

TIVRON: OODBMS 기반의 통합 비디오 검색 시스템

정희원 송용준*, 박동주*, 이동호*, 김형주*

TIVRON: The Integrated Video Retrieval System based on OODBMS

Yong-Jun Song*, Dong-Joo Park*, Dong-Ho Lee*, Hyoung-Joo Kim* *Regular Member*

요약

최근 대용량의 멀티미디어 응용들이 다양해짐에 따라 효율적인 멀티미디어 데이터 검색 기법의 필요성이 날로 증가하고 있다. 특히 비디오 데이터는 많은 활용 분야를 가지지만 비정형성과 구조의 복잡성으로 인한 연구의 어려움때문에 앞으로도 많은 연구가 필요한 실정이다. 본 연구실에서는 지난 몇 년동안 OODBMS를 이용한 MPEG 비디오 데이터의 효율적 저장 관리 연구, 주석 기반 비디오 검색 연구, 그리고 내용 기반 이미지 검색 연구 등 비디오 데이터에 대한 여러가지 연구들을 수행해왔다. 본 논문에서는 이러한 연구 결과들을 확장 통합하여 기존의 연구 결과들뿐만 아니라 복합 비디오 검색 기법까지 지원하는 OODBMS 기반의 통합 비디오 검색 시스템인 TIVRON을 제안한다.

ABSTRACT

The necessity of the efficient methods for multimedia data retrieval is increasing more and more on account of various large scale multimedia applications. Especially video data has a number of practical application areas, but the current technology is not mature enough to cover the hectic characteristics of video data such as being unnormalized and complicated. We have been performing research for past several years on the issues such as the annotation-based video retrieval, the content-based image retrieval, and the technique for storing MPEG video data using OODBMS. By both integrating and extending these results to support the complex video retrieval in addition to the existing functions, this paper proposes TIVRON as an integrated video retrieval system based on OODBMS.

I. 서론

최근 컴퓨터 기술의 급속한 발전으로 인해 기존의 텍스트 위주의 사용자 환경에서 벗어나 이미지, 그래픽, 오디오 및 비디오 데이터 등을 제공하는 멀티미디어 사용자 환경으로 변화하고 있다. 이와 같이 점점 늘어나는 멀티미디어 데이터 환경에서 사용자가 원하는 데이터를 효율적으로 찾을 수 있는 기술은 필수적이다. 특히 비디오 데이터는 많은 활

용 분야를 가지지만 비정형성과 구조의 복잡성으로 인한 연구의 어려움때문에 기술의 발전이 상대적으로 느려서 앞으로도 많은 연구가 필요한 실정이다.

비디오 검색은 방대한 양의 비디오 데이터를 사용하는 VOD, 원격 교육, 전자 도서관 등의 많은 응용에서 필수적인 기술로 현재 여러가지 비디오 검색 기법들이 연구되고 있는데 검색에 사용되는 특징에 따라 크게 다음과 같은 두가지로 분류할 수 있다^[1].

* 서울대학교 컴퓨터공학과 객체지향시스템 연구실(yjsong@oopsla.snu.ac.kr)

논문번호 : 98244-0608, 접수일자 : 1998년 6월 8일

※본 연구는 한국과학기술재단 특정기초연구과제(95-0100-23-04-3) 지원 및 서울대학교 공대연구지원소 관리로 수행되었습니다.

첫째, 주석 기반 검색. 비디오에 대한 의미 정보를 미리 주석으로 작성하여 저장한 후에 사용자가 원하는 비디오를 기술한 문자 질의와의 비교를 통하여 유사한 비디오를 찾는 방법이다. 이 방법은 사람이 인식할 수 있는 모든 내용을 주석으로 처리할 수 있기 때문에 일반적인 비디오 검색 시스템으로의 확장이 가능하다는 장점이 있지만 대용량의 데이터에 대하여 사람이 일일이 주석을 작성해야 하고 사용자에게 따라 비디오의 내용을 다르게 해석하거나 주석을 다르게 부여할 수 있다는 단점이 있다.

둘째, 내용 기반 검색. 비디오로부터 특징적인 내용을 표현하는 데이터를 추출하여 저장한 후에, 사용자가 원하는 비디오에 포함되는 이미지나 비디오 클립 등을 질의로 입력하면 그로부터 추출된 특징 데이터와의 비교를 통하여 유사한 비디오를 찾는 방법이다. 이 방법은 특징 데이터를 자동으로 추출할 수 있어 효율적이라는 장점을 갖지만 그 데이터가 정확한 내용을 표현한다는 것을 보장하기가 현재 기술로는 어렵다는 단점이 있다.

비디오의 내용이나 주제 등의 의미적인 정보를 이용하는 주석 기반 검색과 비디오의 프레임 모양이나 색상 등의 물리적 정보를 이용하는 내용 기반 검색은 비디오 검색 시스템의 구현에 있어 상호 보완적인 역할을 한다. 비디오 검색 시스템의 궁극적인 목적은 사용자가 원하는 비디오를 정확하게 찾는 것이기 때문에 이러한 상호 보완적인 비디오 검색 기법들을 함께 지원한다면 사용자는 원하는 비디오를 보다 자연스럽게 질의할 수 있으며 결국 보다 정확한 비디오들을 검색할 수 있다.

본 연구실에서는 지난 몇년동안 멀티미디어 데이터 지원이 가능한 SOP OODBMS 연구^[4], SOP를 이용한 멀티미디어 데이터의 효율적 저장 관리 연구, 비디오 데이터 모델 연구^[1], MPEG 비디오 데이터의 다양한 조작 연구^[3], 비디오에 대한 주석을 이용하여 원하는 비디오를 쉽게 찾을 수 있는 주석 기반 비디오 검색 연구^[2], 그리고 이미지의 모양 특징을 이용하여 원하는 이미지/비디오를 검색하는 내용 기반 이미지/비디오 검색 연구^[5] 등 다양한 연구들을 수행해왔다.

이러한 연구 결과들을 바탕으로 본 연구에서는 SOP에 구축된 MPEG 비디오 DB로부터 주석 기반 검색 기법과 내용 기반 검색 기법뿐만 아니라 두가지 기법들을 함께 사용하는 복합 비디오 검색까지 지원하는 통합 비디오 검색 시스템인 TIVRON(The Integrated Video Retrieval System based on

OODBMS)을 제안한다.

