멀티미디어 온톨로지 기반의 의미론적 콘텐츠 검색 시스템 설계

황치곤*, 문석재*, 이대성**, 윤창표***

*광운대학교, **부산카톨릭대학교, ***경기과학기술대학교

A Design of Semantic Contents Search System for Multimedia Ontology

Chi-Gon Hwang* · Seok-Jae Moon*, Daesung Lee**, Chang-Pyo Yoon***

*KwangWoon University

**Catholic University of Pusan

***GyeongGi College of Science and Technology

E-mail : duck1052@kw.ac.kr

요 약

최근 멀티미디어 기술과 네트워크의 발전으로 멀티미디어 콘텐츠의 생성은 빠르게 늘어가고 있다. 그러나 이를 찾고 활용하는 기술은 부족하다. 이러한 문제를 보완하고 해결하기 위해 멀티미디어 콘텐츠를 위한 표준안들이 많이 있으나 멀티미디어 데이터 형식과 상호운용성을 표현하기에는 부족하다. 본 논문에서는 멀티미디어 콘텐츠를 상호 운용할 수 있도록 하는 온톨로지 기반의 콘텐츠 검색 시스템을 제시한다. 온톨로지는 MPEG-7의 MDS(Multimedia Description Scheme)와 같은 스키마 구조에 온톨로지 구축을 위한 규칙을 제시하여 구성한다. 이러한 온톨로지를 바탕으로 멀티미디어 데이터가 가지고 있는 연관성을 확장하여 의미론적 검색이 가능한 접근 시스템을 설계한다.

ABSTRACT

With the development of multimedia and network technology, the production of multimedia contents is rapidly increasing. Meanwhile, the technology to search and use the contents is still insufficient. There are standards for multimedia contents to address the problem, but they cannot fully support diverse multimedia data types or ensure their interoperability. In this paper, an ontology-based content search system is proposed to ensure the interoperability of multimedia contents. The ontology is configured by presenting the rules for it using the schema structure of the multimedia description scheme (MDS) of MPEG-7. Based on this ontology, the association of the multimedia data is expanded to design an access system that allows semantic search.

키워드

멀티미디어 콘텐츠, 상호운용, 온톨로지, 의미론적 접근, 메타 정보

Ⅰ. 서 론

최근 멀티미디어 콘텐츠를 효율적으로 운용하기 위한 연구들이 진행되고 있다. 이 연구들 중에 콘텐츠의 메타 데이터를 기술하는 기법들이 있으며, 이에 따라 콘텐츠에 대한 의미 검색이 가능해지고 있다. 지금까지 제시된 멀티미디어 메타 데이터를 표현하는 대표적인 표준으로 MPEG-7, TV-Anytime, MPEG-21, Cablelabs1.1, DIG35, Dublin Core등이 있으며, 이중에 MPEG-7은 멀티미디어 표현보다 멀티미디어 내용을 표현하기 위한 의미론적 관점에서 다루고 있다. 그러나 사용자는 서로 다른 표준을 따르는 콘텐츠가 지속적으로 증가함에 따라 콘텐츠의 검색이 어려워지고 있다.[1] 이러한 문제점을 개선하기 위해 온톨로지를 이용하는 방법이 있으며, 온톨로지는 메타 데이터 사이에서 발생하는 충돌을 해결할 수 있다. 이를 위한 규칙을 제시함으로써 온톨로지를 효율적으로 구성할 수 있다[2]. 효율적인 콘텐츠 검색을 위해 위의 단점을 해결하고 온톨로지에 규칙을 추가할 필요가 있다.

본 논문의 구성은 다음과 같다. 제2장은 본 논

문을 위한 관련연구에 대해 기술한다. 3장은 멀티미디어 온톨로지를 기반으로 제공되는 시스템의 구성, 온톨로지의 구성과 표현을 기술하고, 온톨로지 구성에 따른 충돌 해결을 위한 규칙을 제시한다. 그리고 4장에서는 결론과 향후 연구에 대해 기술한다.

