

主 題

방송용 영상 데이터베이스

한국전자통신연구원 정보검색연구팀 유 성 준

차 례

- I. 개요
- II. 방송에서의 영상 데이터베이스 이용 예
- III. 영상 데이터베이스의 기본 구조
- IV. 관련 기술 동향
- V. 영상 데이터베이스의 설계
- VI. MPEG-7의 범위 (14, 15)
- VII. 결어

1. 개요

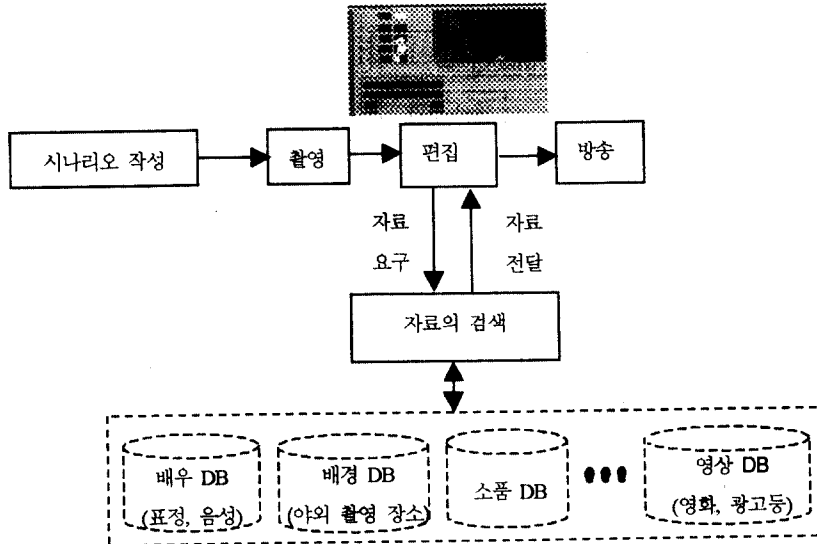
컴퓨터를 기반으로 한 비디오 처리 기술의 발달은 비디오의 촬영, 저장, 편집, 전송, 재생 등의 영역에 디지털 기술을 폭 넓게 사용할 수 있는 길을 열어 놓고 있다. 이러한 디지털 비디오 기술의 발달로 아날로그 미디어에 저장한 뒤에 다시 디지털로 변환하는 것보다는 직접 디지털 미디어에 저장하는 것이 오히려 편리하게 된다. 디지털 영상 시대로 접어들면서, 기존에 비디오 자료를 필름의 형태로 보관하고 있는 것들도 디지털 방식으로 저장하여 보다 손쉽게 검색할 수 있는 기능이 필요하게 되었다. 이는 기존의 자료를 이용하여 다큐멘터리 비디오나 교육용 비디오를 만들 때 등에 절실히 요구되는 기술이라고 하겠다. 아직은 데이터의 양이 방대하고 이를 저장하고 처리하기 위한 저장 매체 및 서버의 가격 등이 고가인 이유로 비디오 저장 및 검색 시스템의 구축이 그 초기 단계에 있지만 향후 디지털 방송과 같은 사회적 환경에 의해서 그에 대한 요구는 증가할 것이다. 이 글에서는 방

송의 제작이나 송출에 필요한 데이터를 효율적으로 보관하고 손쉽게 찾기 위한 방송 데이터베이스 기술의 동향에 대하여 언급한다. 2절에서는 방송에서의 영상 데이터베이스 이용예를, 3절에서는 기본 구조, 4절에서는 관련 기술 동향, 5절에서는 데이터베이스 설계시 고려 사항, 6절에서는 MPEG-7에 대하여 기술한다.

2. 방송에서의 영상 데이터베이스 이용 예

드라마나 다큐멘터리와 같은 방송물을 제작하기 위해서는 먼저 시나리오를 작성하고 필요한 부분을 촬영한 뒤에 기존에 보관하고 있는 영상 자료를 검색하여 필요한 부분에 편집, 삽입하게 된다. 이때 기존의 방법은 영상 자료를 검색하기 위하여 자료실에 필요한 비디오 테이프를 요구하게 된다. 자료실에서는 키워드 정보 검색 시스템을 이용하여 관련 비디오 테이프의 번호를 찾아내고 이 비디오 테이프를 편집실 쪽으로 전달한다. 이러면 편집실에서는 필요한 부분을 발췌, 편집하거나 원하는 비

(그림 1) 방송 데이터베이스의 이용예



디오 테이프가 아니면 다른 요구 사항을 내게 된다. 이렇게 수동으로 자료가 왔다 갔다 하는 시간과 요구 사항의 정확성 문제 때문에 걸리는 시간은 전체 방송 제작 시간에서 많은 비중을 차지하게 된다. 즉, 검색의 정확하고 신속한 검색과 자료의 신속한 전달, 신속한 요구 사항의 수정에 의한 자료의 갱신 등이 자료가 디지털화되고 네트워크가 구축됨으로써 효율적으로 이루어질 수 있을 것이다.