본 논문의 구성은 다음과 같다. 먼저 II장에서는 비디오 검색에 관련된 연구들을 살펴보고, III장에서는 본 연구에서 제안하는 복합 비디오 검색 기법을 소개한다. IV장에서는 TIVRON의 설계로서 관련 클래스 및 구성 모듈들에 대하여 소개하며, 그 구현에 대해서 V장에서 설명한다. 그리고 VI장에서는 TIVRON을 이용한 실제 검색 예를 보이고, 마지막으로 VII장에서 본 논문의 결론 및 향후 연구 계획을 다룬다.

II. 관련 연구

본 연구와 관련된 연구 분야로서 비디오를 표현하는 방법에 관한 비디오 모델 연구와 원하는 이미지와 비디오를 검색하는 주석 기반 비디오 검색, 내용 기반 비디오 검색등의 비디오 검색 연구를 들 수 있다. 지난 10여년동안 많은 관련 연구들이 수행되어 왔는데, 이 장에서는 그 중에서 대표적인 연구들을 소개한다.

먼저 비디오 모델에 관한 연구로서, Gibbs 등은^[13] 객체지향 개념을 도입하여, 스트림기반 멀티미디어 데이터를 모델링하였다. 이들의 주요 관심사는 시간관련 매체에 대해 일반적인 모델의 구성에 있었기에, 영상정보의 주석의 저장, 검색 및 관리에 대해서는 언급하고 있지 않다. Day, Dagtas 및 Iino는^[15] 비디오 데이터를 위해 객체지향 개념 모델을 제안하였는데, 방향 그래프, 개념적 시간 객체, 개념적 공간 객체, 물리적 객체를 이용하여 비디오 데이터에 포함되어 있는 공간적 속성과 시간적 속성을 분리하였다. 김 기병 등은^[1] 비디오 스트림 계층, 특징 계층, 특징 시퀀스 계층, 부 개념 계층 및 개념 계층으로 이루어진 다중 계층 비디오 모델인 MuVi를 제안하여, 내용 기반 검색과 주석 기반 검색을 일관성있게 통합하기 위한 기반을 제시하였다.

다음으로 비디오 검색에 관련된 연구로서, Mackay 등이 개발한 주석시스템인 EVA는^[14] 사용자가 주석을 달고 이를 이용하여 검색할 수 있다. 그러나 이 시스템에서는 사용자간의 주석 공유는 고려하지 않았다. Hirata와 Kato가 만든 내용 기반 이미지 검색 시스템인 QVE는^[9] 이미지로부터 추출한 외곽선 데이터를 검색에 이용하는데, 실제 질의 처리 과정에서 유사성을 검사하기 위해서 데이터베이스에 저장된 각각의 이미지에 대한 이동이나 스케일링, 회전등의 기하학적인 변환을 일일이 고려해

야 한다는 단점이 있다. IBM Almaden 연구소에서 개발한 QBIC은^[10] 이미지의 색상이나 질감, 모양 등의 다양한 속성에 기반한 시각적 질의를 제공하며, 사용자에게 데이터베이스에 있는 이미지에 키워드를 붙일 수 있게 하여 제한적인 주석 기반 검색이 가능하다. Adali와 Subrahmanian 등은^[12] 비디오 데이터에 대한 정형적 모델과 그러한 데이터를 효율적으로 저장하기 위한 공간 데이터 구조를 제공하고, 여러 종류의 비디오 질의를 처리하기 위한 알고리즘을 제시하였다. 또한 이러한 개념에 기반하여 AVIS(Advanced Video Information System)라는 프로토타입을 개발하였다.

그 외에도 많은 관련 연구들이 있는데 아직까지 그 어떤 연구도 표준적인 비디오 데이터 모델과 비디오 검색 방법으로는 자리잡지 못한 실정이다.

III. 복합 비디오 검색 기법

본 연구에서 제안하는 복합 비디오 검색은 내용 기반 비디오 검색과 주석 기반 비디오 검색을 함께 적용하여 사용자의 의도에 보다 가깝게 표현할 수 있고, 보다 정확한 검색 결과를 얻을 수 있다. 이 장에서는 먼저 기본 검색 기법인 주석 기반 비디오 검색과 내용 기반 비디오 검색을 각각 소개한 후에 그것들을 통합 확장한 복합 비디오 검색 기법을 소개한다.

1. 주석 기반 비디오 검색 기법

주석 기반 비디오 검색을 위해서는 먼저 비디오의 내용을 설명하는 주석 데이터를 작성해야 하는데, 본 연구에서는 다음과 같은 과정으로 주석을 작성한다.

1) 선택된 비디오를 재생시키면서 주석을 작성하기 위한 특정 구간으로서 시작 프레임 번호와 끝 프레임 번호를 결정한다.

2) 특정 구간에 대하여 (속성, 값)으로 구성된 주석 원소(들)를 지정한다. 속성은 주석의 의미를 설명하고 값은 실제 내용을 표현하는 것으로서, '속성 = 값'이란 의미를 갖는다. 지정된 주석 원소들은 제목, 프레임 구간 정보 등의 비디오 정보와 함께 하나의 비디오 단위로 DB에 저장 관리된다.

이와 같이 작성된 주석 DB를 이용하여 사용자가 원하는 비디오를 검색하는 과정은 다음과 같다.

1) 특정 비디오 집합 또는 전체 비디오 등과 같이 검색 범위를 결정한다.

2) 사용자가 원하는 비디오의 검색 조건을 다음 질의들을 이용하여 표현한다.

- (1) 주석 원소들 간의 AND, OR, NOT의 논리 연산자를 이용한 논리 질의
 - (2) 검색 조건으로 명시한 특정 비디오 단위와의 시간적 관계를 이용한 시간 관계 질의[8]
 - (3) 검색 조건으로 명시한 비디오 단위 내의 특정 속성값의 공유성을 이용한 의미 관계 질의^[2]
- 3) 검색 범위내에서 모든 검색 조건들을 만족하는 비디오 단위들을 결과로 검색된다.

2. 내용 기반 비디오 검색 기법

내용 기반 비디오 검색에는 일반적으로 다음과 같은 형태의 질의들이 사용된다^[6].

- 주어진 이미지와 유사한 키프레임(key frame)을 포함하는 비디오를 검색하라.
- 주어진 비디오 클립과 유사한 구간을 포함하는 비디오들을 검색하라.
- 특정 비디오 클립 내에서 질의 이미지와 유사한 프레임 검색하라.
- 특정 비디오 내에서 질의 비디오 클립과 유사한 구간을 검색하라.