II. 관련연구

MPEG은 다양한 멀티미디어와 상호작용을 지원하는 콘텐츠를 처리하기 위한 MPEG-4 표준과 콘텐츠에 대한 메타 데이타 표준인 MPEG-7 표준을 제정하였으며, 현재 디지털 콘텐츠의 전자상거래에 필요한 요소기술들과 통합기술의 표준화를 멀티미디어 프레임워크인 MPEG-21이라는 이름으로 진행 중이다. MPEG-7은 데이터 그 자체가 아닌 데이터의 내용에 대한 표현방법인 '메타 데이터(metadata)'에 대해 다루었다[3][4]. MPEG-21은 전자상거래 환경에서의 다양한 네트워크와 단말장치에서 생성, 전송, 인증 등을 실행할 수 있는 응용 애플리케이션을 위해 총 7개의 기술요소로 멀티미디어 프레임워크 표준을 생성하였다[5].

이러한 표준 중 멀티미디어 콘텐츠 생성을 위한 표준보다는 콘텐츠의 의미적 검색을 위해 온톨로지의 개념을 포함한 MPEG-7을 기반으로 한 멀티미디어 온톨로지 구성한다. 그리고 콘텐츠 간에 발생할 수 있는 의미적, 구조적 충돌을 해결하기 위한 규칙을 제시하고, 이를 기반으로 하는 멀티미디어 검색 시스템을 제안한다.

III. 멀티미디어 온톨로지 기반의 검색 시스템

멀티미디어 온톨로지는 메타 정보를 기준으로 하는 개념 구조를 기준으로 콘텐츠들의 연관성을 저장함으로써 구성된다. 이를 기반으로 콘텐츠 검색 시스템을 구성한다.

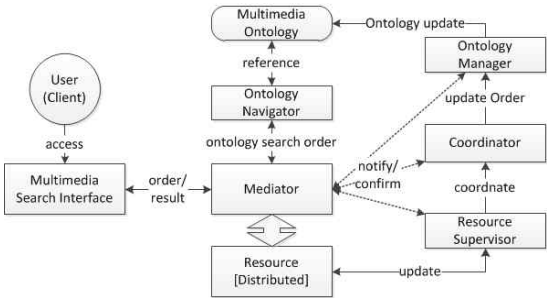

그림 1 온톨로지 기반의 검색 시스템

3-1. 제안 시스템

본 시스템은 사용자의 요구에 따라 멀티미디어 온톨로지를 참조하여 사용자에게 연관 정보를 제공하며, 이에 따라 사용자가 선택한 연관관계에 따라 분산된 콘텐츠 자원들을 검색할 수 있도록 한다. 그림 1은 제안 시스템의 구성과 구성요소 간에 발생하는 작업을 요약하여 나타낸 것이다.

그림에서 표현된 화살표는 각 요소들 간의 상호작용이고, Mediator에 의한 상황에 따른 확인 과정이다. 다음은 시스템에 포함된 각 구성 요소들의 기능이다

- Multimedia Search Interface:이 요소는 사용자의 입력에 따라 작업의 시작을 Mediator에게 알리고 사용자의 행동을 전달한다.
- Mediator:시스템의 전반적인 운영을 담당하며, 온토로지 검색 요구, 사용자의 검색 조건에 따른 컨텐츠 정보 검색, 온톨로지 갱신을 위한 감독한다.
- Ontology Navigator:사용자의 요구 조건을 통한 온톨로지를 참조하여 연관 정보를 추출하여 반환한다.
- Resource Supervisor:분산된 자원의 수정여부를 확인하여 온톨로지의 갱신을 요구한다.
- Coordinator:Mediator를 통한 온톨로지와 Resource Supervisor의 갱신된 정보를 비교하여 온톨로지를 조정한다.
- Ontology Manager:Coordinator에 따라 온톨로지를 갱신한다.

이에 따른 온톨로지의 갱신은 Mediator에 주기적으로 수행하여 콘텐츠의 관계성을 최신의 상태로 유지한다.

