

비디오 문서의 구조 정보를 이용한 메타데이터 모델링에 관한 연구

권재길*

A study on Metadata Modeling using Structure Information of Video Document

Jae Gil Kweon*

요약

비디오 정보는 전자도서관이나 WWW 및 주문형 비디오(VOD) 시스템과 같은 다양한 분야에서 중요한 요소로 부각되고 있으며, 시청각적(audio-visual), 시공간적(spatial-temporal), 의미적(semantics) 정보를 모두 포함하고 있어 사용자에게 다양한 형태의 정보를 제공할 수 있다. 또한 대부분의 경우 비디오 문서 전체를 검색하는 대신 비디오의 원하는 장면만을 검색할 수 있는 기능이 요구되고 있다. 따라서 본 논문에서는 이러한 검색의 다양성을 지원하기 위해 계층적으로 구성되어 있는 동영상 문서의 구조 정보를 이용하여 메타데이터를 모델링하고, 이를 통하여 동영상 문서를 통합적으로 다룰 수 있는 데이터베이스 스키마를 설계한다.

Abstract

Video information is an important component of multimedia system such as Digital Library, World-Wide Web(WWW) and Video-On-Demand(VOD) service system. It can support various types of information because of including audio-visual, spatial-temporal and semantics information.

In addition, it requires the ability of retrieving the specific scene of video instead of entire retrieval of video document. Therefore, so as to support a variety of retrieval, this paper models metadata using video document structure information that consists of hierarchical structure, and designs database schema that can manipulate video document.

* 창원전문대학 전자계산과 전임강사
논문접수: 98.11.11. 심사완료: 98.12.8.

I. 서론

최근 인터넷의 급속한 보급에 따라 이를 통해 교환되는 정보의 양이 증가하고 있으며 컴퓨터를 통해 전자화된 정보는 통신망을 이용함으로써 유통 및 재생산이 매우 용이해지고 있다. 또한 미디어(media)별 처리 기술의 발전, 데이터 입출력 장치와 저장 장치의 발달,

컴퓨터 처리속도의 향상, 통신 기술의 획기적 발전 등으로 컴퓨터를 이용한 멀티미디어(multimedia)의 실현 환경이 구축되고 있으며, 또한 새로운 다양한 미디어 서비스 요구가 증가되고 있다. 특히, OIS(Office Information System), 디지털 도서관(Digital Library), WWW(World-Wide-Web), VOD(Video On Demand) 등 멀티미디어 정보 처리를 요구하는 다양한 응용 분야의 출현으로 대량의 멀티미디어 정보를 효율적으로 저장하고, 처리하며, 검색할 수 있는 기능이 요구되고 있다. 그리고 비디오 정보는 방송, 교육, 출판과 같은 응용분야에서 점점 중요한 요소로 부각되고 있다. 비디오 정보에 대한 초기 연구는 비디오 데이터를 물리적인 세그먼트(segment), 즉 연속적인 프레임(frame) 주기로 분할하는 것이었으나, 최근에는 논리적인 세그먼트 단위로 구조화하려는 방법들이 연구[1, 2]되고 있다. 예를 들면, 비디오 데이터의 내용 자체를 비디오 분석 기술을 이용해 분석하기 보다는 여기에 텍스트나 자연어 형태로 된 설명문(description), 즉 주석(annotation)을 부착해서 이 설명문을 검색하는 방식이 많이 사용되고 있다.

디지털 도서관의 기능은 크게 정보 수집, 정보 관리, 정보 제공으로 나눌 수 있다. 정보 수집은 텍스트나 그래픽뿐만 아니라 소리, 동영상과 같은 멀티미디어 형태의 각종 정보를 수집하는 기능이다. 그리고 정보 관리는 분산되어 존재하는 정보를 분류하여 색인이나 목록과 같은 메타데이터를 이용하여 관리하는 기능이다. 마지막으로 정보 제공은 사용자에게 다양한 인터페이스를 제공하여 정보를 검색할 수 있게 하는 기능이다. 이 중 정보 관리 기능은 디지털도서관을 구축하는데 있어서 핵심적인 기능

으로 다양한 멀티미디어 데이터들의 복잡한 구조를 분석하여 메타데이터를 통합적으로 관리하고 사용자가 이러한 메타데이터를 이용하여 쉽고 체계적으로 원하는 정보에 접근할 수 있도록 하는 방안이 제시되어야 한다.

메타데이터란 “데이터에 대한 데이터” 라고 할 수 있다. 즉, 데이터 또는 데이터 집합을 효율적으로 접근하고 관리할 수 있도록 해주는 데이터에 대한 정보를 총칭한다. 메타데이터를 통한 전자 문서 관리는 정보 검색시 사용자가 보다 정확하고 쉽게 전자 문서를 검색할 수 있도록 하며, 대용량 정보를 취급하는데 유지비용을 절감시킬 수 있다. 특히, 멀티미디어 전자 문서의 메타데이터 모델링은 비정형 디지털 문서의 구조화를 가능케 하는 정형화된 포맷을 제공하기 때문에, 멀티미디어 전자 문서에 대한 메타데이터의 관리는 반드시 필요하다.