본 연구에서는 가장 간단하면서도 검색 정확성이 뛰어난 첫번째 방식의 비디오 질의를 이용하여 내용 기반 비디오 검색 연구를 수행하였다. 이와 같은 비디오 검색을 위하여 먼저 다음과 같은 과정을 통하여 비디오 특징 데이터를 추출한다.

1) 비디오로부터 특징적인 키프레임들을 추출한다. 일반적으로 비디오에서 화면 전환이 일어나는 프레임을 특징적인 키프레임으로 결정한다. 이와 같은 키프레임들을 추출하는 과정을 비디오 파싱(parsing)이라고 하는데, 대표적인 방법으로는 연속된 두 프레임에 대한 픽셀(pixel)/블럭 비교법, 히스토그램 비교법, MPEG 비디오에서의 DCT 계수를 이용하는 방법, 부분대역(subband) 특징 비교법, 그리고 본 연구에서 사용하는 외곽선 데이터를 이용한 특징 비교법 등이 있다. 각 비디오 분할 방법에 대한 자세한 설명은 [7,11]를 참조하기 바란다.

2) 추출된 각 키프레임에 대하여 다음과 같은 과정으로 모양 특징 데이터를 추출한다.

- (1) 칼라 이미지의 키프레임 데이터를 그레이 이미지로 변환한다.
- (2) 그레이 이미지의 외곽선을 추출한다.
- (3) 외곽선 데이터에 대해 웨이블릿 변환을 이

용하여 특징 벡터를 추출한다.

최종적인 비디오 특징 데이터는 (키프레임 번호, 특징 벡터)로 구성된다. 추출된 특징 데이터를 이용하여 사용자가 원하는 비디오를 검색하는 과정은 다음과 같다.

1) 사용자가 원하는 장면을 직접 그리거나 유사한 예제 이미지를 선택하여 질의를한다.

2) 특징 추출 과정의 2)에서와 같은 과정으로 질의 이미지의 특징 데이터를 추출한다.

3) 키프레임의 특징 데이터와의 비교를 통하여 일정 기준 이상의 유사성을 갖는 것들을 검색 결과로 결정한다. 유사성 계산에는 질의 이미지와 키프레임의 특징 데이터 간의 유클리드 거리를 이용한다.

3. 복합 비디오 검색 기법

일반적으로 주석 기반 비디오 검색과 내용 기반 비디오 검색은 독립적으로 사용되지만 사용자의 의도를 명확하게 표현하기 위해서 이러한 두가지 검색 기법을 함께 사용해야만 하는 경우도 있다. 예를 들어, A라는 사람이 앉아 있는 장면을 포함하는 비디오를 찾으려는 경우를 생각해보자. 이 때에는 "이름 = A"라는 주석을 포함하는 비디오 검색과 앉아 있는 사람의 모습을 포함하는 비디오 검색을 함께 사용하는 것이 바람직하다. 이와 같은 경우를 위해 사용자가 기술한 비디오 설명을 주석으로 포함하는 동시에 사용자가 선택하거나 직접 그린 질의 이미지를 키프레임으로 포함하는 비디오를 검색하는 복합 비디오 검색 기법을 제안한다.

주석 기반 검색과 내용 기반 검색을 함께 지원하는 복합 비디오 검색을 처리하기 위해서 두가지 검색 기법을 한꺼번에 처리하는 방법과 순차적으로 처리하는 방법을 생각할 수 있다. 본 연구에서는 각 검색 기법의 독립성을 그대로 유지할 수 있으면서도 첫번째 검색만으로도 그 결과가 만족스러우면 검색 작업을 중간에 종료할 수 있다는 장점을 갖는 순차적 처리 방법을 채택하였는데 간단히 설명하면 다음과 같다.

1) 사용자는 두가지 검색 기법 중에서 어떤 검색 기법을 먼저 사용하는 것이 보다 유리한지 결정한다. 그 판단 기준으로는 다음과 같은 세가지를 생각할 수 있다.

(1) 검색 결과가 적은 것을 먼저 수행한다. 중간 결과의 크기가 적을 수록 전체 검색의 효율성이 좋아지기 때문으로 과거의 검색 경험

으로 예측할 수 있다.

(2) 질의 처리 시간이 짧은 것을 먼저 수행한다. 일반적으로 특징 추출이 필요한 내용 기반 검색이 보다 많은 시간이 걸린다.

(3) 질의를 작성하기가 보다 간단한 것을 먼저 수행한다. 일반적으로 사용자는 보다 쉬운 것을 먼저 수행하는 습성을 따른다.

2) 사용자가 보다 유리하다고 판단한 검색 기법을 A, 다른 검색 기법을 B라 할 때, A를 먼저 수행한다. 만일 A가 주석 기반 검색이면 그 결과로 질의 조건을 만족하는 비디오 단위들이 검색되고, 내용 기반 검색인 경우에는 키프레임들이 검색된다. 여기서의 검색 결과가 충분히 적은 양이거나 사용자에게 만족스러운 경우라면 다음 과정은 생략 가능하다.

3) A의 검색 결과에 대하여 B를 수행한다. 만일 A가 주석 기반 검색이면 결과 비디오 단위들로부터 키프레임들을 얻어서 그것들 중에서 질의 이미지와 유사한 것들을 검색하게 되고, A가 내용 기반 검색인 경우에는 결과 키프레임들을 포함하는 비디오 단위들을 얻은 후에 그것들 중에서 주석 질의를 만족하는 것들을 검색하게 되어 결국 어떤 경우이나 최종적으로는 주석 기반 질의와 내용 기반 질의를 모두 만족하는 비디오들을 검색할 수 있게 된다.

IV. TIVRON 설계

이 장에서는 먼저 비디오 데이터 및 검색에 필요한 여러 데이터를 SOP에 저장 관리하기 위한 클래스들을 설계하고, 그것들을 기반으로 복합 비디오 검색을 수행하는 통합 비디오 검색 시스템인 TIVRON의 구성 모듈들을 소개한다.

1. 관련 클래스 설계

본 연구에서 필요한 데이터로는 크게 나누어 MPEG 비디오 데이터, 주석 데이터, 그리고 내용 특징 데이터가 있는데, 그 각각에 대한 클래스 설계는 다음과 같다. 여기서 복합 비디오 검색은 기존 검색 기법에서의 데이터들만을 사용하기 때문에 별도의 클래스 설계는 필요없다.