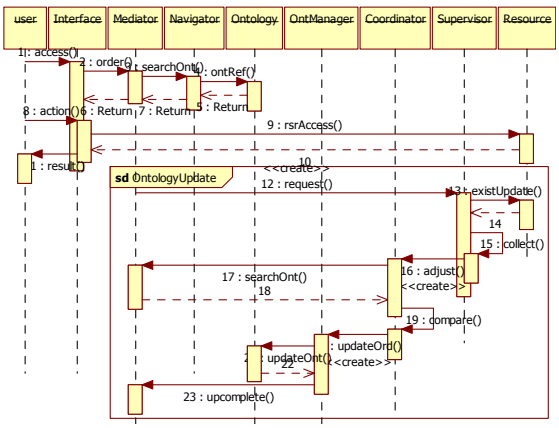

그림 2 검색과 온톨로지 갱신을 위한 시퀀스

시스템의 각 구성요소의 수행 과정은 그림 2와 같다. 과정은 두 개로 나뉜다. 하나는 사용자에 의한 액션으로 일어나는 온톨로지를 통한 콘텐츠 리소스 접근과정으로 시퀀스 1부터 11까지이며, 둘째는 그림2에서 프레임 내에 있는 과정으로, mediator를 통한 주기적인 콘텐츠 변경사항을 확인하여 온톨로지를 갱신시키는 과정으로 시퀀스 12부터 23까지 이다.

3-2. 멀티미디어 콘텐츠를 위한 온톨로지

온톨로지의 구성은 MPEG-7을 기반으로 콘텐츠를 분류하고, 이 분류를 바탕으로 각 개체를 표현하기 위한 클래스(class)와 이 클래스들을 표현하기 위한 속성(property)들로 구성하다. 이 구성에 대한 개념은 그림 3과 같고, 타원은 클래스, 사각형은 속성, 직선화살표는 서브클래스(subClassOf), 점선화살표는 속성의 도메인(domain), 일점쇄선은 속성의 범위(range)에 해당한다.

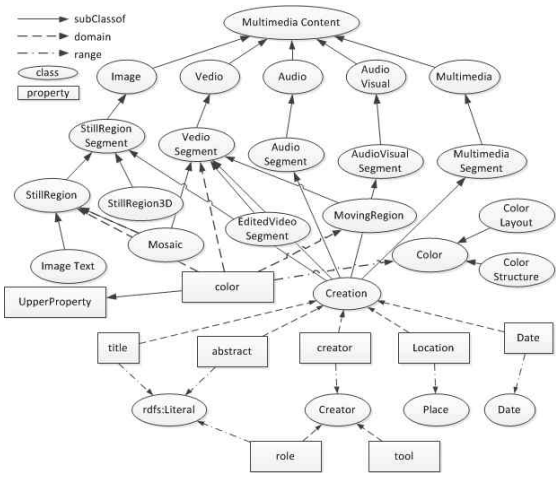

그림 3 멀티미디어 콘텐츠를 위한 온톨로지 구성

```
<complexType name="CreationType">
 <complexContent>
  <extension base="mpeg7:DSType">
   <sequence>
    <element name="Title“
        type="mpeg7:TitleType"/>
    <element name="Abstract“
        type="mpeg7:TextAnnotationType"/>
    <element name="Creator">
     <complexContent>
      ...
     </complexContent>
    </element>
    <element name="CreationLocation“
        type="mpeg7:PlaceType"/>
    <element name="CreationDate“
        type="mpeg7:DateType"/>
   </sequence>
  </extension>
 </complexContent>
</complexType>
```

그림 4 Creation 클래스의 스키마 정의

그림 3에서 Creation 클래스를 예를들어 보면, 이 클래스는 콘텐츠에 대한 제작 정보로 콘텐츠들의 서브클래스이며, 해당 속성으로 title, abstract, creator, location, date를 가지고 있음을 확인할 수 있다. title과 abstract 속성은 rdfs:Literal으로 문자열 형태의 인스턴스를 가질 수 있게 구성되어 있음을 나타낸다. 이는 전체 온톨로지 구조의 일부이다.

