

MPEG-2 압축 동영상 정보 관리 시스템의 설계 및 구현

허진용[†] · 김인홍^{††} · 배종민^{†††} · 강현석^{†††}

요 약

동영상 정보는 자료의 양이 방대하고 다양하여 압축된 형태로 저장하고 검색하는 것이 매우 중요하다. 본 논문에서는 동영상 문서의 구조를 추출하고 동영상 스트림의 인덱싱에 대한 의미를 부여할 수 있는 일반적인 데이터 모델을 제시하고, 이 모델을 이용하여 MPEG-2로 압축된 동영상 자료를 데이터베이스에 저장하고 검색하는 MPEG-2 압축 동영상 정보 관리 시스템(CVIMS, Compressed Video Information Management System)을 제안한다. CVIMS는 MPEG-2 파일에서 I-프레임들을 추출하고, 그 중에서 키 프레임(대표 이미지)을 선택한 후, 선택된 키 프레임에 대한 축소그림(thumbnail)과 캡션 및 그림 설명 정보를 데이터베이스에 저장한다. 그리고 데이터베이스에 저장된 MPEG-2 동영상을 질의어나 키 프레임들을 통해 검색한다.

Design and Implementation of MPEG-2 Compressed Video Information Management System

Jin-Yong Heo[†] · In-Hong Kim^{††} · Jong-Min Bae^{†††} · Hyun-Syug Kang^{†††}

ABSTRACT

Video data are retrieved and stored in various compressed forms according to their characteristics. In this paper, we present a generic data model that captures the structure of a video document and that provides a means for indexing a video stream. Using this model, we design and implement CVIMS (the MPEG-2 Compressed Video Information Management System) to store and retrieve video documents. CVIMS extracts I-frames from MPEG-2 files, selects key-frames from the I-frames, and stores in database the index information such as thumbnails, captions, and picture descriptors of the key-frames. And also, CVIMS retrieves MPEG-2 video data using the thumbnails of key-frames and various labels of queries.

1. 소 개

최근 인터넷과 웹의 발전으로 멀티미디어, 특히 동영상 정보의 검색에 대한 수요가 급증하고 있고, 각종 객

체 지향 데이터베이스 시스템들이 개발되면서 이들을 멀티미디어 시스템에 활용할 수 있게 되었다. 이에 따라 동영상 정보를 효율적으로 저장하고 검색하는 기법에 대한 연구[2, 4, 10, 16]와 동영상 데이터베이스의 모델링에 대한 연구[11, 12]가 많이 이루어지고 있다.

그러나 이러한 멀티미디어 DBMS를 이용하는 동영상 정보의 관리는 검색 대상인 비트맵이나 웹페이지 패턴을 압축하지 않은 상태로 관리하는 방식이다. 그런데 동영상의 특성상 압축되지 않은 자연 영상을 그대로 저장

* 본 연구는 경상대 부설 정보통신 연구센터의 연구비 지원에 의한 것임.

† 준 회원: 경상대학교 전자계산학과 석사과정 졸업

†† 정 회원: 하이콤 테크놀로지 대표이사

††† 종신회원: 경상대학교 컴퓨터학과 교수

논문접수: 1997년 12월 15일, 심사완료: 1998년 3월 25일

하고 검색하거나 전송하기에는 많은 어려움이 따른다.

따라서 동영상 정보 관리 시스템에서 이러한 문제들을 해결하고 실용화하기 위해서는 동영상 정보를 압축하여 저장하고, 압축된 상태에서 실시간으로 검색하여 전송할 수 있는 기술이 요구되며, 이를 반영한 일반적인 동영상 분석을 기술할 수 있는 데이터 모델이 요구된다.

본 논문은 일반적인 동영상 문서에 대한 구조를 추출하고 동영상 스트림의 인덱싱에 대한 의미를 부여할 수 있는 일반적인 데이터 모델을 제시하고, 이 모델을 이용하여 MPEG-2 기술로 압축된 동영상에 대한 텍스트 정보를 추출하여 데이터베이스 시스템에 저장한 후, 질의어나 대표 이미지(키 프레임)를 이용하여 검색할 수 있게 하는 압축 동영상 정보 관리 시스템(CVIMS, Compressed Video Information Management System)을 개발한 내용을 기술한다.

논문의 구성은 다음과 같다. 2장에서는 관련 연구에 대해 소개하고, 3장에서는 본 연구에서 제시한 동영상 문서에 대한 일반적인 데이터 모델(generic data model)에 대해 기술하며, 4장에서는 제시한 모델을 이용하여 개발한 압축 동영상 정보 관리 시스템(CVIMS)의 설계 내용을 기술한다. 5장에서는 실제 구현된 CVIMS의 모양을 나타낸다. 6장에서는 결론 및 향후 연구 과제를 논의한다.

2. 관련 연구

동영상 문서에 대한 일반적인 동영상 데이터 모델에 대한 연구로 [11, 12] 등이 있는데, 현재 동영상 문서의 표준으로 각광받는 MPEG-2와 관련된 내용에 대한 언급이 없어 확장이 필요하다. 또한 동영상 객체의 효율적인 관리를 위해 Dublin Core 기반의 메타데이터를 이용한 확장이 필요하다.

동영상 자료를 저장하고 검색하는 방법은 크게 내용을 기초로 한 검색과 동영상에 대한 캡션(주석)을 이용한 검색으로 나눌 수 있다.

내용을 기초로 한 동영상 정보 검색(content based video retrieval)은 동영상 각각의 프레임에서 모양이나 색깔, 움직임 등을 이용하여 동영상에 대한 의미 정보를 추출하고 이를 기반으로 영상 정보를 검색하는 방법이다[3, 5, 7, 9, 15].

