

XML 기반 멀티미디어 검색시스템의 설계에 관한 연구

윤미희*, 조동욱**

충북과학대학 컴퓨터정보과학과

충북과학대학 정보통신과학과

e-mail:mihee@ctech.ac.kr

A Study on Design of Multimedia Retrieval System based on XML

mihee yoon*, donguk cho**

*Dept of Computer Information Science, Chungbuk Provincial University of
Science and Technology

**Dept of Information Communication Engineering, Chungbuk Provincial
University of Science and Technology

요 약

비디오는 멀티미디어 데이터의 가장 대표적인 형태로, 텍스트나, 이미지, 오디오와 객체의 움직임 같은 풍부한 정보를 담고 있다. 비정형의 멀티미디어 데이터를 다양하고 효율적으로 표현하기 위해서는 XML(eXtensible Markup Language)을 사용하여 저장하고 검색하는 멀티미디어 검색시스템이 필수적이다. 그러므로 멀티미디어 데이터에 대한 검색을 위해서는 멀티미디어 데이터의 내용을 구조적으로 설명하는 메타데이터가 필요하고 이 메타데이터를 XML을 사용하여 표현하며 저장하고 검색하기 위한 멀티미디어 검색시스템이 요구된다. 본 논문에서는 XML 메타데이터 모델링 기법과 이 모델링 기법을 기반으로 한 멀티미디어 검색시스템을 제안한다.

1. 서론

인터넷의 보급과 더불어 내용량 저장장치의 가격이 하락하면서 멀티미디어 데이터의 저장과 공유가 보편화되었다. 또한 멀티미디어 데이터를 활용하는 지리정보시스템, 멀티미디어 의료 정보시스템, 전자신문, 전자도서관, 홈쇼핑, VOD 등의 새로운 응용 분야들이 각광을 받고 있다. 이러한 환경적인 요인은 내용량의 멀티미디어 데이터를 효율적으로 추출, 저장, 검색하기 위한 멀티미디어 데이터 처리기술이 요구된다.[1,2]

비디오는 멀티미디어 데이터의 가장 대표적인 형태로, 텍스트나, 이미지, 오디오와 객체의 움직임 같은 풍부한 정보를 담고 있다. 이러한 멀티미디어의 다양한 정보를 조직적으로 기술하기 위한 정의 언어와 스키마가 필요함에 따라 MPEG(Motion Picture Expert Group)은 멀티미디어 정보를 기술하고 다양한 도구들을 제공하

는 MPEG-7[3]을 발표하였다. 비정형의 멀티미디어 데이터를 다양하고 효율적으로 표현하기 위해서는

XML(eXtensible Markup Language)[4]을 사용하여 저장하고 검색하는 멀티미디어 검색시스템이 필수적이다.

일반적으로 XML 문서의 검색은 요소(Element)단위의 구조검색으로 이루어지며 검색결과와 현재는 조건에 맞는 요소들의 집합으로 표현된다. XML 문서의 특성을 고려한 다양한 구조 검색 방법이 이루어지고 있지만 문서 구조에 대한 사전지식이 필요하다. 그러므로 멀티미디어 데이터에 대한 검색을 위해서는 멀티미디어 데이터의 내용을 구조적으로 설명하는 메타데이터가 필요하

고 이 메타데이터를 XML을 사용하여 표현하며 저장하고 검색하기 위한 멀티미디어 검색시스템이 요구된다.

본 논문에서는 멀티미디어 데이터의 검색을 위한 XML을 기반으로 한 메타데이터 모델링 기법을 제안하고 이 모델을 기반으로 특징기반 검색과 주석기반 검색을 통합하여 다양한 사용자의 의미검색을 지원하기 위한 멀티미디어 검색시스템을 제안한다. 본 논문의 구성은 2장에서는 관련연구로 기존의 멀티미디어 검색시스템에 대해 살펴보고 3장에서는 멀티미디어 데이터의 메타데이터 모델과 이 모델을 XML 모델로 표현하고 5장에서는 XML 기반의 멀티미디어 검색 시스템을 제안한다. 마지막으로 6장에서는 결론 및

향후 연구과제에 대해 기술한다.

