

주석 및 내용 기반 검색을 지원하는 동영상 정보 관리 시스템의 개발†

전미경[○], 허진용*, 김인홍*, 강현석*
*경상대학교 컴퓨터과학과/정보통신 연구센터

Development of A Video Information Management System for Supporting Caption- and Content-based Searches†

Mi-Gyung Jun[○], Jin-Yong Heo*, In-Hong Kim*, Hyun-Syug Kang*
*Dept. of Computer Science/Information and Communication Research Center
Gyeong-Sang National University

요 약

본 논문에서는 동영상 정보의 효율적인 관리를 위해 주석 기반 검색과 내용 기반 검색을 통합적으로 지원하는 통합 동영상 데이터 모델(Integrated Video Data Model, IVDM)을 제안한다. IVDM은 동영상 자료를 계층적으로 구조화하여 상위 수준에서는 의미 단위와 세그먼트 단위로 주석 기반 검색을 지원하고, 하위 수준에서는 이미지 인덱싱을 이용한 내용 기반 검색을 지원한다. 우리는 이 IVDM을 이용하여 MPEG-2로 압축된 동영상 정보를 관리하는 시스템(Integrated Video Information Management System, IVIMS)을 개발한다.

1. 소개

최근 멀티미디어와 인터넷 기술의 발달로 단순 텍스트 형식만이 아닌 멀티미디어 정보, 특히 동영상 정보의 관리에 대한 수요가 급증하고 있다. 이에 따라 동영상 자료를 효율적으로 저장하고 필요에 따라 빠르게 검색할 수 있는 시스템에 대한 연구가 많이 이루어지고 있다[1, 2, 3, 4]. 그런데, 이러한 연구에서 동영상 정보를 압축하지 않은 자연 상태로 관리한다면 동영상의 특성상 저장과 전송에 많은 어려움이 따른다.

한편 이러한 연구들은 주로 동영상 자료를 검색하기 위해 2가지 접근법중 하나를 사용한다[1, 3, 4]. 즉, 내용 기반 검색과 주석 기반 검색으로, 기존의 연구들은 이중 선택하여 사용하는 경향이 있다. 그

런데 이 두 접근법이 갖는 각자의 장점을 함께 수용하고 단점은 상호 보완하여 사용한다면 효과적인 것이다.

본 연구에서는 방대한 양의 동영상 자료를 효율적인 방법으로 저장하고 검색할 수 있도록 주석 기반 검색과 내용 기반 검색을 함께 사용할 수 있는 통합 동영상 데이터 모델(IVDM)을 제시하고, 이 모델을 기반으로 MPEG-2로 압축된 동영상 정보를 객체 관계형 데이터베이스에서 관리하는 시스템(IVIMS)을 개발한다.

본 논문의 구성은 다음과 같다. 2장에서는 동영상 검색 기법에 관련된 연구를 살펴보고, 3장에서는 주석 기반 검색과 내용 기반 검색을 동시에 가능하게 하는 통합 동영상 데이터 모델을 개발한다. 4장에서는 3장에서 제시한 통합 동영상 데이터 모델을 이용하여 동영상 정보 관리 시스템을 설계하고, 5장에서는 구현 결과를 보여준다. 마지막으로 6장에서는 결론과 향후 연구 과제를 논한다.

† 본 연구는 경상대 부설 정보통신 연구센터의 연구비 지원에 의한 것입니다.

2. 관련 연구

동영상 자료를 저장하고 검색하는 방법에는 크게 내용을 기반으로 하는 검색과 주석을 기반으로 하는 검색으로 나눌 수 있다.

내용 기반 검색은 동영상 자료가 가지고 있는 의미를 동영상 각각의 프레임에서 색깔, 모양, 움직임 등을 추출하고, 이를 바탕으로 검색하는 방법이다[5, 6]. 이러한 방법은 특정 정의역에 대해서는 좋은 검색 결과를 나타내지만 동영상 자료가 담고 있는 일반적인 의미를 추출하기가 어렵고, 압축된 동영상일 경우 이미지 추출을 위해 압축을 풀고 검색을 해야 하기 때문에 성능면에서 비효율적이다.