한편 동영상 데이터와 이에 대한 메타데이터를 데이터 베이스에서 통합하여 관리하면 많은 이득을 얻을 수 있다. 즉, 질의 처리 등 데이터베이스 시스템이 제공하는 유용한 기능들을 함께 제공받을 수 있다. 그런데 동영상 문서는 복잡한 구조를 가지며 모델링을 위해 상속성, 확장성 등을 요구하기 때문에 객체 지향 데이터베이스로 관리하는 것이 효과적이다. 이를 위해서는 메타데이터 엘리먼트 집합에 기반한 동영상 문서의 객체 지향 데이터 모델링이 필요하다. 따라서 본 논문에서는 이를 위해 메타데이터를 통합적으로 다룰 수 있도록 메타데이터를 모델링하고, 객체 지향 데이터베이스 스키마를 설계한다.

II. 관련 연구

디지털 도서관에서 다루어지는 데이터의 형태는 텍스트, 이미지, 동영상 등을 들 수 있으나 디지털 도서관의 구축의 핵심은 동영상 데이터라 볼 수 있다. 따라서 디지털 도서관에 있어서 동영상 문서를 효율적으로 접근하고 검색하기 위해서는 메타데이터 혹은 카탈로그 정보를 효율적으로 저장하고 관리하는 기능이 필수적이다. 비디오 정보에 대한 초기 연구는 비디오 데이터를 물리적인 세그먼트, 즉 연속적인 프레임 주기로 분할하는 것이었으나 최근에는 논리적인 세그먼트 단위로 구조화하려는 방법들

이 연구되고 있다. 예를 들면, 비디오 데이터의 내용 자체를 비디오 분석 기술을 이용해 분석하기 보다는 비디오 데이터를 논리적인 세그먼트로 구조화한 후 여기에 텍스트나 자연어 형태로 된 상세 설명을 부착해서 이 텍스트 형태의 설명 정보를 검색하는 방식을 사용한다. 이러한 방식을 비디오 데이터의 내용 기반 검색 혹은 주석 기반 검색이라 한다.

동영상 문서의 관리 시스템을 구축하기 위한 모델링 방법으로 E-R(Entity-Relationship)모델, 객체 지향 모델 등을 사용하고 있다. 대표적으로는 교육용 VOD 시스템으로 미국 버클리 대학에서 개발된 Plateau(3), 범용 데이터 시스템으로 노르웨이 공대에서 개발한 VideoSTAR(1), 동영상 데이터의 모델링과 검색을 위해 동영상 객체 개념을 도입한 일본 고베 대학의 OVID(4) 등이 있다. 이들 중 Plateau는 E-R 모델링을 사용하여 분산시스템 환경에서 멀티미디어 데이터베이스 시스템을 구축했으며 동영상 문서 관리를 위한 최소 단위는 프레임 시퀀스를 사용했다. 그러나 교육 분야에 적합하도록 모델링되었기 때문에 해당 응용분야에는 전문성을 띠고 있으나 디지털 도서관의 확장된 데이터를 다루기 위한 확장성이 부족하였다.

한편 OVID는 객체지향 모델링을 사용했으며 적용 분야는 NOD(news on demand)였다. 이는 동영상 객체들을 시스템의 기본 단위로 사용하여 스키마가 없는 동적 스키마 구조와 속성값을 가짐으로써 사용자가 원하는 정보를 원하는 시점에서 속성들 정의함으로써 추가 할 수 있다. 그러나 동영상 데이터베이스가 확장 될 수록 스키마가 없는 메타데이터의 관리가 복잡해진다. 또한 사전에 질의에 사용되는 속성들을 계층 구조로 구축해 두어야 하기 때문에 데이터베이스를 확장하고자 할 때 다른 질의를 적용하기 어려워 확장이 어렵고 문서 관리의 단위가 동영상 객체이기 때문에 문서에 대한 접근점을 다양화 할 수 없다.

OVID의 비디오 데이터 모델은 비디오 문서 구조의 모델을 명백하게 지원하지 못했다. 또한 OVID는 사용자에게 정의된 속성값으로 비디오 객체를 검색하게 하는 기존의 SQL에 기초한 질의어인 Video SQL을 제공하지만, 비디오 객체들 사이에 시간적인 관계성을 기술하기 위한 표현은 제공하지 못한다.

VideoSTAR는 E-R 모델링을 사용했으며 주된 적용 분야는 VOD였다. 이는 범용 동영상 데이터 관리 시스템을 위해 개발된 프로토타입 시스템으로 일반적인 메타데

이터를 사용했다. 따라서 동영상 정보의 공유와 재사용이 용이하게 설계되었으며 메타데이터 모델링을 사용하는 분야에 범용으로 사용될 수 있다. 그러나 모델이 복잡하고 의존적인 부분에 대해서는 고려하지 않았기 때문에 이 모델을 이용하고자 하는 시스템에서는 응용에 의존적인 부분을 다시 수정해야 하는 단점이 있다. 그러나 문서 관리의 최소 단위가 프레임 시퀀스이기 때문에 사용자가 문서를 다양한 형태로 접근할 수 있다.