2장 관련연구

VIDEO-STAR[9]는 비디오와 메타데이터의 공유와 재사용을 지원하기 위해 관계형 데이터베이스 모델을 사용한 데이터베이스 시스템이다. 메타데이터의 구성은 분할기법을 통해 비디오데이터를 구조화하고, 구조화된 비디오데이터에 대해 people, location, event의 메타데이터를 구성하고, 이를 다시 기본, 일차, 이차 컨텍스트로 구분하여 재사용과 공유를 용이하게 하며, 고정된 속성을 통해 사용자가 쉽게 질의를 구성할 수 있다. 그러나 메타데이터가 people, location, event의 내용만으로 한정되어 있어 다양한 의미검색이 불가능하고, 질의를 처리하기 위해서는 비디오데이터베이스를 검색할 때 'Decompose', 'Map to composition'등의 대응연산을 수행하여 해당 비디오데이터와 메타데이터를 관련시켜야 하므로 질의처리가 복잡하다.

AVIS(Advanced Video Information System)[4]는 미국 메릴랜드 대학에서 개발된 비디오데이터베이스 시스템으로 비디오 내용에 나타나는 객체, 사건 행위 유형에 대한 메타데이터를 정의하고, 이들을 비디오세그먼트와 연계시킴으로써 효율적인 검색 방법을 제안했다. 그러나 AVIS는 상위 레벨의 내용을 중심으로 하는 주석기반의 메타데이터만을 제공하고, 하위레벨을 위한 메타데이터가 없으며 명확한 질의언어를 제공하지 않는 단점이 있다.

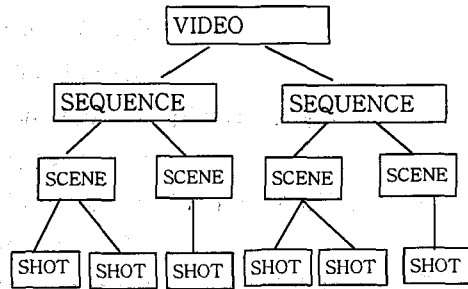
Venus[8]는 특징기반 검색을 위해 개발된 비디오 데이터베이스 시스템이다. 이미지 처리 기술을 사용하여, 비디오의 각 프레임에 나타나는 객체들의 시간관계, 공간관계를 메타데이터로 구축하였다. Venus는 특징기반 질의언어인 CVQL(Content-based Video Query Language)을 제공한다. 그러나 비디오 내용에 대한 상위 레벨의 의미를 위한 메타데이터가 없으므로 다양한 질의를 구성할 수 없다.

OVID[3]는 비디오데이터를 모델링하고 검색하기 위해 객체지향 접근 방법을 사용한 객체지향 비디오데이터베이스 시스템이다. OVID는 의미 있는 장면을 비디오 객체라는 개념을 도입하고 각 비디오 객체마다 포함된 내용을 설명하는 속성과 속성값을 부여하여 메타데이터를 구성하였다. OVID는 비디오객체를 검색하기 위해 Video SQL이라는 사용자 인터페이스를 제공한다. OVID는 비디오 객체모델이 스키마를 사용하지 않았기 때문에 사용자는 각 비디오객체마다 완전한 속성과 속성값을 지정해야 하고 비디오데이터의 구조가 일반화되어 있지 않아 질의형식을 지정할 수 없다.

3. 메타데이터 모델링

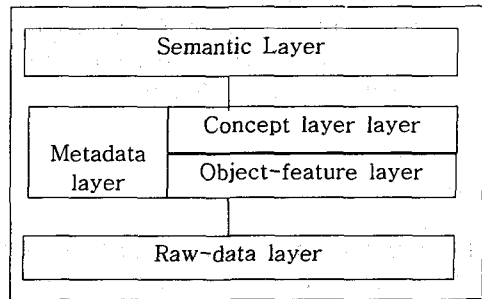
본 논문에서는 비디오데이터의 메타데이터 모델을 위해 세 개의 계층을 가진 객체지향 메타데이터 모델을 제안한다. 세 개의 계층을 가진 객체지향 메타데이터 모델(Three layered Metadata Model : TMM)은 원시데이터 계층

(raw-data layer), 비디오의 의미검색을 위해 특징과 주석을 통합하는 메타데이터 계층과, 질의 재형성을 위한 의미 계층으로 구성되며, 모델링을 위해 통합모델링 언어(UML)를 사용하였다. 한편의 비디오는 필름이론에 의해 시퀀스, 장면, 화면, 프레임으로 구성된다. 비디오의 계층적 구조는 (그림 1)과 같다.