1.1 MPEG 비디오 데이터 관련 클래스

DBMS가 MPEG 데이터와 같이 가변적이며 비정형적인 대용량의 멀티미디어 데이터를 처리하기 위해서는 BLOB 구조에 기반을 둔 타입이 필요하다.

SOP에서는 BLOB타입으로 LargeObj 클래스를 제공하는데, 본 연구에서는 그것을 기반으로 설계된 다음과 같은 클래스 스키마를 이용하여 MPEG 비디오 데이터를 지원한다(그림 1 참조).

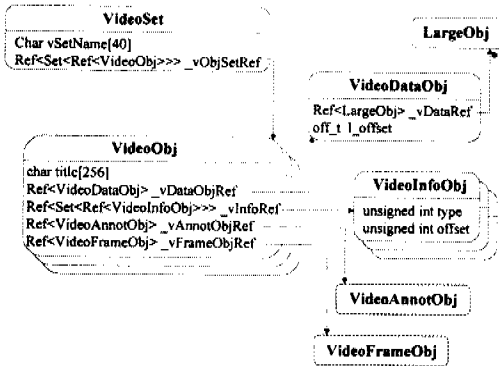


그림 1. MPEG 비디오 데이터 관련 클래스 스키마

- VideoDataObj : LargeObj 객체를 이용하여 물리적 데이터를 저장 관리한다.
- VideoInfoObj : MPEG 데이터를 논리적으로 나누어 관리한다. MPEG 데이터는 논리적으로 이미지 프레임들의 연속으로 볼 수 있는데, 각 프레임 타입과 물리적 데이터 내에서의 위치 등의 프레임 정보를 저장 관리한다.
- VideoObj : 하나의 비디오에 관련된 모든 정보들을 전체적으로 관리한다. 물리적 데이터와 논리적 데이터를 저장하기 위해 VideoDataObj 객체와 VideoInfoObj 객체를 각각 포함한다. 그 비디오에 대한 모든 주석과 내용 특정 데이터를 저장하기 위해 VideoAnnotObj 객체와 VideoFrameObj 객체를 각각 포함하는데 자세한 설명은 뒤에서 다룬다.
- VideoSet : DB내의 모든 비디오들에 관련된 정보를 관리한다. 따라서 VideoObj 객체들의 집합으로 구성되며, 새로운 비디오를 추가하거나 기존의 비디오를 삭제하는 기능, 비디오를 선택하는 기능 등을 제공한다.

1.2 주석 데이터 관련 클래스

본 연구에서는 주석 기반 비디오 검색을 위해 각 비디오마다 주석이 부여된 비디오 단위들을 관리하는 부분과 부여된 모든 주석들을 관리하는 부분으로 나누어 처리하는데, 하나의 비디오 단위에 속성과 값으로 구성되는 주석 데이터를 여러개 부여할 수 있으며, 그 주석의 값 또한 집합값을 가질 수 있

다. 이러한 주석 데이터와 관련된 클래스들은 다음과 같다(그림 2 참조).

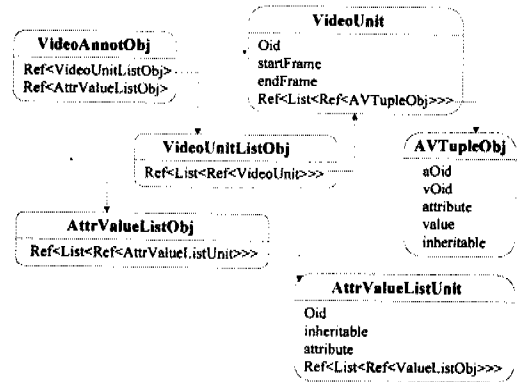


그림 2. 주석 데이터 관련 클래스 스키마

- VideoUnit : 비디오 내의 특정 구간을 하나의 비디오 단위로 관리하는데 그 구간에 대한 모든 주석들을 저장하며, 실제 검색의 대상 및 결과가 된다.
- AVTupleObj : 실제 주석 데이터를 저장하는 것으로서 속성과 값 및 그 식별자들, 그리고 상속결정인자로 구성된다. 상속결정인자는 두 개의 서로 다른 비디오 단위가 시간적으로 서로 겹칠 때 주석 데이터의 상속 여부를 결정한다.
- VideoUnitListObj : 하나의 비디오 내의 모든 비디오 단위들을 관리한다.
- AttrValueListObj : 하나의 비디오에 부여한 모든 속성들과 값들을 저장 관리한다. 사용자들이 주석 데이터를 공유할 수 있도록 하여 주석 데이터의 불일치를 제거하며, 속성과 값에 각각 고유한 식별자를 부여한다.
- AttrValueListUnit : 속성이 가질 수 있는 값들을 관리한다. 각 값은 ValueListObj 객체로 표현된다.
- VideoAnnotObj : 하나의 비디오에 부여된 모든 주석 관련 데이터를 관리하는 것으로서 VideoUnitListObj 객체와 AttrValueListObj 객체로 구성된다.

1.3 내용 특정 데이터 관련 클래스

본 연구에서는 내용 기반 비디오 검색을 위해 각 비디오마다 모든 키프레임들과 그 특정 벡터들을 다음과 같은 클래스들을 이용하여 저장 관리한다(그림 3 참조).

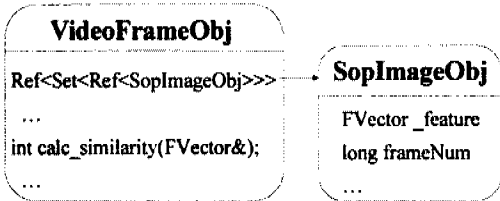


그림 3. 내용 특징 데이터 관련 클래스 스키마

- SopImageObj : 하나의 키프레임에 대한 관련 데이터를 저장 관리하는 것으로서 키프레임으로부터 추출한 특징 데이터와 프레임 번호를 포함한다.
- VideoFrameObj : 하나의 비디오 내의 모든 키프레임 객체들을 관리하는 것으로서 SopImageObj 객체의 집합을 포함하며 질의 이미지와의 유사성 계산을 위한 calc_similarity 멤버 함수를 갖는다.