[5]의 연구에서는 비디오 데이터의 계층적인 구조를 기반으로 한 메타데이터 모델을 제안하고 있다. 특히, 이 연구에서는 비디오 데이터의 특성을 7가지 관점에서 추출하고 메타데이터를 크게 내용 기반 정보와 내용 독립 정보로 분류한 후, 각각을 더 세분화해서 다양하게 분류하고 있다. 이 연구에서 제안된 메타데이터 모델은 다양하게 분류된 메타데이터와 비디오 데이터의 논리적인 계층 구조를 지원할 수 있다. 하지만 비디오 데이터와 메타데이터 모델에 대한 색인 구조에 대한 언급이 없으며, 각 계층 구조간의 효율적인 접근 방법을 제공하지 못한다.

[6]의 연구는 이미지 분석 기법과 비디오 스트림을 샷의 집합으로 분할하는 기법을 이용한 세그멘테이션 기반 모델로서 뉴스 비디오 프로그램을 도메인으로 사용하여 각 이미지 프레임의 공간적 구조와 전체프로그램의 시간적 구조를 기반으로 한 모델을 제공한다. 또한 내용기반 비디오 색인 방법을 제안하였다. 하지만 비디오 데이터에 대한 체계적인 구조가 없으며 본 논문에서와 같이 비디오 문서에 대한 다양한 레벨의 검색을 지원하지 못한다.

III. 동영상 문서의 구조 정보 모델링

동영상 자체는 시청각적(audio-visual), 시공간적(spatial-temporal), 의미적(semantic) 정보를 모두 포함하고 있어 사용자에게 다양한 형태의 정보를 제공할 수 있다. 또한 비디오 정보 자체는 원시 데이터이지만 대부분의 경우 비디오 문서 전체를 검색하는 대신 비디오의 각 원하는 장면만을 검색할 수 있는 기능을 요구하고 있다. 그런데 기존의 연구들은 비디오 데이터의 구조 정보에 대한 메타데이터 모델 및 질의어를 충분히 제공하지

못하고 있다. 예를 들어 영화 비디오 문서의 경우 “해뜨는 장면”을 찾고자 할 때 기존의 물리적인 색인 방법만으로는 이러한 질의를 검색할 수 없다. 즉, 이와 같은 질의의 결과를 얻기 위해서는 우선적으로 비디오 데이터의 계층적인 구조 정보가 있어야 한다. 따라서 이 장에서는 매우 큰 용량하면서도 계층적 구조를 가지고 있는 동영상 구조를 고찰하고, 이러한 구조 정보를 효과적으로 관리하여 문서 전체가 아닌 작은 단위의 검색을 가능하게 하여 사용자가 원하는 특정 영역에 바로 접근할 수 있게 해줌으로써 검색의 정확성 및 다양성을 증대시키기 위한 동영상 구조 정보 모델을 설계한다.

3.1 계층적인 동영상 구조 및 메타데이터 모델링

일반적으로 문서는 목시적이든 명시적이든 다양한 구조 정보를 가지고 있다. 이러한 구조정보가 검색되는 관점에서는 대개 무시되어져 왔는데 마크업 언어가 많이 이용되면서 문서구조에 의한 검색 및 활용이 높아지고 있다. 이러한 구조 정보를 이용한 내용기반의 검색은 문서 전체가 아닌 작은 단위의 검색을 가능하게 하여 사용자가 원하는 특정 영역에 바로 접근할 수 있게 해줌으로써 검색의 정확성 및 다양성을 증대시킬 수 있다. 따라서 동영상 문서 역시 계층적으로 구성되어 있어 이 구조를 분석하여 관리하면 동영상에 접근하는 접근점을 다양화시킬 수 있다.

동영상 문서는 구현 방법에 따라 여러 계층으로 나눌 수 있으나 본 논문에서는 구조적 성분을 프레임 시퀀스(frame sequence), 장면(scene), 샷(shot), 프레임(frame)으로 구성되는 것으로 가정하며, 이들은 서로 계층적인 관계성을 가진다. 그림1은 동영상 구조의 메타데이터 모델을 개체지향 개발 방법론인 OMT를 확장한 XOMT[11]로 나타내었으며, ◇은 part-of 관계성을, —●은 일대다의 관계를 나타낸다. 프레임은 하나의 정지 영상에 해당하며, 이들은 하나 혹은 여러 개가 모여 샷을, 그리고 이들은 하나의 장면을 구성한다. 예를 들어 뉴스 프로그램에서 각 뉴스 기사와 스포츠 소식, 일기 예보 등은 각각 하나의 장면으로 볼 수 있으며, 일기 예보를 전하는 장면에서 호수 사진과 구름 사진 등은 하나의 샷이 된다. 또한 프레임 시퀀스는 하나 혹은 여러개의 장면으로 구성되며, 이들이 모여 하나의 동영상을 구성한

다. 따라서 이장에서는 이러한 구조적 성분의 최소단위인 프레임 중 키 프레임을 추출한 후 캡션 및 그림 설명 형태의 주석을 부여하고, 키 프레임이 가지는 이미지 타입, 프레임 번호, 시작 주소와 끝 주소 등의 물리적인 특성을 정의한다. 그리고 이들을 메타데이터로 활용하여 텍스트 형태로 데이터베이스에 저장하고, 관리하여 이를 통한 검색을 지원하는 메타데이터 모델링을 제안한다.