캡션(주석)을 기초로 한 동영상 정보 검색(caption

based video retrieval)은 동영상에 대해 캡션 정보 및 그림 설명 정보를 주석 형태로 삽입하여, 이를 기반으로 영상 정보를 검색하는 방법이다[2, 10, 16].

동영상에 대한 주석 형태의 텍스트 정보를 이용할 경우, 사용자는 세밀한 부분까지 원하는 영상 정보를 검색할 수 있는 장점을 지니고 있다. 반면, 추가적으로 주석을 부가하기 때문에 사용자간에 이러한 주석 정보를 공유하기 어려운 단점이 있다. 본 연구에서는 캡션을 이용한 영상 정보 검색을 바탕으로 사용자간에 주석을 공유하고 일관성을 유지할 수 있도록 주석을 캡션과 그림 설명 정보로 나누어 관리하는 시스템을 개발한다. 즉, 캡션은 미리 정해진 형태에 맞게 주석을 기술하고, 그림 설명 정보는 사용자가 직접 동영상을 보면서 필요한 정보를 기술하는 형태를 사용한다.

최근 MPEG으로 압축된 동영상 자료에 대한 효율적인 인덱스 생성 기술에 대해 활발한 연구가 있었다[3, 15]. 이들은 주로 압축된 동영상에서 내용을 분석하여 자동으로 키 프레임을 선택하여 인덱스를 구성하는 방법이다.

그런데 실제 동영상의 내용을 분석하여 검색 시스템을 만든다고 할 때, 세부적인 부분까지 내용이 파싱되어 검색에 사용되는 단계에는 아직 미치지 못하기 때문에 효율성이 다소 떨어진다. 또한 키 프레임을 추출하기 위해서 압축된 동영상을 모두 풀어서 파싱할 경우, 시간이 너무 걸리는 단점이 있다. 비록 최근 소개된 알고리즘들[5, 10]은 이러한 단점을 보완하기 위해 실제 압축된 형태의 MPEG 자료를 분석하고 편집하는 방법에 대해서 기술하고 있지만, 여전히 키 프레임을 추출하여 검색하는데 시간이 많이 걸리며 검색의 형태도 제한을 받는다.

따라서 본 연구에서는 기존에 소개된 동영상 데이터 모델[11, 12]을 MPEG-2와 Dublin Core 기반의 메타데이터를 포함하는 확장된 동영상 데이터 모델을 개발하고, 이 모델을 기반으로 동영상 정보를 검색할 수 있는 시스템을 구현한다.

3. 확장된 동영상 데이터 모델(Enhanced Generic Video Data Model)

동영상 자체와 동영상의 내용과 구조에 대한 기술인 메타데이터를 통합적으로 데이터베이스를 통해 관리하기 위해서는 동영상에 대한 일반적인 데이터 모델이 제

공통되어야 한다[12]. 7번째, 이러한 일반적인 모델이 각각의 응용마다 존재할 필요는 없다. 동영상 데이터 모델의 주 목적은 공유를 위해 적어도 하나의 일반적인 표준을 만드는 것이다. 몇몇 응용은 특정한 개념을 가지고 이러한 일반적인 데이터 모델을 확장할 수 있다. 우리가 제안하는 일반적인 동영상 데이터 모델은 추상화의 레벨을 정의하여 동영상 문서에 대한 구조를 정의하기 위해 세그멘테이션 접근법을 사용한다.

그림 1에 나타난 객체도(OMT)[8]는 [11, 12]의 연구를 확장한 것으로 아래와 같은 내용을 추가하여 보완하였다.

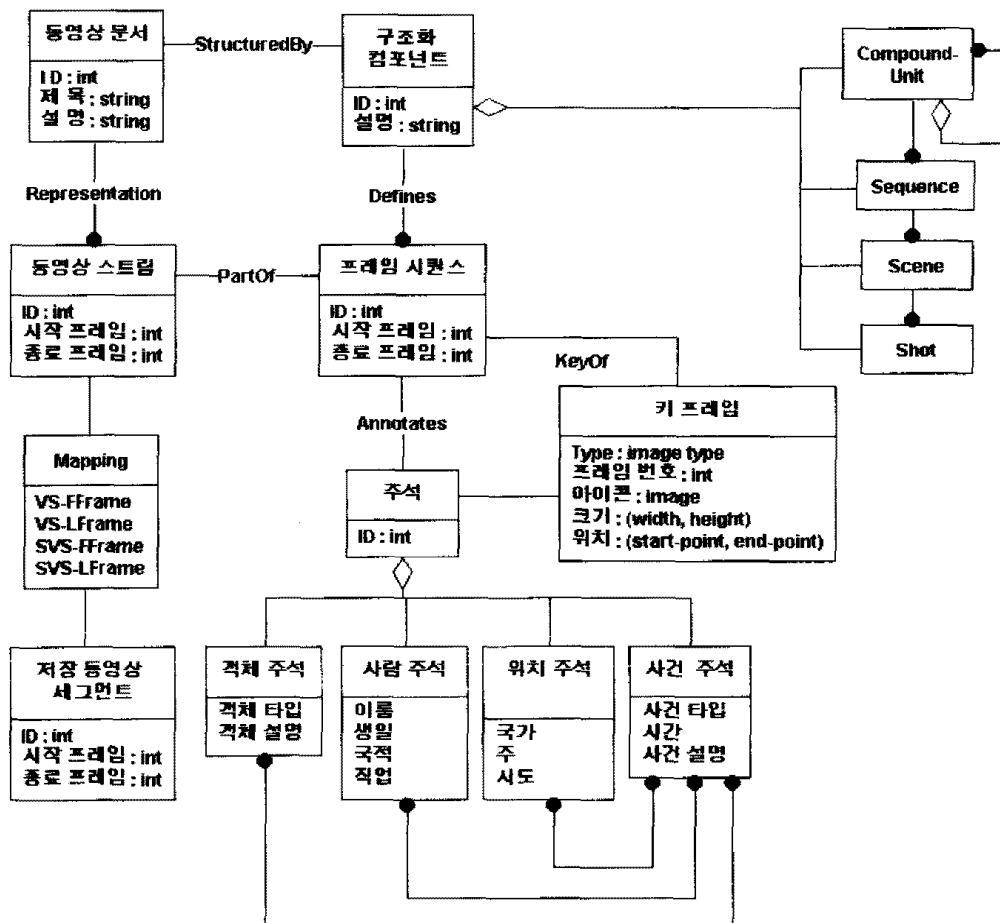
- (1) 키 프레임 관리 모듈
- (2) 객체 주석의 관리에 Dublin Core 메타데이터 활용