(그림 1)비디오의 계층적 구조

본 논문에서는 동일시간, 동일 장소에서 발생하는 화면을 처리의 기본 단위로 한다. 비디오를 위한 메타데이터 모델은 (그림 2)와 같다. 각 계층에 대한 정의는 다음과 같다.



(그림 2) 세 개의 계층을 가진 통합 객체지향 메타데이터 모델

(정의 1) 원시데이터계층(Raw-data layer)

물리적인 비디오스트림을 표현하는 계층

(정의 2) 메타데이터계층(metadata layer)

비디오에 대한 정보와 비디오스트림에 포함된 의미를 기술하기 위한 계층으로 개념계층(concept layer)과 객체특징계층(object-feature layer)의 부계층(sublayer)으로 구성된다.

(정의 2-1) 개념 계층(concept layer)

비디오 정보와 비디오스트림에 나타나는 추상적인 의미를 표현하기 위한 계층

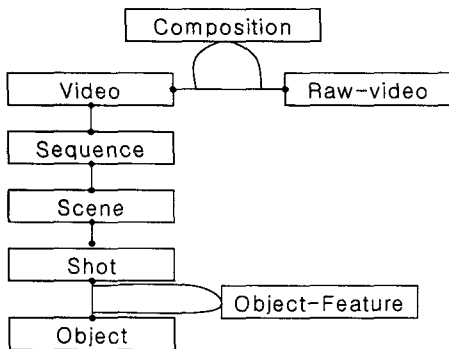
(정의 2-2) 객체특징계층(object-feature layer)

장면에 나타나는 객체들의 특징을 기술하기 위한 계층

(정의 3) 의미 계층(semantic layer)

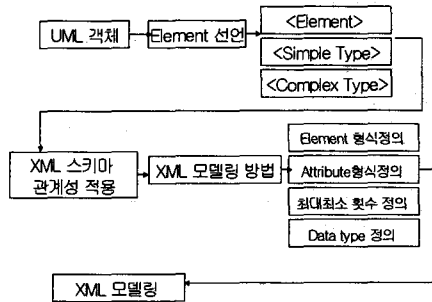
메타데이터계층에 나타나는 의미들 사이의 연관관계를 표현하는 계층으로 IF - THEN규칙의 형태로 표현된다. 본 논문에서는 IF - THEN규칙을 지식이라 칭하고 지식베이스를 구성한다.

메타데이터계층의 개념계층은 주석기반질의를 지원하고 객체특징계층(object-feature layer)은 특징기반질의를 지원한다. 주석기반질의는 한편의 비디오에 대한 질의나 기사의 질의, 인물이나 사건과 같은 질의가 가능하고, 한 장면에 나타나는 관심 있는 객체의 특징질의는 객체특징계층을 이용하여 객체의 시간관계질의, 공간관계질의, 한 장면에 나타나는 객체들간의 시공간상관관계의 질의가 가능하다. 또한 메타데이터계층을 이용하여 유사성질의를 지원할 수 있다. 의미계층은 사용자의 주관적인 관점으로 쓰여진 질의를 처리하기 위한 계층으로 질의의 결과가 존재하지 않을 경우 이 계층의 지식을 사용하여 질의를 재형성 할 수 있다. 원시대데이터계층은 원시대데이터객체로 구성되고 메타데이터계층은 비디오, 시퀀스, 장면, Object, 객체특징 객체로 구성된다. (그림 3)은 비디오의 메타데이터를 위한 UML 모델링기법이다.