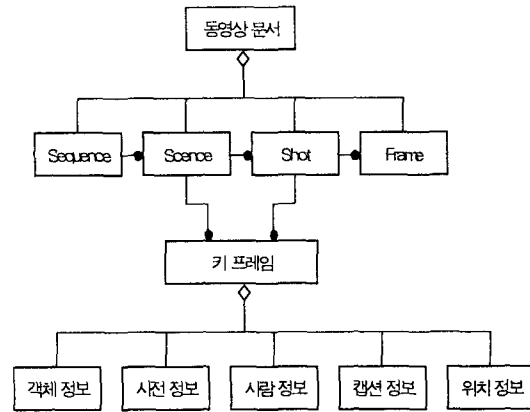


그림 1. 메타데이터 모델링
Fig 1. metadata modeling

3.2 MPEG-2의 구조

일반적으로 멀티미디어 DBMS를 이용하는 동영상 정보의 관리는 검색 대상인 비트맵이나 웨이브 패턴을 압축하지 않은 상태로 관리하는 방식이다. 그런데 동영상의 특성상 압축하지 않은 자연 영상을 그대로 저장하고 검색하거나 전송하기에는 많은 어려움이 따른다. 따라서 동영상 정보관리 시스템에서 이러한 문제들을 해결하고 실용화하기 위해서는 동영상 정보를 압축하여 저장하고, 압축된 상태에서 실시간으로 검색하여 전송할 수 있는 기술이 요구되며, 정보 검색의 다양성과 정확성을 제공하기 위하여 압축된 동영상 문서의 계층적인 구조를 분석하여 각 구조정보에 대한 메타데이터를 구축하는 기법이 요구된다. 동영상을 압축하는 기법에는 여러 가지가 있으나 본 논문에서는 현재 널리 사용되고 압축의 표준으로 자리잡아 가고 있는 MPEG-2를 활용한다. 그리고 MPEG-2의 계층적인 구조를 분석하여 메타데이터를 설계하기 위하여 MPEG-2에서 동영상 문서의 압축 원리와 구성을 알아야 한다.

동영상의 압축 기술에는 크게 공간적 압축 기술과 시

간적 압축기술이 있다. 공간적 압축기술은 한 화면에서의 공간적 중복성을 줄이기 위한 압축 기술이고, 시간적 압축 기술은 연속 화면간의 시간적인 중복성을 제거하기 위한 제거하기 위한 압축 기술이다. 실제로 동영상에 있어서 공간적인 중복성 보다는 시간적인 중복성이 훨씬 많으므로 시간적인 중복성을 어떻게 찾아내는 문제가 고효율로 압축하는 척도가 된다. 이러한 시간적 압축기술로 움직임 추정을 이용하는데 이는 주위 영상들의 움직임을 추정해서 부호화하는 신호를 줄임으로써 압축하는 방식이다. 즉, 영상이 변화없는 부분은 그대로 두고 움직임이 있는 부분만을 부호화하여 다음 장면을 보여주는 방법으로 앞 프레임의 현재 프레임의 예측에 사용하며 예측된 정보와 함께 예측오차와 함께 정보를 인코딩하여 보여준다. 움직임 추정은 사용되는 참조 픽처에 따라 앞 프레임을 참조하는 순방향 예측과 뒤의 프레임을 참조하는 역방향 예측이 있다. MPEG에서는 세가지의 픽처 형식을 지원하는데 움직임 추정을 사용하지 않는 I-픽처와 순방향 예측만을 사용하는 P-픽처 그리고 순방향 예측과 역방향 예측을 모두 사용하는 B-픽처가 있다. MPEG-2에서는 프레임 픽처뿐만 아니라 필드 픽처까지 지원함으로써 압축의 효율성을 높였으며, 프레임 단위로 압축할 경우 픽처란 말 대신 프레임이란 말을 쓰기도 하는데 여기서는 프레임으로 통일 하였으며, 3가지 종류의 프레임을 가진다.

프레임

시간적인 압축기술은 사용하지 않고 공간적인 압축 기술만을 사용해서 압축한 프레임으로 다른 프레임의 참조 프레임 역할을 하며 시간적인 정보를 사용하지 않기 때문에 다른 프레임들 보다 많은 비트수를 필요로 한다. 그러나 시간적인 정보를 사용하지 않기 때문에 단독으로 부호화가 가능하고 임의의 접근을 할 수 있는 프레임이 될 수 있다. 따라서 캡션 정보 및 설명 정보를 이용한 메타데이터를 구축하여 동영상 구조 정보를 이용한 내용기반의 검색이 가능하도록 하기 위해 여기에서는 I-프레임을 사용한다.

P-프레임

가장 최근의 I-프레임이나 P-프레임을 참조 프레임으로 움직임 추정 기술을 사용해서 부호화한 프레임이다. 즉, 순방향 예측만을 사용해서 다음 화상을 추정하는 프

임으로 P-프레임은 움직임 추정을 사용했기 때문에 I-프레임보다 비트 발생수가 적다.