이렇게 확장된 동영상 데이터 모델(Enhanced Generic Video Data Model : EGVDM)은 동영상

데이터의 구조화(structuring), 동영상 데이터의 사용자 중심 주석화(annotation), 동영상 데이터의 공유 및 재사용(sharing and reusing) 기능을 제공하는 프레임워크이다.

EGVDM에서 동영상 데이터는 저장 비디오 세그먼트(stored video segment)라 명명된 프레임의 연속적인 그룹이다. 그리고 동영상 문서(video document)는 하나 혹은 그 이상의 저장 비디오 세그먼트로 사상되는 비디오 스트림에 의해 표현된다.

중요하고 유연성있는 동영상 정보 단위는 프레임 시퀀스인데, 이는 동영상 문서에서 특정 프레임 시퀀스를 추출한 것이다. 프레임 시퀀스는 하나의 프레임이 될 수도 있고 전체적인 동영상 문서가 될 수도 있다. 프레임 시퀀스가 우리의 데이터 모델에서 가장 중요한 부분을 차지하는데, 이는 동영상 매체의 구조 단위(structural unit)와 동영상에 대한 주석(annotation)



(그림 1) 객체도를 이용한 확장된 동영상 정보에 대한 일반적인 데이터 모델
 (Fig. 1) An enhanced generic data model for video information using a OMT

을 연결하는 수단이다. 프레임 시퀀스는 비디오 스트림과 PartOf 관계로 정의되고 (first frame, last frame)의 쌍으로써 비디오 스트림을 참조한다. 이를 기초로 각 기능에 대해 설명한다.

3.1 구조(Structure)

동영상 문서는 전체적인 동영상 문서와 개별적인 프레임의 2가지 추상화 레벨로 나눌 수 있다. 대부분의 응용에서 전체적인 동영상 문서는 추상화의 레벨이 너무 높다. 반면에 하나의 프레임은 너무 낮다. 이는 하나의 프레임은 상영시간이 너무 짧기 때문이고 또한 매우 작은 동영상 문서라 하더라도 꽤 많은 프레임으로 구성되기 때문이다.

국립 영상 제작소에서 제작한 대한뉴스의 분석을 통해, 장면과 뉴스 아이템같은 것으로 추상화하는 것이 동영상 정보를 참조하는데 용이하고, 쉽게 그 내용을 이해할 수 있다는 것을 알게 되었다. 다른 연구에서는 동영상 정보를 브라우징 하는 것이 매우 어렵다는 것을 보여주고 있다[13]. 본 연구에서는 구조적 추상화가 동영상을 브라우징하는데 매우 유용하다는 것을 보여주고 있다.

제안된 EGVDM은 필름 이론(film theory)[6]과 동영상 매체의 세그멘테이션에 근간을 두고 동영상 매체의 프레임 시퀀스와 연관된 구조적 구성 성분의 개념으로 만들어 졌다. 구조적 구성 성분의 개념은 다시 복합 단위(compound unit), 시퀀스(sequence), 장면(scene), 샷(shot)으로 세분화되며, 이들 부 클래스들은 서로 계층적인 관계성(hierarchical relationship)으로 정의된다. 샷은 하나 혹은 그 이상의 연속된 프레임으로 구성되고, 시공간적으로 연속된 행위(action)로 나타난다. 장면은 여러개의 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.

이러한 EGVDM은 동영상 정보를 필요로 하는 다양한 응용에 적용할 수 있는 일반적인 하부 구조를 제공하며, 뒷 장에서 이를 적용하여 구현한 CVIMS를 기술한다.

3.2 주제의 인덱싱(Thematic Indexing)

동영상 문서의 구조는 동영상 매체를 근간으로 하여 추출한 것이지만 동영상 매체의 모든 텍스트 정보를 표

현하는 데는 적합치 못하다. [11, 14]의 연구에서 논의된 것처럼 동영상 매체의 내용에 대한 상세한 기술은 임의의 프레임 시퀀스에 대한 기술만으로도 적합하다는 것을 알 수 있다. 따라서 EGVDM에서는 임의의 프레임 시퀀스에 대해 주석을 부여할 수 있도록 하였다. 이러한 주석은 임의의 구조적 구성 성분에 독립적으로 구성되어 있어, 주제의 인덱싱(thematic indexing)은 동영상 문서의 구조를 기술하는데 충분하다.