2. TIVRON의 구성

본 연구에서의 TIVRON은 그림4에서와 같이 관련 데이터를 저장 관리하는 클라이언트/서버 구조의 SOP를 기반으로 MPEG 비디오 브라우저 시스템, 주식 기반 비디오 검색 시스템, 내용 기반 비디오 검색 시스템, 그리고 복합 비디오 검색 시스템으로 구성된다.

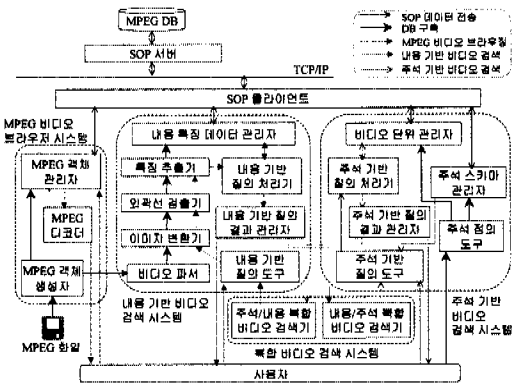


그림 4. TIVRON의 전체 구조

2.1 SOP OODBMS

SOP는 본 연구실에서는 지난 5년여에 걸쳐 개발한 ODMG-93 표준을 따르는 국내 최초의 다중 사용자용 OODBMS로서 ODL(Object Definition Language) 및 OQL(Object Query Language)을 지원하며 지속성 객체 지향 응용 프로그램의 작성을 위해 C++ 및 Smalltalk, CLOS 바인딩을 제공한다.

SOP는 효율적인 객체의 저장 및 관리 기능을 제

공하는 객체 지향 저장 시스템(Soprano)[4], 스키마 관리 및 진화를 담당하는 스키마 관리자, 사용자 질의를 처리하는 OQL 처리기, 그리고 객체 지향 데이터베이스 언어 전위처리기 등의 엔진 모듈과 그래픽 환경에서의 스키마 정의 및 질의 처리 등을 지원하기 위한 다양한 사용자 지원 도구들로 구성되어 있다.

2.2 MPEG 비디오 브라우저 시스템 구성

MPEG 비디오 브라우저 시스템은 다음과 같은 모듈들로 구성된다.

- MPEG 객체 생성자 : MPEG 화일로부터 MPEG 관련 객체들을 생성하여 MPEG 객체 관리자에게 넘겨준다. 여기서 물리적 데이터뿐만 아니라 각 프레임의 타입 및 위치 등의 논리적 데이터까지 생성한다.
- MPEG 디코더 : 선택된 MPEG 데이터에 대하여 재생, 정지, 빨리보기, 이전/다음 프레임 보기, 임의 장면 보기 등 다양한 조작들을 처리한다.
- MPEG 객체 관리자 : MPEG 관련 객체를 SOP DB로 구축하고 전체 비디오 목록을 유지한다. 또한 비디오 목록이나 검색 결과로부터 선택된 비디오에 관련된 데이터를 SOP로부터 메모리로 읽어들이어 관리한다.

2.3 주식 기반 비디오 검색 시스템 구성

주식 기반 비디오 검색 시스템은 다음과 같은 모듈들로 구성된다.

- 주식 정의의 도구 : 비디오의 특정 구간(속성, 값)으로 표현되는 주식 원소(들)을 지정하여 주석을 정의하도록 한다.
- 주식 스키마 관리자 : 정의된 주식 원소들을 저장 관리한다. 이 도구를 이용하여 주석을 참조함으로써 주석 작성의 일관성을 유지하거나 효율적인 질의 작성이 가능하다.
- 주식 기반 질의 도구 : 검색 범위와 주식 원소(들)로 구성된 검색 조건으로 주식 기반 질의를 작성한다. 작성된 질의는 주식 기반 질의 처리기로 넘겨진다.
- 주식 기반 질의 처리기 : 질의가 올바르게 표현되었는지 검사한 후에 올바른 경우라면 SOP가 관리하는 비디오 DB의 지정된 검색 범위내에서 주어진 검색 조건을 모두 만족하는 비디오 단위들을 검색한다.
- 주식 기반 질의 결과 관리자 : 질의와 그 처리 결과로 검색된 비디오 단위들을 저장 관리한다. 추

후에 사용자가 참조함으로써 자주 검색되는 질의에 대해서 발생하는 질의의 중복 작성 및 실행을 제거할 수 있다. 또한 검색 결과를 사용자가 원하는 결과에 보다 가깝게 정제하기 위해 주석/내용 복합 비디오 검색기의 입력으로 사용할 수도 있다.

- 비디오 단위 관리자 : 지정된 비디오의 모든 비디오 단위들을 관리하여 하나의 연속된 비디오로 추상화한다. 또한 주석 작성 과정에서 생성된 비디오 단위를 SOP DB로 구축하고, 질의 처리에 필요한 비디오 단위를 SOP로부터 읽어들이어 관리하는 등 주석 기반 비디오 검색에서 필요한 SOP와의 모든 통신에 작업을 담당한다.

2.4 내용 기반 비디오 검색 시스템 구성

내용 기반 비디오 검색 시스템은 다음과 같은 모듈들로 구성된다.

- 비디오 파서 : 연속하는 두 프레임에 대한 외곽선 데이터의 특징 비교법을 이용하여 비디오의 특징적인 키프레임들을 추출한다. 이 작업은 많은 시간이 필요하므로 오프라인으로 수행된다.
- 이미지 변환기 : 추출된 키프레임들을 그레이 이미지로 변환한다.
- 외곽선 검출기 : 변환된 그레이 이미지의 외곽선을 추출한다.
- 특징 추출기 : 추출된 외곽선 데이터에 웨이블릿 변환을 적용하여 특징 데이터를 계산한다.
- 내용 기반 질의 도구 : 원하는 이미지를 직접 그리거나 유사한 예제 이미지를 선택하여 질의 이미지를 표현한다. 질의 이미지는 일련의 특징 추출 과정을 거친 후에 특징 데이터 형태로 내용 기반 질의 처리기에게 전달된다.
- 내용 기반 질의 처리기 : 질의 이미지의 특징 데이터와 유사도가 일정 수준 이상의인 특징 데이터를 갖는 키프레임들을 검색한다.
- 내용 기반 질의 결과 관리자 : 질의와 그 처리 결과로 검색된 키프레임들을 관리하여 사용자가 참조할 수 있도록 한다. 또한 검색 결과를 사용자가 원하는 결과에 보다 가깝게 정제하기 위해 내용/주석 복합 비디오 검색기의 입력으로 사용할 수도 있다.
- 내용 특징 데이터 관리자 : 지정된 비디오의 모든 내용 특징 데이터를 관리하여 하나의 연속된 비디오로 추상화한다. 또한 특징 데이터 생성 과정으로부터 생성된 키프레임 데이터를 SOP DB로

구축하고, 질의 처리에 필요한 특징 데이터를 SOP로부터 메모리로 읽어들이어 관리하는 등 내용 기반 비디오 검색에서 필요한 SOP와의 모든 통신 작업을 담당한다.