B-프레임

순방향 예측뿐 아니라 역방향 예측까지 포함해서 보다 더 세밀한 예측을 한 프레임이다. 때문에 일반적으로 I-프레임이나 P-프레임보다 발생 비트수가 적은 것이 특징이다. 하지만 역방향 예측에서는 앞으로 발생할 화상을 참조 하기 때문에 프레임 재 순서화가 필요하고 이로 인한 지연이 발생할 수 있다.

3.3 키 프레임의 색인 구조

MPEG-2 동영상 압축 파일은 앞에서 본 바와 같이 크게 3종류, 즉 I-프레임, P-프레임, B-프레임으로 구성된다. 따라서 본 논문에서는 MPEG-2로 압축된 동영상 문서 파일내의 모든 I-프레임들을 키 프레임의 후보가 될 수 있다고 가정한다. 이 가정에 따라 압축된 동영상에서 I-프레임을 추출하여 사용자가 직접 키 프레임을 선택할 수 있는 방법을 제공한다. 또한 동영상에 대한 검색은 실제 프레임으로 검색하는 것이 아니라 각 키 프레임에 대한 캡션 정보 및 물리적인 정보를 작성하여 이를 키 프레임과 함께 구조화하여 데이터베이스에 저장한 다음 이를 메타데이터로 활용하여 검색을 수행한다.

[그림 2]에서 보는 바와 같이 첫 번째 단계는 MPEG-2파일 스트림에서 키 프레임의 후보가 될 수 있는 I-프레임을 추출하여 I-프레임 테이블을 구축하는 것이다. 두 번째 단계는 I-프레임 테이블에서 편집자가 직접 키 프레임이 될 수 있는 I-프레임을 선정하여 그림 설명 정보 및 프레임의 물리적인 속성값을 메타데이터로 입력하면 키 프레임에 대한 jpeg, bmp, gif 등의 이미지 타입을 값으로 가지는 ImageType, 프레임의 번호를 나타내는 FrameNo, 축소 그림을 나타내는 ImageIcon, 크기를 나타내는 Height 및 Width, 키 프레임이 속한 클립의 시작 주소와 끝 주소를 나타내는 StartLoc, EndLoc, 동영상 파일의 시작으로부터 키 프레임까지의 가지 거리를 나타내는 Offset 그리고 그림 설명 정보를 가지는 그림 2와 같은 인덱스 테이블을 구성한다.

이 방법은 자동 인덱싱 기법들 보다는 효율성이 뒤떨어지지만 동영상에서 자동으로 키 프레임을 추출한 후 영

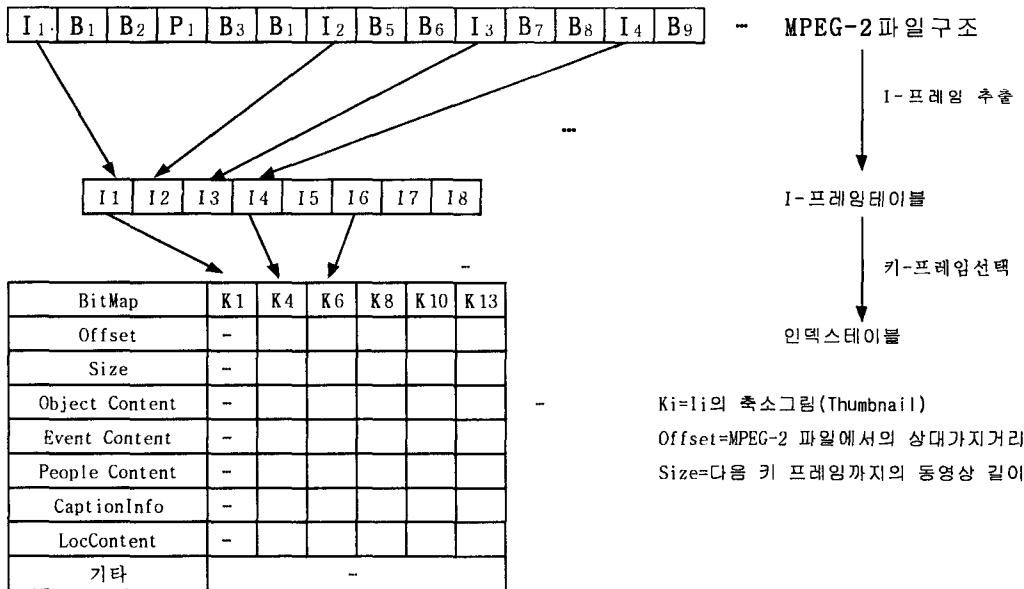


그림 2. 키 프레임의 색인 구조
Fig 2. Index structure of key frame

상 인식을 통해 키 프레임의 내용을 처리하는 기술은 아직 완전하게 구현하기 어렵기 때문에 본 논문에서는 중간 과정으로 이 키 프레임에 대한 상세 정보 및 물리적 정보를 메타데이터로 이용하여 검색하는 방법을 사용한다.