그림 1에서와 같이 프레임 시퀀스는 주석 개체와 키 프레임 개체로 나누어 진다. 키 프레임 개체는 프레임 시퀀스에서 특정 대표 이미지를 추출하여 이것에 대한 이미지, 이미지 타입, 프레임 번호, 크기, 위치 정보로 구성되며, 주석 개체는 객체 주석(object annotation), 사람 주석(person annotation), 위치 주석(location annotation), 사건 주석(event annotation)의 부 클래스로 나누어 진다. 각 주석 개체에 대한 타입과 의미는 각 응용의 상황에 맞게 정의될 수 있다. 그러나 EGVDM에서는 일반적인 주석 타입들의 집합을 부 클래스에서 보는 것 처럼 제공하고 있으며, 이러한 주석 개체와 키 프레임 개체를 이용하면 효율적인 검색이 가능하다.

객체 주석은 객체의 타입, 객체 설명으로 구성되는데 본 연구에서는 Dublin Core 기반의 메타데이터를 이용하여 객체 주석을 정의하였다. 즉, 객체 주석은 제목(title), 주제(subject), 식별자(identifier), 관계성(relation), 권한(right), 언어(language), 문서형태(format) 등으로 정의될 수 있다. 이렇게 함으로써 문서의 제작자와 제공자에게 효율적인 기술 수단을 제공하고, 다양한 검색 도구들 사이의 상호 운용성을 쉽게 하며, 분류 시스템들간의 사상이 가능하도록 한다.

3.3 동영상 매체의 공유와 재사용(Sharing and Reusing)

같은 동영상 매체가 서로 다른 동영상 문서에서 사용될 수 있다. 따라서 EGVDM에서는 동영상 문서의 내용을 논리적 개념으로 정의함으로써 이러한 요구를 충족하도록 하였다. 즉, 동영상 스트림과 물리적으로 데이터베이스에 저장되는 매체의 연결을 저장된 동영상 세그먼트상의 사상들의 집합으로 구성하여 같은 동영상 매체에 대해 데이터베이스에서 어떻게 표현되는가에 따라 다양한 방식으로 사용하는 것을 허용한다.

이렇게 EGVDM에서는 저장된 동영상 매체의 서로 다른 사용이 명백하게 표현되기 때문에 데이터베이스는

시상화 동영상 매체의 이미 부분이 사용되는데에 대한 정확한 정보를 갖게 된다. 또한 동영상 스트림과 관련된 주석과 구조적 정보를 이용하여 같은 동영상 매체에 저장되어 사용되는 서로 다른 동영상 매체에도 접근이 가능하게 된다.

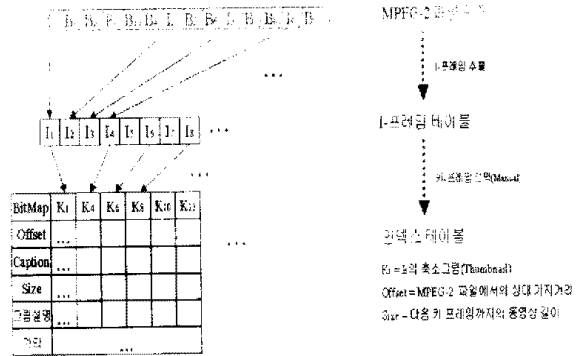
4. 압축 동영상 정보 관리 시스템(CVIMS)의 설계

이 장에서는 앞에서 제시한 동영상 데이터 모델에 기반하여 MPEG-2로 압축된 동영상 자료에서 키 프레임을 추출한 후 캡션 및 그림 설명 형태의 주석을 부여하고, 이를 텍스트 형태로 데이터베이스에 저장하여 관리할 수 있게 하는 압축 동영상 정보 관리 시스템(CVIMS)을 설계한다.

4.1 MPEG-2를 위한 인덱스 구조

MPEG-2에서는 동영상을 고효율로 압축하기 위해 움직임 추정 기술을 사용하는데, 이는 실제로 공간적인 중복성보다는 시간적인 중복성이 훨씬 많기 때문이다. 따라서 움직임을 얼마나 잘 추정했는지가 압축의 관건이다. 움직임을 추정하는데 사용되는 참조 픽처(reference picture)에 따라 앞의 프레임을 참조하는 순방향 예측(forward prediction)과 뒤의 프레임을 참조하는 역방향 예측(backward prediction)이 있다.

MPEG-2 동영상 압축 파일은 크게 3종류, 즉 I-프레임, P-프레임, B-프레임으로 구성된다(1). 이중 I-프레임은 시간적인 압축 기술은 사용하지 않고, 공간적인 압축 기술만을 사용해서 압축한 프레임이다. 따라서 I-프레임은 단독으로 복호화가 가능하고 임의 접근을 할 수 있기 때문에 기준 프레임이 될 수 있다. 그래서 CVIMS에서는 MPEG-2로 압축된 비디오 파일내의 모든 I-프레임들을 키 프레임의 후보가 될 수 있다고 가정한다. 이 가정에 따라 CVIMS는 MPEG-2로 압축된 동영상에서 I-프레임을 추출하여 사용자가 직접 키 프레임으로 선택할 수 있는 방법을 제공한다. 또한 각 동영상에 대한 검색은 실제 프레임으로 검색하는 것이 아니라, 각 키 프레임에 대한 캡션 정보를 작성하여 이것을 키 프레임과 함께 구조화하여 데이터베이스에 저장한 다음, 이러한 캡션 정보를 이용하여 검색을 수행한다. MPEG-2 동영상에 대한 인덱스 구조 및 캡션과의 관계는 그림 2와 같다.