2.5 복합 비디오 검색 시스템 구성

복합 비디오 검색 시스템은 다음과 같은 두개의 모듈로 구성된다.

- 주석/내용 복합 비디오 검색기 : 주석 기반 비디오 검색의 처리 결과에 대하여 내용 기반 비디오 검색을 수행하도록 한다.
- 내용/주석 복합 비디오 검색기 : 내용 기반 비디오 검색의 처리 결과에 대하여 주석 기반 비디오 검색을 수행하도록 한다.

V. TIVRON 구현

TIVRON은 본 연구실에서 기존에 개발한 시스템들을 확장 통합하여 구현되었다. 이 장에서는 먼저 각 시스템의 기반 연구 및 확장 작업에 대하여 설명한 후에 최종적으로 복합 비디오 검색을 지원하는 검색 시스템으로의 통합에 대하여 설명한다.

1. MPEG 비디오 브라우저 시스템 구현

본 연구실에서는 SOP 상에 MPEG 비디오 데이터를 저장 관리하며 사용자가 선택한 비디오 데이터에 대한 다양한 조작 기능을 제공하기 위한 연구를 수행하여 그 결과로 최근에 확장 용이 MPEG 비디오 브라우저 시스템을 발표하였다³⁾.

MPEG 비디오 브라우저 시스템에서는 MPEG 디코딩을 위하여 UC.Berkeley에서 개발한 MPEG-1 디코더 라이브러리 버전 2.2를 기반으로 사용하였다. 이 라이브러리는 C와 X윈도우 프로그램으로 개발되어 본 시스템 개발 환경인 C++ 및 Motif용 프로그램과의 통합 작업이 필요하였다. 또한, MPEG 데이터의 저장 관리를 위해 SOP와 통합하였으며, 이전/다음 프레임 보기, 임의 장면 보기 등의 새로운 기능들을 추가하였다. 기존의 MPEG 비디오 브라우저 시스템은 SOP에서 저장 관리되는 전체 비디오 목록을 보여주고 원하는 제목의 비디오를 선택하는 단순한 비디오 선택 방법을 제공하는데, 대용량의 비디오 DB로부터 사용자가 원하는 비디오를 효율적으로 찾기 위해서는 본 연구에서와 같은 여러가지 효율적인 비디오 검색 기법의 추가가 필요하다.

2. 주석 기반 비디오 검색 시스템(VIRON-II) 구현

본 연구실에서는 SOP 상의 주석 기반 비디오 검색 연구를 수행하여 그 결과로 VIRON이라는 주석 기반 비디오 검색 시스템을 발표하였다[2]. 일반적으로 주석 기반 검색 시스템은 여러 사용자들이 각자 주관적으로 주석을 부여하기 때문에 주석들 간에 의미의 통일성을 잃을 수 있다는 단점이 있는데, 그로 인한 폐해를 줄이기 위해서 VIRON에서는 다음과 같은 기법들을 제공한다.

- 동적인 속성 스키마 지원 : 속성들을 관리하기 위하여 객체 지향 개념에 기반한 계층적 구조의 하나의 속성 스키마를 구성하고 모든 사용자가 공동으로 그것을 참조하여 주석 데이터를 정의하도록 함으로써 주석 작성의 일관성을 유지할 수 있다.
- 비디오 인덱스 스트림의 구성 : 한 비디오에 대한 모든 주석 정보를 단일한 시간축을 갖는 비디오 인덱스 스트림으로 맵핑시켜 관리함으로써 비디오 단위의 중복을 제거할 수 있다.

기존의 VIRON은 비디오 데이터를 비롯하여 주석 데이터, 비디오 관리 데이터 등 거의 모든 데이터를 SOP의 LargeObj에 저장하여 결국 SOP를 단순한 저장 시스템으로 이용했을 뿐 실제적인 데이터 해석 및 처리 등 많은 부분을 응용 프로그램에서 직접 처리해야만 했다. 그 결과 SOP가 제공하는 효율적인 데이터 저장, 검색 및 관리 기능 등을 거의 사용하지 못하여 시스템의 성능 저하 문제가 발생하였고, 시스템도 커져서 유지 보수 및 확장에 어려움이 따랐다.

이와 같은 단점들을 해결하기 위해서 본 연구에서는 주석 기반 검색과 관련된 데이터를 그림 2에서와 같이 SOP 상의 클래스로 정의하여 SOP의 효율적인 저장, 검색 및 관리 기능을 이용하도록 하였으며, 기존의 VIRON이 포함하고 있던 자체적인 비디오 재생기 모듈을 MPEG 비디오 브라우저 시스템으로 대체함으로써 보다 구조화된 시스템인 VIRON-II로 재구성하였다.

3. 내용 기반 비디오 검색 시스템(SCARLET-V) 구현

본 연구실에서는 최근에 SOP 상의 내용 기반 이미지 검색 연구를 수행하여 그 결과로 SCARLET이라는 이미지 검색 시스템을 발표하였다⁵⁾.

SCARLET은 모양 특징을 이용하여 SOP 상의

이미지 DB로부터 질의 이미지와 유사한 이미지들을 검색하는 시스템으로서, 이미지의 모양 특징 추출을 위해서 그레이 이미지로의 변환, 외곽선 추출, 웨이블릿 변환 과정을 차례로 거치며 그 결과로 추출된 특징 데이터는 이미지와 함께 SopImageObj 객체로 생성되어 SOP에서 저장 관리된다. 사용자가 질의 도구를 이용하여 원하는 이미지를 직접 그리거나 비슷한 예제 이미지를 선택하면 그 이미지의 모양 특징 데이터를 앞에서 설명한 과정으로 추출하여 SOP 이미지 DB 상의 유사한 이미지들을 검색하여 보여지게 된다.