3.4 동영상 구조 정보를 위한 SGML DTD

위에서 정의한 동영상 문서 메타데이터가 될 수 있는 프레임의 상세 정보와 물리적 정보의 상위 엘리먼트를 KeyFrame으로 하고, 내용 모델을 키프레임의 물리적 정보를 나타내는 FrameInfo 및 상세정보를 나타내는 FrameDesc로 정의 하였다. 그리고 각각이 가지는 세부적인 속성 및 하위 엘리먼트로 정의하여 그림 3과 같이 SGML DTD를 설계했다

[그림 3]에서 엘리먼트 KeyFrame은 키 프레임에 대한 물리적인 정보를 가지는 하위 엘리먼트 FrameInfo와 프레임에 대한 상세 정보를 가지는 하위 엘리먼트 FrameDes로 구성된다. FrameInfo의 속성값은 이미지 파일의 파일 타입을 나타내는 ImageType, 키 프레임의 번호를 나타내는 FrameNo, Thumbnail 정보를 나타내는 ImageIcon, MPEG-2 파일의 시작점으로부터 키 프레임까지의 상대가치거리를 나타내는 Offset, 키 프레임

에서 다음 키 프레임까지의 동영상 길이를 나타내는 Size, 그리고 키 프레임을 클릭 했을 때 사용자에게 보여 줄 파일의 시작 위치와 끝 위치를 나타내는 StartLoc 와 EndLoc으로 구성된다.

키 프레임의 상세 정보를 기술하는 FrameDes는 다섯 개의 하위 엘리먼트를 가지며, 이들은 키 프레임의 상세 정보를 기술하기 위한 속성값을 가진다. ObjectContent는 키 프레임에 나타나는 객체들을 기술하는데 사용되는 엘리먼트이고, EventContent는 키 프레임의 상황을 설명하는데 사용되며, PeopleContent는 프레임에 사람이 나타나는 경우 사람에 대한 상세 설명을 하는데 사용되는 엘리먼트이다. 그리고 CaptionInfo는 촬영 날짜, 감독 등의 정보를 기술하는 엘리먼트이며, LocContent 장면을 찍은 장소 및 구체적인 국가를 기술하는 엘리먼트로 사용되며 이들 메타데이터를 활용하여 동영상 문서를 프레임별 혹은 클립별로 다양한 검색을 하게 해 준다.

3.5 동영상 구조 메타데이터 객체 다이어그램

[그림 3]의 SGML DTD에 기초하여 XOMT를 이용하여 동영상 구조 메타데이터의 객체 다이어그램을 보인다.

```

<!DOCTYPE VAIM[
<ELEMENT KeyFrame -- (FrameInfo, FrameDes*) >
<ELEMENT FrameInfo -- EMPTY >
<ATTLIST FrameInfo
ImageType CDATA #REQUIRED
FrameNo CDATA #REQUIRED
ImageIcon CDATA #REQUIRED
Height CDATA #REQUIRED
Width CDATA #REQUIRED
StartLoc CDATA #REQUIRED
EndLoc CDATA #REQUIRED >
<ELEMENT FrameDes -- (ObjectContent, EventContent,
PeopleContent, CaptionContent, LocContent) >
<ELEMENT ObjectContent -- (Type, Notation) >
<ELEMENT Type (#PCDATA) >
<ELEMENT Notation (#PCDATA) >
<ELEMENT EventContent -- (Time, EventDesc, EventType) >
<ELEMENT Time (#PCDATA) >
<ELEMENT EventDesc (#PCDATA) >
<ELEMENT EventType (#PCDATA) >
<ELEMENT PeopleContent -- (Name, Birthday, Country, Job) >
<ELEMENT Name (#PCDATA) >
<ELEMENT Birthday (#PCDATA) >
<ELEMENT Country (#PCDATA) >
<ELEMENT Job (#PCDATA) >
<ELEMENT CaptionInfo -- (NewsId, Date, Director, FileLoc) >
<ELEMENT NewsId (#PCDATA) >
<ELEMENT Date (#PCDATA) >
<ELEMENT Director (#PCDATA) >
<ELEMENT FileLoc (#PCDATA) >
<ELEMENT LocContent -- (Nation, City, CityArea) >
<ELEMENT Nation (#PCDATA) >
<ELEMENT City (#PCDATA) >
<ELEMENT CityArea (#PCDATA) >
]
    
```

그림 3. SGML DTD
Fig 3. SGML DTD

IV. 동영상 문서의 메타데이터를 위한 데이터베이스 스키마 설계

디지털 도서관을 위한 동영상 문서는 데이터베이스에서 체계적으로 관리할 필요가 있다. 그런데, 동영상 정보에 대한 다양한 사용자의 요구를 수용하기 위해서는 객체지향 데이터베이스를 사용하면 효과적이다. 이럴 경우 질의 처리 등 데이터베이스 시스템이 기본적으로 제공하는 기능뿐만 아니라, 확장성(extendable), 상속성(inheritance), 구성 요소의 반복 사용(reusability), 그리고 선택적 사용(optional) 등을 이용할 수 있다.