주석 기반 검색은 동영상 자료가 담고 있는 의미를 사람이 먼저 파악한 후 이를 자연어를 이용하여 표현하고, 이를 바탕으로 검색하는 방법이다[1, 3]. 이 방법은 자동화된 방법으로 알아내기 힘든 동영상 자료의 여러 가지 의미를 쉽게 모델링할 수 있게 하고 이를 검색에 이용할 수 있게 한다. 반면에 사용자의 관점에 따라 주석을 다르게 부여하거나 다르게 해석할 수 있으므로 동영상 자료에 대한 일관성을 잃기 쉽다. 특히 포괄적인 내용의 주석이 아닌 아주 세분화된 부분까지 주석을 부여하고자할 때는 더욱더 일관성을 유지하기가 어려워진다.

따라서 이러한 두 기법을 통합적으로 사용할 수 있는 방안의 모색이 필요하다. 이때 사용자의 요구에 적합한 형태로 두가지 검색 기법을 통합적으로 지원하기 위해서는 무엇보다 통합 데이터 모델의 개발이 필요하다.

최근, 이 두 기법을 통합하는 형태의 검색을 위해 김기병 등[2]은 다중 계층 비디오 모델(Multi-Layered Video Model; MuVi)를 제안하였다. MuVi 모델은 각 계층별로 독립성을 유지하여 내용 기반이나 주석 기반 검색의 특정 방법에 종속되지 않는 질의 처리기를 구현한 것으로 비디오 데이터의 검색에 관련된 모델을 제시한 것이다. 그러나 본 논문은 동영상을 관리할 수 있는 통합적인 모델을 제시하는 것으로, 검색에서는 MuVi 모델의 일부를 수용하며, 단계별로 종속성을 가지지만 좀 더 사용자의 요구에 근접하기 위한 방법을 제시한다.

한편 본 연구팀에서는 최근 동영상 문서의 효율적인 관리를 위해 일반적인 동영상 데이터 모델을 제안하고, 이를 기초로 주석 기반 검색만을 지원하는 MPEG-2 압축 동영상 문서 관리 시스템의 개발에

대한 연구를 수행한 바 있다[1].

그러나 본 논문에서는 주석 기반 검색과 내용 기반 검색을 동시에 지원하고 관리하는 통합 동영상 데이터 모델을 제시하고, 이 모델을 기반으로 클라이언트/서버 환경에서 객체 관계형 데이터베이스를 이용하여 압축 동영상 정보를 관리하는 시스템을 개발한다. 이때 웹에서의 사용을 위해 플러그인 기법도 이용한다.

3. 통합 동영상 데이터 모델(IVDM)

동영상을 효율적으로 검색하고 관리하기 위해서는 압축된 동영상 자체, 그에 관련된 주석, 그리고 이미지 분석 결과를 통합적으로 데이터베이스화하여 공유할 필요가 있다. 이를 위해서는 일반적인 표준 모델을 만들어 다양하고 방대한 양의 압축 동영상을 관리하는 것이 요구된다. 따라서 이 장에서는 동영상 정보의 관리를 위한 통합 동영상 데이터 모델(Integrated Video Data Model, IVDM)을 개발한다.

이러한 IVDM의 의의는 동영상 데이터를 구조화함으로써 다양한 동영상 데이터에 대해 상위 수준에서는 자유로운 주석 기반 검색을 지원하고 하위 수준에서는 용이한 내용 기반의 검색을 지원한다는 것이다. 그리고 이러한 작업으로 동영상 정보를 공유하여 사용할 수 있게 한다는 것이다.

IVDM은 동영상의 계층 구조와 인덱싱 구조로 구성된다.

3.1 동영상의 계층 구조

동영상은 저장 동영상 세그먼트라고 하는 프레임의 연속적인 그룹으로 동영상 데이터베이스에 저장된다. 따라서 동영상은 하나 혹은 그 이상의 저장 동영상 세그먼트로 사상되는 동영상 스트림에 의해 표현된다.