SCARLET을 비디오 검색 시스템으로 확장하기 위해서는 먼저 비디오 데이터로부터 내용으로 추정되는 데이터를 검출하는 비디오 파싱 과정을 추가해야 하는데, 본 연구에서는 Cornell 대학에서 제안한 특징에 기반한 방법으로 구현한 비디오 파서를 그림 4에서와 같이 이미지 변환기의 입력부에 두었다. 또한 비디오 파싱 과정으로 검출된 키프레임들로부터 특징 데이터를 추출하여 SOP에 저장 관리해야 하는데, 이를 위해서 그림 3에서와 같이 VideoFrameObj 클래스를 추가하고 기존의 SopImageObj 클래스에 이미지 데이터를 삭제하는 대신 프레임 번호 정보를 추가하였다. 마지막으로 검색 결과로 얻어진 비디오를 조작할 수 있도록 MPEG 비디오 브라우저 시스템을 연결시켰다. 이와 같이 내용 기반 비디오 검색 시스템으로 확장된 SCARLET을 SCARLET-V라 한다.

4. 통합 비디오 검색 시스템의 구현

본 연구에서는 이제까지 살펴본 MPEG 비디오 브라우저와 VIRON-II, SCARLET-V의 독립성을 유지하면서 복합 비디오 검색 시스템으로의 상호 작동이 가능하도록 통합하여 SOP 기반의 통합 비디오 검색 시스템인 TIVRON을 구현하였다.

먼저 각 시스템은 모두 SOP에 기반을 두고 있기 때문에 그림 1에서와 같은 클래스 스키마를 정의하여 모든 관련 데이터를 SOP 상에서 저장 관리하도록 함으로써 자연스러운 데이터 통합이 가능하였다. 다음으로 MPEG 비디오 브라우저에서 사용자가 VIRON-II와 SCARLET-V를 이용할 수 있도록 함으로써 주석 기반 비디오 검색과 내용 기반 비디오 검색을 가능하도록 하였으며, VIRON-II와 SCARLET-V의 검색 결과 비디오를 MPEG 비디오 브라우저를 이용하여 조작할 수 있도록 하였다. 마지막으로 VIRON-II와 SCARLET-V를 함께 이용하

여 복합 비디오 검색을 지원한다. VIRON-II로부터 얻은 검색 결과에 대하여 SCARLET-V를 수행할 수 있도록 함으로써 주석/내용 복합 비디오 검색을 지원하며, 또한 SCARLET-V의 검색 결과에 대해서 VIRON-II를 수행할 수 있도록 함으로써 내용/주석 복합 비디오 검색을 지원한다.

VI. 복합 비디오 검색 예

본 장에서는 TIVRON을 이용한 복합 비디오 검색을 간단히 보여준다. 예를 들어, 사용자가 "C++_lect_1"이라는 비디오 내에서 "Bill이라는 이름을 갖는 화자의 상반신을 보여주는 모든 장면"을 찾고자 하는 경우를 생각해보자. 이 경우에는 "Bill" 값을 갖는 '이름' 속성을 주석 데이터로 포함하는 동시에 상반신을 표현하는 특징 데이터를 함께 포함해야 하므로 복합 비디오 검색 기법을 사용해야 한다. 여기서는 주석 기반 비디오 검색을 먼저 사용하는 경우를 살펴보기로 하자.

먼저 화자가 Bill이라는 주석을 포함하는 모든 비디오 단위를 찾아야 하는데, 그림 5에서 TIVRON을 이용한 질의와 그 처리 결과를 보여준다. 뒤에 있는 윈도우가 질의 도구로서 검색 범위를 포함하는 From 영역에 비디오 제목으로 "C++_lect_1", 그리고 검색 조건을 포함하는 Where 영역에 속성과 그 값으로서 "speaker = Bill"로 표현된 질의를 볼 수 있다. 그 처리 과정은 그림 4의 주석 기반 비디오 검색 흐름과 같은데, 그 처리 결과로 5개의 비디오 단위가 검색되었다는 것을 그림 5의 앞에 있는 결과 윈도우에서 보여준다.