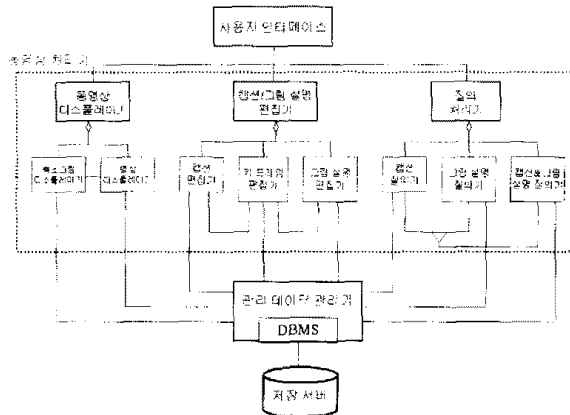
(그림 2) MPEG-2 파일과 인덱스 정보의 구조 (Fig. 2) MPEG-2 File and Index Information Structure

이 방법은 비록 2장에서 소개한 자동 인덱싱 기법들(3, 7, 15) 보다는 효율성이 뒤떨어지지만, 기술적으로 동영상에서 키 프레임을 추출한 후 영상 인식을 통해 키 프레임의 내용을 처리하는 기술이 아직 완전하게 구현되기 어렵기 때문에 중간 과정으로 이 키 프레임에 캡션을 붙여서 검색하도록 하는 것이 실용적이라 판단하여 본 연구에서는 이 방식을 사용한다.

4.2 CVIMS의 전체 구조

CVIMS는 사용자 인터페이스와 MPEG-2 동영상을 인덱싱할 수 있는 캡션 및 그림 설명 편집기, 사용자의 다양한 질의를 처리하는 질의 처리기, 질의의 결과를 디스플레이하는 영상 디스플레이기, 인덱스 자료와 동영상 자료를 관리하는 데이터베이스, 그리고 MPEG-2 동영상을 저장하는 저장 서버로 구성된다. 그림 3에서는 CVIMS의 구성 요소들과 각 구성 요소들이 갖는 부클래스들 및 관계성들을 나타내고 있다. 그림 3에서 보는 바와 같이 CVIMS는 크게 사용자 인터페이스, 동영상 처리기, 관리 데이터 관리기로 구성된다. 동영상 처리기는 다시 질의 처리기, 캡션/그림 설명 편집기, 그리고 동영상 디스플레이기로 구성된다. 질의 처리기는 질의의 형태에 따라 캡션 질의기와 그림 설명 질의기, 그리고 캡션과 그림 설명을 조합한 질의기로 나누어 진다. 이는 사용자의 질의를 받아 들어 관리 데이터 관리자가 관리하는 각각의 개체를 탐색하여 원하는 결과를 가져온다. 캡션/그림 설명 편집기에서는 사진 디코딩된 I-프레임의 리스트에서 키 프레임을 선택하고, 각각의 키 프레임에 대한 캡션 정보 및 그림 설명 정보를 작성하여 데이터베이스에 저장한다. 동영상 디스플레이기는 질의의 결과를 디스플레이해 주는 부분으로

축소그림(thumbnail picture) 형태의 아이콘으로 출력하는 축소그림 디스플레이와 실제 동영상을 디스플레이해 주는 영상 디스플레이기로 나누어진다. 이는 데이터베이스에서 축소그림과 관련된 정보 및 동영상이 저장된 저장 서버에서의 파일명과 가지거리를 가져오고, 실제 동영상은 저장 서버에서 가져온다. 마지막으로 관리 데이터 관리기는 데이터베이스에 저장되는 정보를 관리하는 부분으로, 편집기에서 작성한 다양한 인덱스 정보 및 캡션/그림 설명 정보를 관리한다.



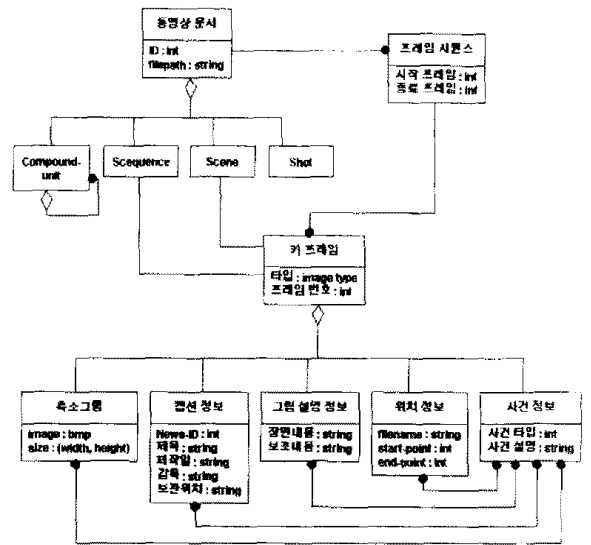
(그림 3) CVIMS 시스템의 설계도
(Fig. 3) CVIMS System Design

4.3 관리 데이터의 구조

그림 4은 3장의 EGVDM을 근간으로 하여 관리 데이터 관리기에서 관리하는 동영상 데이터베이스의 인덱스 정보에 대한 구조를 키 프레임 중심으로 나타낸 것이다. 즉, 동영상 문서에서 프레임 시퀀스를 추출하고, 이것에 대한 키 프레임을 추출하여 키 프레임에 대한 주석 정보와의 관계를 객체도(OMT)[8]로 나타낸 것이다. 키 프레임을 중심으로 인덱스에 관련된 정보는 축소그림 형태의 비트맵 이미지로 관리하고, 캡션, 그림 설명 및 위치에 대한 정보는 각각 하나의 객체로 관리한다. 또한 사건 정보는 축소 그림, 캡션 정보, 그림 설명 정보, 위치 정보에서 각각 발생할 수 있는 사건들의 타입과 이에 대한 설명을 기술한 객체로 실제 동영상 문서에 대한 질의의 수행은 사건 정보에 기술된 내용을 통하여 이루어진다.

