어를 지원하는 실험을 하였다. 또한 사람과 기계간, 기계와 기계간의 의미적 상호 운용성이 보장되는 프레임워크를 제안하였으며, 실험연구를 통해 본 논문에서 제안한 기반 구

조의 효용성을 나타내었다.

영상매체의 색인 용어들을 클래스화하여 색인 온톨로지로 구성한 내용이 사람이나 지능형 개체에 의해 저장 및 회수될 수 있으며, 시맨틱 색인 용어를 기반으로 한 서술논리 표현의 응용으로 추론엔진을 통한 색인 정보의 완성된 확장이 가능해지게 되었다.

향후 연구과제로 MPEG-7과 같은 멀티미디어 콘텐츠 기술 표준을 이용하여, 사용자가 원하는 자료를 빠르고 효율적으로 검색할 수 있도록 하는 연구와 연계하는 방안을 고려할 수 있다. 이러한 추가적인 연구들을 활용하면 기존 영상검색 시스템의 단점을 보완한 효율적인 검색 시스템 구축이 가능해질 것이다.

참 고 문 헌

- [1] 황금숙, “디지털 도서관론(개정증보),” (주)학문사, 고양시, 2003.
- [2] Cees G.M. Snoek and Marcel Worrying, “Multimodal Video Indexing: A Review of the State-of-the-art,” *Multimedia Tools and Applications*, Vol.25, pp. 5-35, 2005.
- [3] Laurence Nigay and Joelle Coutaz, “A design space for multimodal systems: concurrent processing and data fusion,” *Proceedings of the INTERCHI '93 conference on Human factors in computing systems*, pp. 172-178, 1993.
- [4] World Wide Web Consortium, <http://www.w3c.org>
- [5] 김성태, 도승철, 이준규, 박수민, 양정진, “위치 기반 멀티 에이전트를 활용한 긴급 구조 시스템 모델,” *한국지능정보시스템학회 춘계학술대회*

논문집, pp. 168-178, 2004.

[6] Jaikumar Vijayan, "Web Services, Internet Collaboration Pose Security Challenges for 2002," IT World, Canada, 2002.

[7] 박영택, 최중민, "온톨로지 추론 개요와 연구동향," *정보과학회지*, 제24권, 제4호, pp. 17-23, 2006.

[8] Jian Zhou, Li Ma, Qiaoling Liu, Lei Zhang, Yong Yu, and Yue Pan, "Minerva: A Scalable OWL Ontology Storage and Inference System," *Proceedings of the First Asian Semantic Web Conference*, pp. 429-443, 2006.

[9] 임수민, 정진우, 손진현, 이동호, "고수준 의미 정보 자동 추출을 위한 이벤트 온톨로지 기반 비디오 주석시스템," *한국정보과학회 가을 학술발표논문집*, 제35권, 제2호, pp. 246-251, 2008.

[10] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Huang Qian, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: The QBIC system," *IEEE Computer*, Vol.28, No.9, pp. 23-32, 1995.

[11] J. M. Martinez, "Overview of the mpeg-7 standard. technical report 5.0," ISO/IEC, Singapore, 2001.

[12] C. Fellbaum, "WordNet: An Electronic Lexical Database," Bradford Book, 1998.

[13] L. Hollink, A.T. Schreiber, J. Wielemaker, and B. Wielinga, "Semantic annotation of image collections," *In Proc. of Knowledge Markup and Semantic Annotation Workshop, USA*, 2003.

[14] A. Hoogs, J. Rittscher, G. Stein, and J. Schmiederer., "Video content annotation using visual analysis and a large semantic

knowledgebase," *In Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.2, pp. 379, 2003.

[15] 정민영, 박성한, "비디오 서버에서 온톨로지를 이용한 의미 기반 장면 검색," *전자공학회 논문지*, 제45권, 제5호, pp. 427-432, 2008.

[16] R. Brunelli, O. Mich, and C.M. Modena, "A survey on the automatic indexing of video data," *Journal of Visual Communication and Image Representation*, Vol.10, No.2, pp. 78-112, 1999.



황 우 연

2001년 3월~2008년 2월 가톨릭대학교 컴퓨터공학과 학사
 2008년 3월~현재 가톨릭대학교 컴퓨터공학과 석사과정
 관심분야 : 에이전트 시스템, 온톨로지, 웹 서비스



양 정 진

1985년 이화여자대학교 전자계산학과 학사
 1992년 University of Connecticut 공학석사
 1998년 University of Connecticut 공학박사

1999년 University of Connecticut Post Doc.
 1999년~2000년 University of Hartford 조교수
 2001년~현재 가톨릭대학교 컴퓨터·정보공학부 조교수, 부교수
 관심분야 : 지능형 (다중)에이전트 시스템, 유비쿼터스 컴퓨팅, 시맨틱 웹 서비스