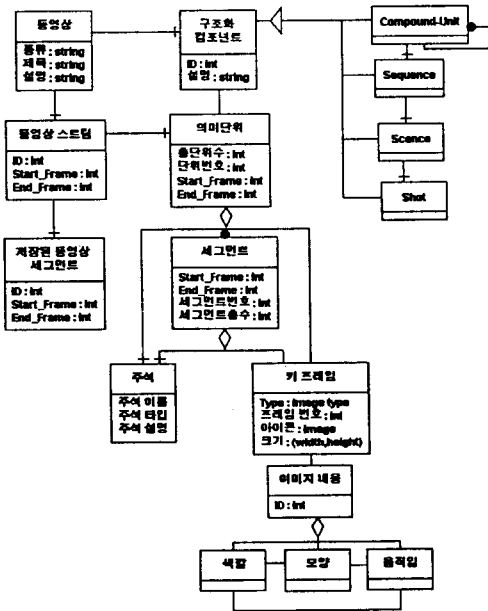
우리의 IVDM은 동영상 문서의 의미 단위와 연관된 구조적 구성 성분의 개념으로 만들어졌다. 구조적 구성 성분의 개념은 복합 단위(compound unit), 시퀀스(sequence), 장면(scene)과 샷(shot)으로 세분화되며, 이들 부클래스들은 서로 계층적인 관계로 정의된다. 샷은 하나 혹은 그 이상의 연속된 프레임으로 구성되고, 시간과 공간적으로 연속된 행동(action)으로 나타난다. 장면은 여러 개의 샷이 모여서 이루어지며, 시퀀스는 이러한 장면들이 모여서

이루어진다. 관련된 시퀀스들의 모임이 다시 복합 단위를 구성하게 되며, 복합 단위는 임의의 레벨로 자기 자신을 참조할 수 있다.

동영상 검색 구조는 주석 기반의 검색을 지원하는 단계와 내용 기반 검색을 지원하는 두가지 단계로 나누어진다.

주석 기반 검색을 지원하는 단계에서는 동영상을 종류별로 분류해서 한 종류 내에서도 의미 단위를 기준으로 나눈다. 예를 들면, 종류가 드라마일 경우에는 매일 방송 회분으로 나누고, 교육용 자료일 경우에는 단원별로 나누어 주석을 부여한다. 이것은 구체적인 검색이 아닌 주제별이나 객관적으로 드러나는 사실을 기준으로 검색을 하고자할 때 사용될 수 있는 범위이다. 이 의미 단위는 동영상의 종류마다 크기가 다르고, 같은 동영상이라도 크기가 다를 수 있다. 이 의미 단위에 주석이 부여되며 다시 세그먼트 단위로 나누어진다. 세그먼트 단위에서도 적절한 주석을 부여할 수 있다.

내용 기반 검색을 위한 단계는 세그먼트 단위에서 특정 키프레임을 추출해내어 이미지 분석을 적용시킨다. 그래서 색깔, 모양, 움직임 등의 특징을 추출해 낸다. <그림 3-1>은 OMT[7]의 재체도로 기술한 동영상의 계층 구조이다.



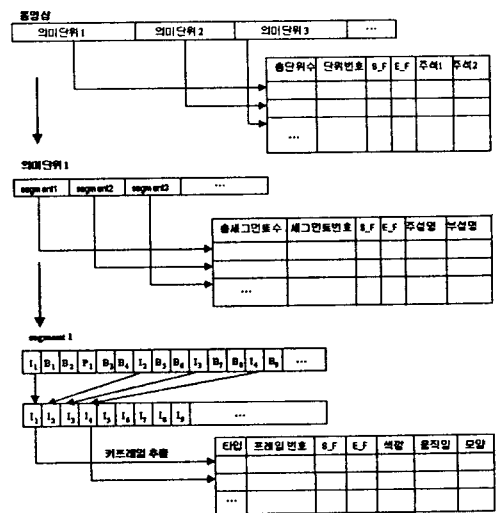
<그림 3-1> IVDM 동영상의 계층 구조

3.2 동영상의 인덱스 구조

동영상을 인덱싱하는 방법은 앞 절에서 설명했던 바와 같이 주석 기반과 내용 기반 검색 모두를 지원 가능하도록 하기위해 2가지 방법을 통합하여 사용한다.

먼저 전체적인 동영상에 대해 종류, 제목, 설명을 부여한다. 그리고 이를 동영상 스트림으로 표현되는 의미 단위로 나눈다. 이것은 주석 기반으로 큰 주제별로 검색할 때 유용하다. 의미 단위의 속성으로는 하나의 동영상이 몇 개로 나누어 관리되는지를 나타내는 총의미 단위수와 현재 의미 단위 번호, 시작 프레임 번호와 종료 프레임 번호를 속성으로 가진다. 이 단계의 주석은 아주 객관적인 사실들만을 동영상의 성격에 맞게 부여하여야 한다. 예를 들어, 종류가 드라마인 경우 방영 회수가 총단위수가 될 수 있다. 부가적인 주석은 방영 날짜, 시간, 제작자, 텔런트, 시나리오 작가 등이 나올 수 있다.