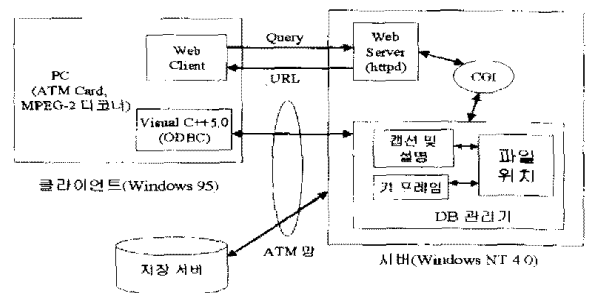
4.4 CVIMS의 개발 환경

CVIMS는 클라이언트/서버 환경으로 구축되었으며,



(그림 4) 동영상 정보의 인덱스 구조
(Fig. 4) Video Information Index Structure

서버는 MPEG-2 동영상에서 I-프레임을 추출하고, 여기서 키 프레임을 선정하여 키 프레임에 대한 다양한 정보를 기록, 데이터베이스에 구조화하여 관리한다. 클라이언트는 사용자의 질의를 받아들여 처리하는 역할을 수행한다. 클라이언트는 사용자와의 인터페이스를 위해 사용자의 요구 즉, 키 프레임에 대한 캡션 및 그림 설명 정보를 편집하는 모듈과 사용자의 질의를 처리하는 모듈, 그리고 결과를 보여주는 디스플레이 모듈로 구성되며, 자체 클라이언트와 서버의 연결은 VC++의 ODBC 드라이버를 이용하여 연결하였고, MPEG-2 파일은 ATM 망을 통하여 저장 서버에서 제공하도록 구성하였다. 또한 웹(Web)을 통하여 검색할 수 있도록 CGI(Common Gateway Interface) 기법을 이용하여 웹 서버와 데이터베이스를 연결하여 동영상을 검색하고 디스플레이할 수 있는 환경도 제공한다(그림 5 참조).



(그림 5) CVIMS의 클라이언트/서버 환경
(Fig. 5) CVIMS Client/Server Environment

5. 압축 동영상 정보 관리 시스템(CVIMS)의 구현 결과

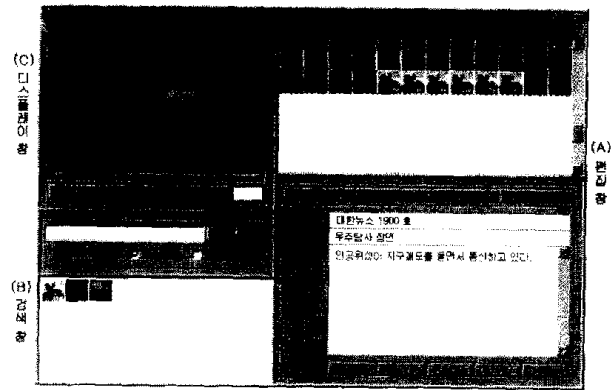
이 장에서는 4장에서 설계한 CVIMS의 구현 결과를 사용자 인터페이스 화면을 중심으로 알아 본다. 사용자는 자체 클라이언트 환경의 사용자 인터페이스를 통해서 CVIMS에 접근이 가능하다. 그림 6은 사용자가 CVIMS를 수행시켰을 때의 화면을 나타내고 있다. 구현 언어는 Visual C++ 5.0을 이용하였고, 관계형 DBMS인 SQL Server를 사용하였다. 동영상 정보의 효율적인 관리를 위해서는 객체지향 데이터베이스 시스템이 효과적일 것으로 판단되었으나 현재 본 연구팀에게 사용가능한 환경이 SQL Server이며 개발하고 있는 CVIMS가 프로토타입 수준으로, SQL Server를 이용한 정보 관리로도 기본 기능을 수행하는데 크게 문제가 없다고 판단되어 그대로 사용하였다. 화면의 구성은 크게 3개로 나누어지며, 편집을 할 수 있는 편집 창과 질의를 할 수 있는 검색 창, 그리고 동영상을 상영해주는 디스플레이 창이다.

5.1 편집 창

편집 창은 그림 6(A)에 나타나 있는데, 각각을 기능 중심으로 설명하면, 먼저 편집자는 파일 열기에서 편집할 MPEG-2 동영상 파일을 불러온다. 그러면 I-프레임에 대한 정보가 축소그림 형태의 비트맵 이미지로 I-프레임 리스트 창에 나타나게 된다. 편집자는 이 I-프레임 테이블에서 키 프레임이 될 수 있는 I-프레임과 그것의 적용범위를 마우스를 통해서 설정한다. 그런 다음 화면에 있는 플레이 버튼을 클릭하면 동영상 디스플레이 화면에 선택한 영역만큼의 동영상이 디스플레이되고, 편집자는 이 영상을 보면서 캡션 편집 창과 그림 설명 편집 창에서 선택한 키 프레임에 대한 캡션 정보 및 그림 설명 정보를 편집한다. 편집을 마치게 되면 입력 버튼을 눌러 키 프레임에 대한 비트맵 이미지, 위치 정보, 그리고 캡션 및 그림 설명 정보를 데이터베이스에 저장한다.

5.2 검색 창

검색 창은 편집 창에서 편집자가 작성한 키 프레임에 대한 캡션 및 그림 설명 정보를 이용하여 일반 사용자가 원하는 동영상을 검색하는 시스템이다. 그림 6(B)는 검색 부시스템에 대한 창이다. 검색 부시스템



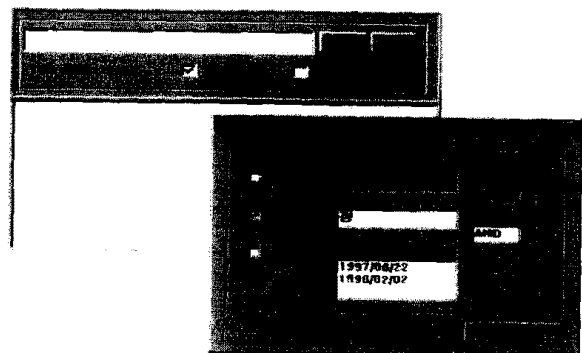
(그림 6) 사용자 인터페이스의 전체 화면
(Fig. 6) User Interface Total Screen

의 기능은 크게 기본 검색과 확장 검색으로 나누어지며, 기본 검색은 일반 사용자가 캡션 및 그림 설명 정보를 이용하여 원하는 정보를 찾는 검색이다.