따라서 이 장에서는 XOMT도를 기초로, 동영상 문서의 메타데이터를 위한 데이터베이스 스키마 설계한다. 그림 5는 동영상 문서의 메타데이터를 위한 데이터베이스 스키마이다. 이는 객체 모델링으로 기술되었는데, 클래스는 클래스명과 속성을 갖는 네모, 클래스들간의 상속 관계(is-a)는 작은 세모, 부품 관계(part-of)는 마름모로 표현하였다.

본 논문에서의 데이터베이스 스키마는 동영상 문서의 메타데이터를 위한 설계로서 KeyFrame 클래스를 상위 클래스로 두고, 각각 16개로 나뉘어진 메타데이터 엘리

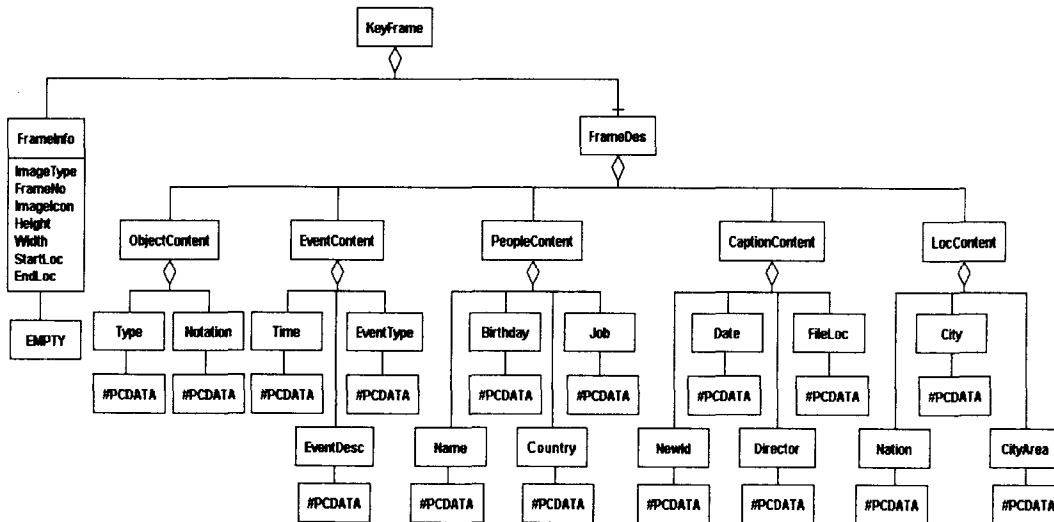


그림 4. 동영상구조 메타데이터 객체 다이어그램
Fig 4. metadata Object diagram of Video structure

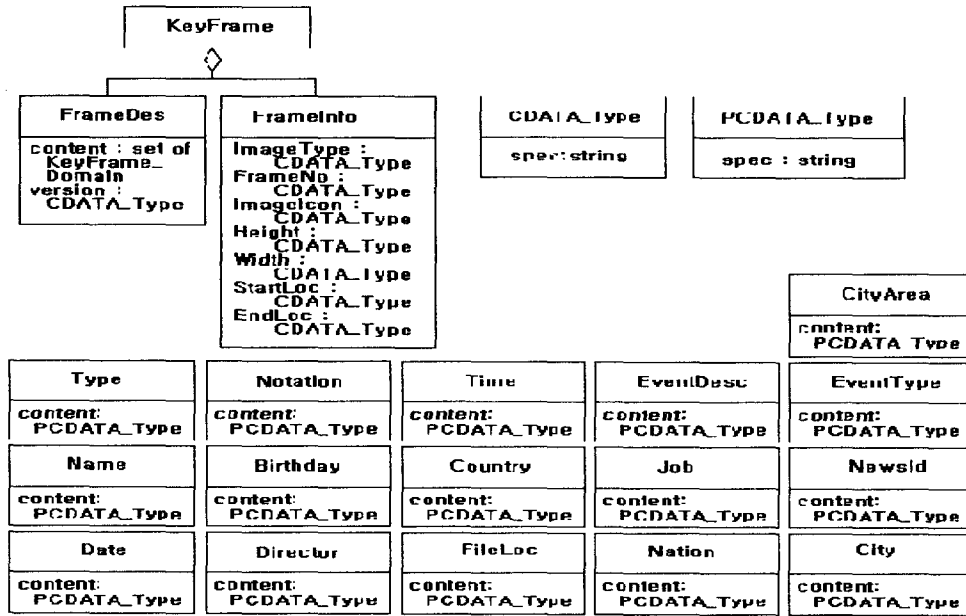


그림 5. 데이터베이스 스키마
Fig 5. Database Schema

먼트를 각각 독립된 하위 클래스로 설계함으로써 동영상 메타데이터는 필요한 정보를 상위 클래스로부터 상속받는 기준으로 설계하였다. 이는 메타데이터를 보다 확장시키고자 할 때와 동영상 문서 저장을 하고자할 때 본 설계를 사용하고자 함이다.

여기서 KeyFrame 클래스는 16개의 엘리먼트 메타데이터를 표현할 수 있는 FrameDes과 동영상의 매체 정보를 가지고 있는 FrameInfo 클래스를 하위 구조로 가지며, FrameDes는 set of KeyFrame_Domain로 [그림

6]과 같은 도메인 클래스 타입 계층으로 연결시켜 16개의 메타데이터 엘리먼트를 클래스로 가진다.