좀 더 세부적인 검색을 위해 의미 단위를 더 작은 크기의 세그먼트로 나누어 주석을 부여한다. 세그먼트에 부여하는 주석도 의미 단위에 추가하는 방법과 같이 주석의 수와 타입은 자유롭다. 차이점은 의미 단위보다 장면 설명 위주의 주석이 많을 것이다. 이보다 더 세부적으로 동영상을 나누어 주석을 부여하는 것은 주석을 다는 사람의 주관을 배제하기 힘들 뿐만 아니라 사람이 일일이 수작업을 하는데서 오는 오버헤드가 커서 비효율적이다.



따라서 더 구체적인 검색에서는 객관적인 이미지 분석 기법을 이용하기 위해 세그먼트에서 키프레임을 추출해 낸다. 추출된 키프레임에서 색깔, 모양, 움직임 벡터를 분석해서 인덱스화하여 검색에 이용한다.

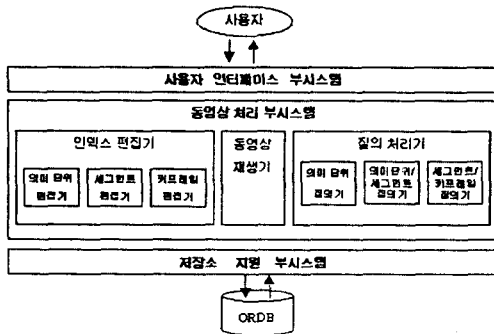
이렇게 동영상 데이터베이스에 저장되는 동영상 인덱스는 <그림 3-2>와 같다.

4. 통합 동영상 정보 관리 시스템 (IVIMS)의 설계

3장에서 제시한 IVDM에 기반하여 실제로 MPEG-2로 압축된 동영상 데이터를 관리하는 시스템(Integrated Video Information Management System, IVIMS)을 개발한다. 이 시스템은 MPEG-2로 압축된 동영상에 대해 상위 단계에서 주석을 부여하여 주석 기반 검색을 지원하고, 하위 단계에서 이미지 분석 기법을 적용하여 내용 기반 검색을 지원한다.

4.1 IVIMS의 구조

IVIMS는 <그림 4-1>처럼 사용자 인터페이스 부시스템, 동영상 처리 부시스템, 저장소 지원 부시스템으로 구성되고, 동영상 처리 부시스템은 인덱스 편집기, 동영상 재생기, 질의 처리기로 구성된다. 인덱스 편집기는 검색을 위해서 동영상 자료에 대한 주석을 미리 인덱스화하기 위한 모듈인데, 의미 단위 편집기, 세그먼트 편집기, 그리고 키프레임 편집기로 나뉘어진다. 각 편집기는 의미상 순서대로 계층을 이룬다. 동영상 재생기는 하나의 모듈로 이루어지고, 질의 처리기는 의미 단위 질의기, 의미 단위



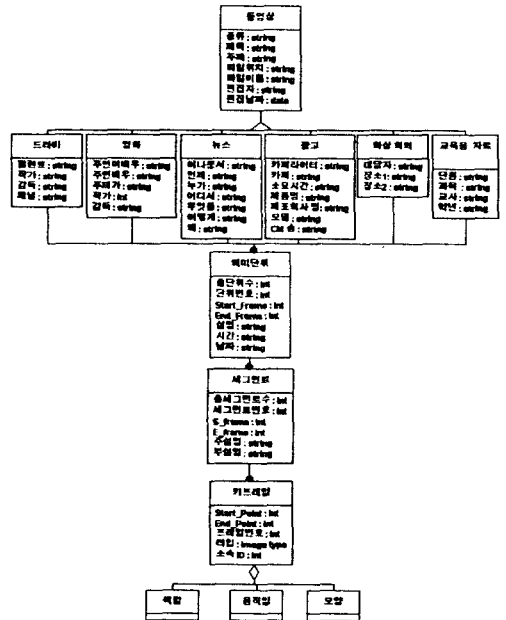
<그림 4-1> IVIMS의 구조도

/세그먼트 질의기, 세그먼트/키프레임 질의기로 나누어진다. 의미 단위 질의는 의미 단위 인덱스 수준에서 이루어지고, 의미 단위/세그먼트 질의기는 의미 단위와 세그먼트 인덱스를 조합하여 이루어지며, 세그먼트/키프레임 질의기는 세그먼트와 키프레임의 인덱스 조합으로 검색이 이루어진다.