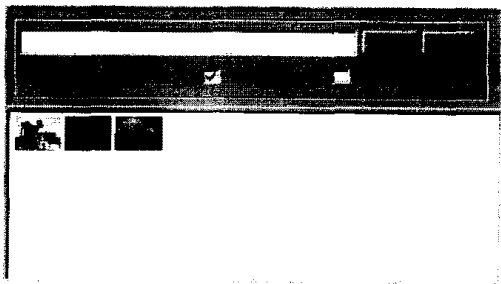
여기서는 and 연산 및 or 연산을 지원하며 체크 박스를 이용하여 캡션 정보로 검색할 것인지, 혹은 그림 설명 정보로 검색할 것인지를 결정한다. 또한 캡션 및 그림 설명 정보를 둘다 선택하여 검색할 수 있는 복합 기능도 제공한다. 확장 검색은 캡션 편집기에서 작성한 캡션 정보를 이용한 리스트 형태의 검색을 지원한다.

그림 7은 검색 창에서 기본 검색을 클릭한 후 검색 조건 및 검색 키워드를 입력한 경우의 화면을 나타낸 것이다. 이 창에서는 체크 상자 및 리스트 상자를 이용하여 검색할 항목을 설정하며, 선택된 항목에 대해 각각 키워드를 입력한다.

검색 창에서 사용자의 질의를 처리하는 아래와 같은 SQL 문을 통하여 실제 데이터베이스에 질의를 하며, 질의를 처리한 결과는 그림 8과 같이 축소그림 형태로 나타나게 된다.



(그림 7) 검색어 입력 화면
(Fig. 7) Keyword Input Screen



(그림 8) 검색 결과 출력 화면
(Fig. 8) Retrieval Result Output Screen

- ▶ 단순 질의 : 각 테이블의 키워드를 통한 질의
- 1) 캡션 테이블에 저장된 각각의 키워드를 이용한 질의

```
select * from caption_info
  where title = 검색어 [and(or) author
    = 검색어]
    [and(or) madeday = 검색어]
```

- 2) 그림 설명 테이블에 저장된 항목을 통한 질의

```
select * from picture_desc
  where main1_content = 검색어 [and(or)
    main2_content = 검색어] [and(or)
    content = 검색어]
```

- ▶ 복합 질의 : 캡션 및 그림 설명 테이블을 조합한 질의

Caption_info와 Picture_desc 테이블을 조인 웹을 통한 검색 또한 동영상에 대한 캡션 정보를 이용하여 검색을 수행하는데, 먼저 원하는 항목을 선택한 후 선택한 항목에 대한 검색어를 입력하는 형태이다. 이때 입력된 검색어는 CGI 프로그램을 통해서 데이터 베이스에 질의어의 형태로 전달되며, 검색된 결과는 html 문서로 웹상에 나타나게 된다. 여기에서 원하는 동영상의 축소그림을 클릭하면 이 축소그림을 키 프레임으로 갖는 동영상이 디스플레이되는데, 디스플레이되는 동영상 시스템은 Netscape의 플러그-인 형태로 구동된다.

5.3 디스플레이 창

디스플레이 창은 검색 창에서 검색한 결과 정보를 축소그림 형태의 비트맵 이미지로 나타내 주고, 여기서 선택한 키 프레임의 비트맵이 가리키는 실제 동영상을 디스플레이해 주는 화면이다. 그림 6(C)가 동영상 디

스플레이 창을 나타낸 것이다. 검색에서 찾은 정보가 축소그림 형태의 비트맵 이미지로 나타나며, 사용자가 여기서 특정 비트맵을 클릭하면 이 비트맵이 대표하는 실제 동영상을 디스플레이 창에 출력해 준다. 또한 편집 창에는 이 비트맵이 대표하는 실제 영상에 대한 캡션 및 그림 설명 정보를 출력한다. 이때의 편집 창은 단순히 캡션 정보와 그림 설명 정보를 보여주지만 하고, 사용자가 직접 편집할 수는 없다.

6. 결론 및 향후 연구 과제

동영상 문서는 그 형태가 다양하고 방대하기 때문에 이러한 문서를 효율적으로 관리할 수 있는 일반적인 데이터 모델이 필요하다. 따라서 본 연구에서는 [9, 10]의 연구에서 소개된 동영상 데이터 모델을 확장하여 EGVDM(Enhanced Generic Video Data Model)을 개발하였고, 이를 기반으로 MPEG-2로 압축된 동영상을 관리할 수 있는 압축 동영상 정보 관리 시스템(CVIMS)의 원형(prototype)을 설계하고 구현하였다.

본 연구에서 구현된 CVIMS는 관리 정보를 관리하기 위해 관계형 DBMS인 SQL Server를 사용하였고, 질의의 형태도 대부분 텍스트와 축소그림 위주의 검색을 주로 지원하고 있지만, 키워드 검색뿐만 아니라 텍스트 형태의 그림 설명 정보를 이용한 내용 검색도 일부 지원하고 있다. 또한 단순한 검색 형태이지만 웹을 통해서도 동영상의 검색이 가능하다.