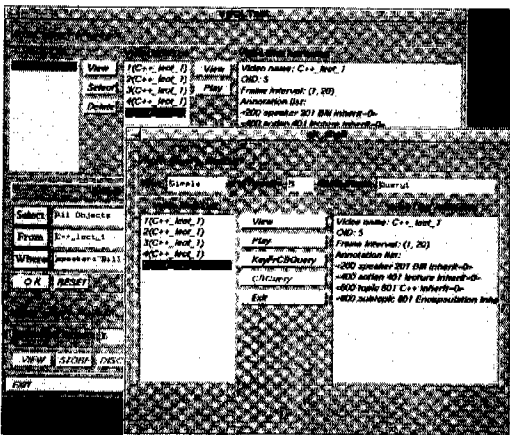


그림 5. TIVRON에서의 주석 기반 검색 및 결과

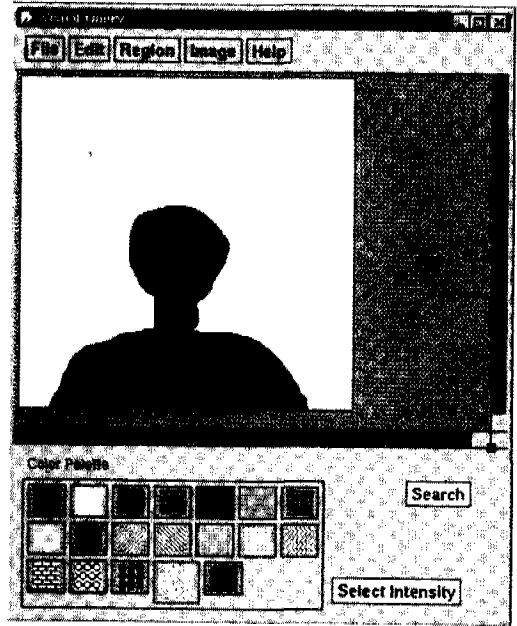


그림 6. 사용자 그림에 의한 내용기반 검색

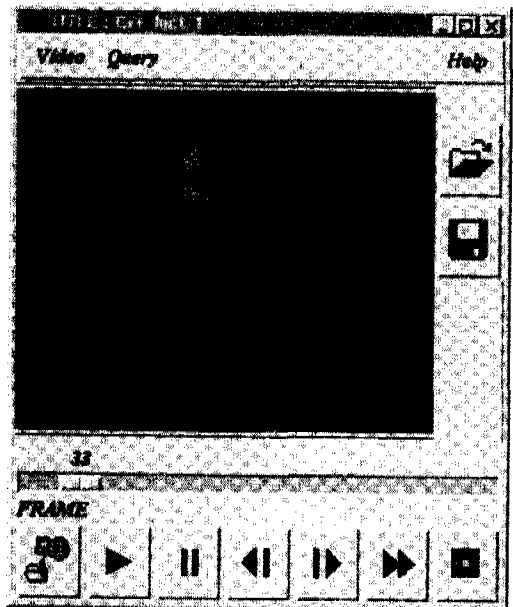


그림 7. TIVRON의 검색 결과 조작 예

다음으로 이 비디오 단위들이 포함하는 모든 프레임들을 대상으로 상반신이 나타나는 장면을 찾는 검색을 수행해야 한다. 그림 5의 결과 윈도우에서 CBQuery 버튼을 누름으로써 검색 결과에 대한 내용 기반 비디오 검색이 실행되는데, 그림 6에서는 사용자 그림에 의한 질의로서 사용자가 질의 도구

를 이용하여 상반신을 직접 그린 질의 이미지를 보여준다. 그림 6에서 Search 버튼을 누름으로써 내용 기반 검색이 수행되는데, 그 처리 과정은 그림 4의 내용 기반 비디오 검색 흐름을 따른다. 그 결과로 결국 C++_lect_1라는 제목의 비디오 내에서 Bill이라는 이름을 갖는 화자의 상반신을 보여주는 모든 키프레임들을 보여주게 된다.

사용자는 검색 결과 중에서 원하는 것을 선택하여 TIVRON 내의 MPEG 비디오 브라우저로 조작할 수 있다. 그림 7은 결과 중의 한 장면을 선택하여 MPEG 비디오 브라우저로 조작하는 예를 보여주는데, 현재 C++_lect_1라는 제목의 비디오 중에서 33번째 프레임들을 보여주고 있다.

VII. 결론 및 향후 연구

본 논문에서는 OODBMS 상에 구축된 MPEG 비디오 데이터베이스로부터 사용자가 원하는 비디오를 보다 효율적으로 찾기 위해 비디오 검색 분야에서의 일반적인 두가지 기법인 주석 기반 검색과 내용 기반 검색을 독립적으로 지원할 뿐만 아니라 두 기법들을 함께 사용한 복합 비디오 검색까지 지원하는 통합 비디오 검색 시스템으로서 TIVRON을 제안하였다.

비정형적이며 대용량의 비디오 데이터베이스를 효율적으로 저장 관리하기 위해서는 OODBMS의 사용이 바람직하다. 본 연구실에서 개발한 SOP OODBMS는 비디오 데이터의 효율적인 저장 관리를 지원하는데, 그 결과로 구축된 비디오 데이터베이스를 이용하면 여러 응용 프로그램의 개발 및 확장이 용이하다는 장점을 갖는다. 본 연구에서는 SOP 상의 MPEG 비디오 데이터 및 검색에 관련된 데이터베이스를 구축하기 위하여 통합적인 클래스 스키마를 설계하였고, 주석 기반 비디오 검색을 위해서 기존의 주석 기반 비디오 검색 시스템인 VIRON을 VIRON-II로 효율적으로 재구성하였으며, 내용 기반 비디오 검색을 위해서 기존의 내용 기반 이미지 검색 시스템인 SCARLET을 내용 기반 비디오 검색 시스템인 SCARLET-V로 확장하였다. 또한, 그러한 두가지 검색 기법을 함께 지원하는 복합 비디오 검색을 위해서 각 검색 시스템의 검색 결과에 다른 검색 시스템을 적용하는 순차적 처리 방법을 구현하였으며, 마지막으로 이러한 여러가지 검색 기법의 결과를 MPEG 비디오 브라우저를 이용하여 다양하게 조작할 수 있도록 하였다.

하지만 본 연구에서의 TIVRON은 하나의 비디오 데이터 모델을 기반으로 개발된 것이 아니라 여러 시스템들을 통합한 것이기 때문에 새로운 기법의 비디오 검색 기법의 추가나 다른 타입의 멀티미디어 데이터 검색 시스템으로의 확장이 어렵다는 단점이 있으며, 내용 기반 검색에서는 비디오 데이터의 가장 큰 특징 중의 하나인 시간성을 고려하지 않았다는 단점이 있다.

향후 과제로는 일관된 비디오 데이터 모델을 연구하여 그것을 기반으로 여러가지 다양한 검색 기법들을 쉽게 추가할 수 있는 비디오 검색 시스템으로의 재구성 작업과 비디오의 시간성에 기반한 내용 기반 검색의 추가 작업이 필요하다. 아울러 현재 검색에 관련된 데이터로의 효율적인 접근에 필수적인 인덱스를 고려하지 않았는데, 본 연구의 검색 기법에 따른 효과적인 인덱스의 연구 및 적용 방법 연구도 바람직하다.

참고 문헌

- [1] 김기병, 김형주, "내용 기반 검색 및 주석 기반 검색을 통합하는 비디오 데이터 모델의 설계 및 구현", *한국정보과학회논문지(C)*, 3(2), pp. 115-126, 1997
- [2] 김기욱, 김형주, "비디오 주석 시스템의 설계 및 구현", *한국정보과학회논문지*, 24(6), June 1997.
- [3] 송용준, 김형주, "OODBMS 기반의 확장 용이 MPEG 비디오 브라우저", *한국정보과학회논문지(C)*, 4(6), December 1998.
- [4] 안정호, 이강우, 송하주, 김형주, "Soprano: 객체 저장 시스템의 설계 및 구현", *한국정보과학회논문지(C)*, 2(3), September 1996.
- [5] 이동호, 송용준, 김형주, "SCARLET: 웨이블릿 변환을 이용한 내용 기반 이미지 검색 시스템의 설계 및 구현", *한국정보과학회논문지(C)*, 3(4), August 1997.
- [6] A. Gupta, R. Jain, "Visual Information Retrieval," *Communications of the ACM*, 40, pp. 71-79, May 1997.
- [7] G. Ahanger, T.D.C. Little, "A Survey of Technologies for Parsing and Indexing Digital Video," MCL Technical Report, 11-01-95,