4.2 IVIMS의 데이터베이스 스키마 설계

IVIMS의 동영상 관리를 위한 데이터베이스 스키마 구조는 <그림 4-2>와 같다.

제일 상위에 있는 동영상은 종류, 제목, 주제, 파일 위치, 파일 이름 및 편집자를 속성으로 가지고, 그 하위 개체로 드라마, 영화, 뉴스 등이 있는데 이 모든 개체들은 동영상의 속성을 상속받는다. 동영상의 종류 속성을 가지고 하위 단계의 개체가 무엇이 될지 판단하게 된다. 드라마, 영화, 뉴스 등은 0개 이상의 의미 단위로 구성되고, 의미 단위는 0개 이상의 세그먼트로, 다시 세그먼트는 0개 이상의 키프레임으로 구성된다. 여기서 의미 단위까지의 속성으로는 객관적으로 드러나는 확실한 정보만 그 속성으로 정의하고 더 자세한 부가 설명이 필요한 경우는 다시 세그먼트라는 개체를 만들어 주석을 부여한다.



<그림 4-2> IVIMS의 데이터베이스 스키마 구조

<<의미 단위 검색>>

```
Select F_location, F_name, Start_frame,
       End_frame from Meaning_unit
where <검색필드> <관계> '검색값' [and(or)
<검색필드> <관계> '검색값' [and(or)
<검색필드> <관계> '검색값'];
```

<<세그먼트 검색>>

```
Select F_location, F_name, Start_frame,
       End_frame from Segment
where <검색필드> <관계> '검색값' [and(or)
<검색필드> <관계> '검색값' [and(or)
<검색필드> <관계> '검색값']] and
M_explanation like '주설명' and
S_explanation like '부설명';
```

6. 결론 및 향후 연구 과제

최근 방대한 양의 멀티미디어 정보가 쏟아져 나오고 있는데, 그 중에서도 동영상 데이터가 큰 비중을 차지한다. 따라서 동영상 정보에 대한 검색 요구가 많아지고 있다.

본 연구에서는 이러한 동영상 정보를 효율적으로 검색하기 위해 이전 연구들의 분리된 2가지 접근법 즉, 주석 기반과 내용 기반 검색의 장점을 모두 수용하는 통합 동영상 데이터 모델(IVDM)을 제시하였다. 그리고 이 모델을 기반으로 MPEG-2로 압축된 동영상 정보를 관리하는 시스템을 설계하고 구현하였다.

현재, 본 논문에서 기술한 통합 동영상 정보 관리 시스템(IVIMS)의 실제 구현은 주석 기반의 검색만 지원하고 있으며 내용 기반 검색에 대한 구현은 진행 중이다. 이와 함께 계속 더 연구가 진행되어야 할 부분은 본 논문에서는 인덱스 정보만 데이터베이스화해서 저장했지만 앞으로는 3장에서 설계된 내용인 동영상 데이터 자체도 검색에 편리한 형태로 구조화하여 DB에 저장한다면 훨씬 더 빠른 검색이 가능하리라 생각된다.

참고문헌

[1] 허진용, 김인홍, 배종민, 강현석, "MPEG-2 압축 동영상 정보 관리 시스템의 설계 및 구현", 정보처리학회 논문지(계제예정)

[2] 김기병, 김형주, "내용 기반 검색 및 주석 기반 검색을 통합하는 비디오 데이터 모델의 설계 및 구현", 정보과학회 논문지(C) 제3권 제2호, pp. 115-126, 1997.

[3] 김기욱, 김형주, "비디오 주석 시스템의 설계 및 구현", 정보과학회 논문지, 제24권, 제6호, pp. 588-597, June 1997.

[4] S. Smoliar and H. Zhang, "Content-Based Video Indexing and Retrieval," IEEE MultiMedia, pp. 62-72, 1994.

[5] H. Aoki, S. Shimotsuji, and O. Hori, "A Shot Classification Method of Selecting Effective Key-Frames for Video Browsing," ACM Multimedia 96, pp. 1-10, 1996.

[6] I. A. R. Center, "Query by Image and Video Content: The QBIC System," IEEE Computer, 28(9), pp. 23-32, Sept. 1995.

[7] J. Rumbaugh, M. Blaha, W. Premeriani, F. Eddy, and W. Lorensen, *Object-Oriented Modeling and Design*, Prentice-Hall, 1991.