본 연구는 [11, 12]의 연구에서 소개된 모델에 키 프레임 관리 모듈 및 객체 주석의 관리를 위해 Dublin Core 메타데이터를 추가하여, 일반적인 동영상 문서를 위한 동영상 데이터 모델을 개발하고, 초기 단계의 동영상 정보를 저장하고 검색할 수 있는 시스템을 구현했다는 데 의의가 있다. 특히, 본 연구에서 개발한 동영상 인덱서, 동영상 정보 저장기, 동영상 정보 검색기 등은 대규모 영상을 관리해야 하는 영상 도서관, 국립 영상 제작소, 방송국 등에서 필수적인 기술이다.

현재 구현된 CVIMS는 I-프레임의 테이블에서 사용자가 직접 키 프레임을 선정해야 한다. 그런데, MPEG-2로 압축된 비디오 파일은 I-프레임의 개수가 너무 많기 때문에 사용자가 이 모든 것을 일일이 선택해서 편집해야 하는 어려움이 있다. 또한 현재 구현된 CVIMS는 관계형 데이터베이스를 사용하였고, 실제 동영상 문서는 저장 서버에 하나의 파일로 존재하기 때문에 관리

상 어려움을 대포하고 있다. 이를 극복하기 위해 현재 키 프레임은 자동으로 추출할 수 있도록 내용을 기반으로 한 검색 알고리즘에 대한 연구를 진행 중이고, 동영상 문서를 직접 관리할 수 있도록 객체지향 데이터베이스에서의 구현도 진행 중에 있다.

참 고 문 헌

[1] D. Gall, "MPEG: A Video Compression Standard for Multimedia Applications," CACM, 34(4), pp.47-58, April 1991.

[2] E. Oomoto and K. Tanaka, "OVID: Design and Implementation of a Video-Object Database System," IEEE Trans. on Knowledge and Data Engineering, 5(4), pp.62-72, 1994.

[3] 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.

[4] H. Frater and D. Paulissen, Multimedia Mania, Abacus, 1993.

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

[6] J. Monaco, "How to Read a Film," The Art, Technology, Language, History and Theory of Film and Media, Oxford University Press, 1981.

[7] J. Meng and S. Chang, "CVEPS - A Compressed Video Editing and Parsing System," ACM Multimedia 96, pp.43-53, 1996.

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

[9] J. Smith and S. Chang, "Searching for Images and Videos on the World-Wide Web," Technical Report #459-96-25, Center for Telecommunications Research, Columbia University, New York, August 1996.

[10] M. Davis, "Media Streams: Representing Video for Retrieval and Repurposing," Ph. D. Thesis, Massachusetts Institute of Technology, 1995.

[11] R. Hjelmsvold, "Video Information Contents and Architecture," Proceedings of the 4th International Conference on Extending Database Technology, pp.28-31, March 1994.

[12] R. Hjelmsvold and R. Midtstraum, "Modelling and Querying Video Data," Proceedings of the 20th VLDB Conference, 1994.

[13] S. Stevens, "Next Generation Network and Operating System Requirements for Continuous Time Media," Proceedings of the Second International Workshop for Network and Operating System Support for Digital Audio and Video, November 1991.

[14] T. Smith, "If You Could See What I Mean... Descriptions of Video in An Anthropologist's Notebook," Master's Thesis, MIT, 1992.

[15] V. Kobla and D. Doermann, "Compressed Domain Video Indexing Techniques Using DCT and Motion Vector Information in MPEG Video," SPIE, Vol.3022, pp.200-211, 1977.

[16] 김기욱, 김형주, "비디오 주석 시스템의 설계 및 구현," 정보과학회논문지, 24권 6호, pp.588-597, June 1997.



허진용

1994년 2월 경상대학교 전자계산
학과 졸업(학사)
1998년 2월 경상대학교 전자계산
학과 석사과정 졸업
관심분야: 멀티미디어, 멀티미디어
정보 검색, SGML, 전
자 도서관



김인홍

1981년 2월 경북대학교 전자공학
과 졸업(학사)
1985년 2월 한국과학기술원 전자
계산학과 이학석사
1996년~현재 경상대학교 전자계
산학과 박사과정 재학중

1984년 한국전자통신연구원 연구원

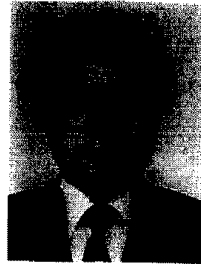
1990년 Maspar Computer Corporation(미)
1996년 Jinnes Technologies, Inc.(미)
1997년~현재 하이콤 테크놀러지 대표이사
관심분야: 멀티미디어 통신 및 응용, ATM, 웹 에이전트



배 종 민

1980년 2월 서울대학교 수학교육
학과 졸업(학사)
1983년 2월 서울대학교 대학원
계산통계학과 이학석사
(전산학)
1996년 2월 서울대학교 대학원
계산통계학과 이학박사
(전산학)

1982년~1984년 한국전자통신연구원 연구원
1984년~현재 경상대학교 컴퓨터과학과 교수
관심분야: 병렬 프로그래밍, SGML, 정보 검색, 전자
도서관



강 현 석

1981년 2월 동국대학교 전자계산
학과 졸업(학사)
1983년 2월 서울대학교 계산통계
학과 이학석사(전산학)
1989년 2월 서울대학교 계산통계
학과 이학박사(전산학)

1981년~1984년 한국전자통신연구원 연구원
1984년~1993년 전북대학교 전자계산학과 부교수
1993년~현재 경상대학교 컴퓨터과학과 교수
관심분야: 객체지향 데이터베이스, 소프트웨어 공학,
멀티미디어