[그림 6]은 동영상 문서의 메타데이터를 위한 데이터베이스 스키마의 속성 도메인 클래스 타입 계층이다.

V. 결론 및 향후 연구과제

비디오 문서의 논리적인 계층구조의 특성은 비정형적

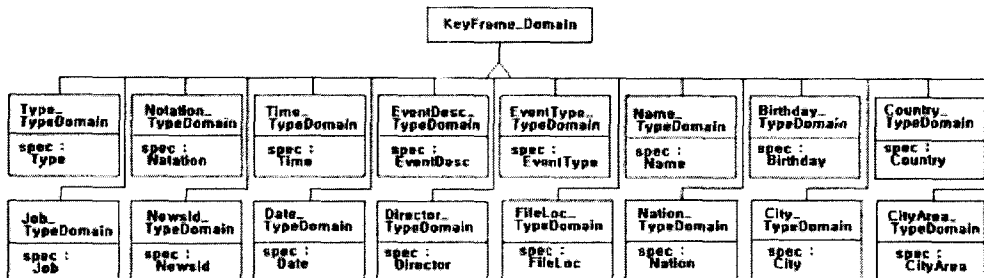


그림 6. 속성 도메인 클래스 타입계층
Fig 6. attribute domain class type hierarchical

인 비디오 문서에서 사용자가 원하는 요소를 추출하는데 매우 유용하다. 본 논문에서는 이와 같은 특성을 지원하기 위해 비디오 문서의 구조적인 요소를 이용하여 메타데이터를 모델링하였다.

본 논문에서 제안한 메타데이터 모델은 디지털 도서관에서 나타나는 동영상 문서의 일반적인 특징들을 모델링하였다. 이는 디지털 도서관에서 일반적으로 등장하는 동영상 문서를 데이터베이스 시스템을 사용하여 효과적으로 관리할 수 있게 하며, 동영상에 대한 메타데이터와 동영상의 구조 정보를 통합적으로 관리할 수 있도록 하는 프레임워크를 제공한다.

본 연구에서는 SGML DTD로 기술되는 동영상 문서의 관리를 위해 객체 지향 데이터베이스를 이용하였는데, 이로써 메타데이터의 확장성, 수정성, 선택성, 반복성을 지원할 수 있었다. 그러나 설계된 데이터베이스 스키마는 실제 응용에 사용하기 위해서는 아직도 보완할 점이 남아 있다. 시스템의 구현과 함께 제안한 스키마를 확장하는 것으로 새로운 메타데이터가 필요할 경우에 이를 위한 확장이 필요하다. 또한 최근에 시간 관계성을 지원하는 HyTime[12]이 제안되었는데, 보다 완전한 하이퍼미디어 응용을 구축하기 위해서는 이에 대한 지원도 요구된다.

참고문헌

- [1] R. Hjelqvold and R. Midstraum, "Modeling and Querying Video Data," In Proc. of the 20th VLDB Conference, Santiago, Chile, pp. 686-694, Sep. 1994.
- [2] Ramesh Jain and Arun Hampapure, "Metadata in Video Databases," SIGMOD RECORD, Vol. 23, No.4, pp. 27-33, 1994
- [3] L.A. Rowe, J.s. Boreczky, and C.A. Eads, Indexs for User Access to Large Video Database, In Proc. of the IS&T/SPIE Symposium on Electronic Imaging Science and Technology, Conf. on Storage and Retrieval for Image Video Databases II, San Jose, CA, Feb. 1994
- [4] E. Oomoto and K. Tanaka, "OVID: Design and Implementation of a Video-Object Database System," IEEE Transaction on Knowledge and Data Engineering, Vol. 5, No. 4, pp. 629-643, April 1994.
- [5] Park, Y., Yongkeol Kim, Seongil Jin and Wan Choi, "Hierarchical Structure-based Metadata Model for Video Database Application," the ISCA 13th international conference Computer And Their Applications'98, pp.242-245, 1998.
- [6] HongJiang Zhang, et al., "Automatic parsing and indexing of news video," IEEE Multimedia Systems, pp. 256-266, 1995.
- [7] F. Kappe, "Aspects of the Modern Multimedia Information System," Ph. D. Dissertation, Graze University, Austria, 1991.
- [8] S. Weibel, J. Godby, and E. Miller, "OCLC/NCSA Metadata Workshop Report," Office of Research, OCLC Online Computer Library Center Inc., 1995.
- [9] L. Burnard, E. Miller, L. Quin, and C. Sperberg-McQueen, "A Syntax for Dublin Core Metadata," Office of Research, OCLC Online Computer Library Center Inc., 1996.
- [10] E. Herwijnen, Practical SGML(2nd Ed.), Kluwer Academic Publishers, 1994.
- [11] 강현석, 배종민, 박인호, 한애노, 장원호, 임혜정, 중간보고서 - 객체지향 데이터베이스 설계도구 개발, 한국전자통신연구소, 1996. 6.
- [12] ISO, Hypermedia/Time-based Structure Language : HyTime(ISO 10744), 1992.

권재길

한국 OA학회 논문 제2권 제3호참조