그림 4은 Creation 클래스의 ComplexType에 대한 스키마 정의로 그림 3의 Creation 클래스를 위한 타입으로 <sequence>에 의해 필드들을 <element> 태그로 나열한다. 각 element에서 name는 필드의 이름이고, type은 필드의 기술 타입로 변환됨을 나타낸다. 그림 5는 그림 3에서 정의한 속성과 개체를 RDF로 표현하는 것이다. 이에 예로 사용한 것은 color 속성과 Color 클래스이다. 속성은 3개의 도메인과 1개의 범위를 가진 것을 기술하고 있다. 클래스는 해당 클래스의 식별자와 해당 클래스의 하위클래스를 표현하고 있음을 표현한다.

```
<rdf:Property rdf:ID="color">
  <rdfs:label>color of a video or audio segment
  </rdfs:label>
  <rdfs:domain rdf:resource="#Mosaic"/>
  <rdfs:domain rdf:resource="#VedioSegment"/>
  <rdfs:domain rdf:resource="#MovingRegion"/>
  <rdfs:range rdf:resource="#Color"/>
</rdf:Property>
<rdfs:Class rdf:ID="Color">
  <rdfs:label>Color</rdfs:label>
  <rdfs:subClassOf rdf:resource="#ColorLayout"/>
  <rdfs:subClassOf rdf:resource="#ColorStructure"/>
</rdfs:Class>
```

그림 5 구성한 온톨로지를 위한 RDF 적용예

이에 따라 온톨로지는 의미적 연관성을 가지며 데이터 간 발생하는 충돌을 해결할 수 있는데, 각 충돌의 해결을 위한 다음 규칙을 포함한다.

규칙 1. 클래스 간에 같은 이름을 가짐으로 발생하는 충돌을 해결한다.

규칙 2. 연관성을 가지는 클래스 간에 구조가 달라서 발생하는 충돌을 해결한다.

규칙 3. 클래스와 인스턴스 사이에서 발생하는 구조적 충돌을 해결한다.

규칙 4. 클래스와 속성 사이에서 발생하는 구조적 충돌을 해결한다.

규칙 5. 속성과 인스턴스 사이에서 발생하는 구조적 충돌을 해결한다.

규칙 6. 속성 사이에서 발생하는 이름의 차이에 의해 발생하는 충돌을 해결한다.

규칙 7. 속성 사이에서 존재하는 구조적 차이에 의해 발생하는 충돌을 해결한다.

이와 같은 방법으로 구성된 온톨로지는 클래스

간의 연관성과 충돌 문제를 해결할 수 있다.

Ⅳ. 결론 및 향후 연구

본 연구는 콘텐츠 정보를 효율적으로 검색하기 위한 방안으로 멀티미디어 온톨로지의 구성방안과 충돌해결을 위한 규칙을 제시하고, 이를 바탕으로 콘텐츠를 검색하기 위한 시스템 모델을 제시하였다. 이에 따라 사용자에게 의미 기반 검색을 제공할 수 있으며, 콘텐츠의 검색에서 온톨로지를 RDF 형식으로 구성함으로써 데이터베이스 형태의 매핑으로 구성된 온톨로지에 비해 온톨로지의 접근에 소요되는 시간과 기억공간을 줄일 수 있다.

제안하는 방식은 단일 온톨로지를 구성하여 분산된 콘텐츠의 연관 정보를 이용하여 검색할 수 있는 방안을 제시하였다. 최근 분산된 온톨로지에 대해 통합 온톨로지가 연구되고 있다. 이에 따라 통합 온톨로지 구축과 이를 이용하여 다양한 콘텐츠를 이용할 수 있도록 확장하고, 제시된 규칙을 정규화 할 필요가 있다.

참고문헌

[1] 김선경, 신판섭, 임해철, “멀티미디어 콘텐츠의 강화된 의미 검색을 위한 온톨로지 설계,” 한국인터넷정보학회, 13권, 1호, pp.107-115, 2012. 02.

[2] 황치곤, 신효영, 정계동, 최영근, “협업에서 서비스 상호운용을 위한 온톨로지 설계 및 구현,“ 한국정보통신학회논문지, 16권, 12호, pp.2777-2783, 2012. 12.

[3] MPEG-7 and the Semantic Web http://www.w3.org/2005/Incubator/mmsem/XGR-mpeg7/, 2007.8

[4] J.Hunter, "Adding Multimedia to the Semantic Web - Building and Applying an MPEG-7 Ontology", Chapter 3 of "Multimedia Content and the Semantic Web", Eds. Giorgos Stamou and Stefanos Kollias, Wiley, 2005

[5] 손소현, 홍기채, 문병주,“ 멀티미디어 프레임워크 표준: MPEG-21,” 전자통신동향분석, 17권, 6호, 2002. 12.