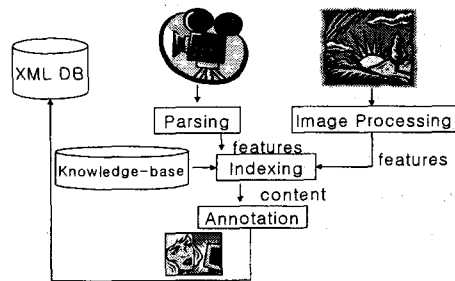
(그림 3) 메타데이터 모델링

UML객체를 이용한 XML 모델링 과정은 (그림 4)와 같다.



(그림 4) XML 모델링

비디오데이터에 대한 메타데이터의 구성은 (그림 5)와 같다. 이미지데이터는 전처리과정을 통해 특성을 추출하고 추출된 특성을 기반으로 의미를 부여한다. 비디오는 샷 검출과 자막데이터를 기반으로 세그멘테이션 처리과정을 거치고, 세그멘테이션된 결과를 가지고 주석자는 샷 장면, 시퀀스를 구분하고 메타데이터의 각각의 계층에 적당한 키워드를 삽입하여 XML 데이터베이스에 저장한다.



(그림 5) 메타데이터 처리

4. XML 기반 멀티미디어 검색시스템

멀티미디어 검색 시스템은 생성된 메타데이터의 XML 데이터베이스를 이용하여 사용자가 원하는 이미지나 비디오의 장면 또는 샷을 검색한다. 검색은 웹서버와 XML 엔진 비디오 엔진, 데이터베이스가 담당한다. XML 엔진은 XSLT 프로세서, 파일시스템, DOM, XML 데이터베이스를 위한 API등으로 구성된다. XSLT 프로세서는 XML 문서를 HTML 문서로 변환할때 사용되고, 저장된 XML 문서의 갱신이나 질의 및 XSLT 컴포넌트는 W3C XML 표준 인터페이스를 이용한다[8]. 사용자의 질의는 XML 문서에 스타일 언어인 XSLT를 적용시켜 폼을 구성하여 사용자의 입력을 받고 폼을 통해 입력 받은 값은 Xquery를 사용하여 XML 데이터

베이스에 대한 질의를 통하여 결과 값을 산출한다. 산출된 결과 값은 사용자가 원하는 형태로 보여주기 위해 적당한 XSLT 문서를 적용하여 HTML 문서로 만들어져 미디어 플레이어나 웹브라우저를 통해 브라우징된다. 비디오의 경우 계층적인 구조로 구성되었기 때문에 시퀀스, 장면, 샷을 기반으로 검색하고 추출할 수 있다.

5. 결론 및 향후 연구과제

효율적인 멀티미디어 검색시스템을 위해서는 자동화된 멀티미디어 분석기법과 효율적인 검색을 위한 인덱싱 기법과 자동적인 지식 발견에 의한 지식베이스 구성 기법을 개발하는 것이 요구된다.

참고문헌

- [1] R.Hjesvold and R.Midstraum, "Modeling and Querying sequence Data", In Proc of the 20th VLDB conference, Sandiago, Chile, Sep. 1994
- [2] P. Senthil Kumar, G. Phanendra Babu, "Intelligent multimedia data : data indices + interface", Multimedia systems, Vol 6, pp.395-407, 1998.
- [3] Moving Picture Experts Group, "Overview of MPEG-7 standara(Version 6.0)," ISO/IEC JTIC1/SC29/WG11 N4509, Pattaya, Thailand, December, 2001.
- [4] "Extensible Markup Language (XML) 1.0", World Wide Web Consortium Recommendation. (<http://www.w3.org/TR/REC-xml>)
- [5] J. Monaco, "How to Read a Film. The Art, Technology, Language, History and Theory of Film and Media", Oxford University Press, 1981.
- [6] 김채미, 최학열, 김삼석 공저, "전문가와 함께하는 XML Camp", 마이트 Press, 2001.
- [7] Duckett Jon, Ozu Nik, williams Kevin, Mohr Stephen, Stephen, Carl Jurt, Griffin Oliver, Norton Francis, Stokes Rees Ian, and Tennison Jeni, Professional XML Schemas, Wrox Pr Inc, 2001.
- [8] 남윤영, 황인준, "XMARS : XML 기반 멀티미디어 주석 및 검색 시스템", 정보처리학회 논문지 B 제9-Brnjs 제 5호, 2002.10.