디지털 스튜디오에서 사용하는 비디오 데이터는 단일 미디어 형태로 존재하거나 또는 다른 멀티미디어 데이터와 결합된 형태로 존재하기도 한다. 영화를 만들기 위하여 촬영한 필름을 디지털화하거나 디지털 카메라로 찍은 내용을 디지털 스튜디오에서 편집하기 위해서는 수동이나 자동으로 비디오 데이터들을 임의의 조각으로 자르게 된다. 이렇게 획득한 비디오 데이터는 단일 미디어 형태로 존재하는 것이다. 그러나 이미 합성된 멀티미디어 데이터도 편집에 이용할 수 있을 것이다.

아래의 그림 1은 방송국에서의 데이터베이스 이용 예를 나타낸 것이다. 여기에는 영상 데이터베이스 외에도 배우에 대한 자료, 배경 자료, 소품 관련 자료 등도 저장하고 있음을 나타내고 있다. 이들은 모두 방송 제작을 할 때에 검색할 수 있다.

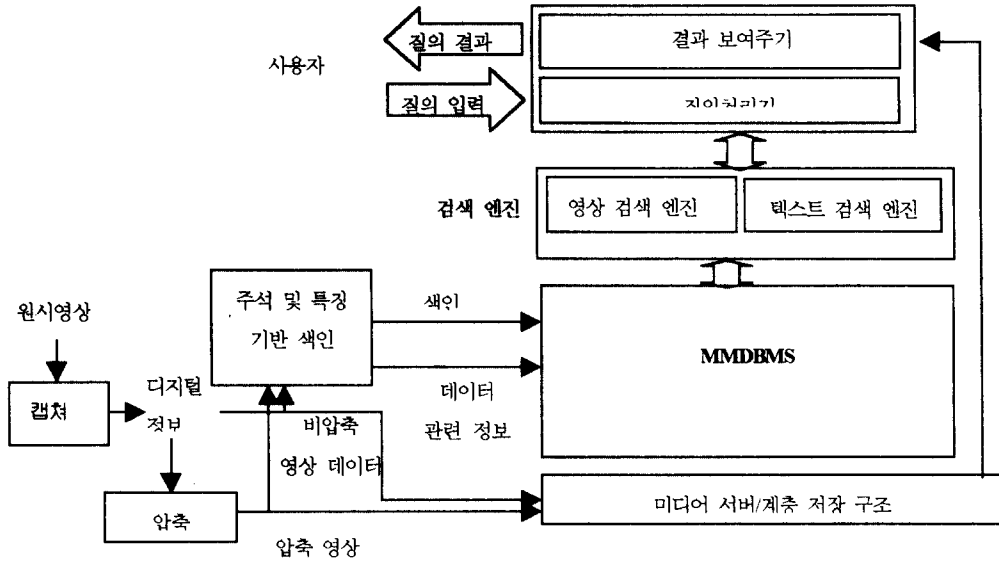
3. 영상 데이터베이스의 기본 구조

그림 2는 영상 데이터베이스의 내부 구조를 나타낸 것이다. 즉, 영상 데이터베이스 내에는 영상 자료를 디지털화해서 입력하는 기능, 입력한 자료를 압축하는 기능, 데이터에 주석 색인을 붙이거나 추가하고, 특징을 추출하는 기능, 저장 기능, 검색 기능 등이 이 시스템을 이루는 주요 기능이다.

4. 관련 기술 동향

방송에서 이러한 데이터베이스를 이용하는 국내의 예는 MBC, KBS, SBS, 아리랑 TV 등이 자체적으로 이미 방송된 프로그램을 세그멘테이션하고 품질이 낮은 데이터로 바뀌어서 인터넷으로 서비스하거나 광고와 같은 일부 데이터들을 DB화해놓고 이용하는 것 외에 주로 실험실에서 이루어진 몇 가지 연구들이 있다. 한국전자통신연구원의 정보검색 연구팀에서는 비디오 및 이미지 검색 시스템 구축 솔루션을 개발하고 있다. 대한 뉴스 및 수천장의 이미지를 저장하여 검색하고자 하는 시스템을 개발하면서 내부 구성 요소인 키팩트 기반 텍스트 정

(그림 2) 영상 데이터베이스의 기본 구조



보 검색 기능, 형태소 분석 기반 색인 기능, 사전 구축, 특징 기반 상표 검색 기능, MPEG의 장면 전환 인식 기능, MPEG의 비디오 요약 기능, MPEG 비디오 특징 기반 검색 기능, JPEG 기반 이미지 검색 기능 등을 개발하고 있다. 향후 음향 및 그래픽 데이터의 검색 기능 및 웹 기반 검색 기능으로 확장할 계획으로 있다.

이 절에서는 뉴스 데이터베이스 예를 먼저 기술하고 이외에 제반 영상 데이터베이스 시스템이나 구축 도구 등에 대하여 간략히 기술한다.

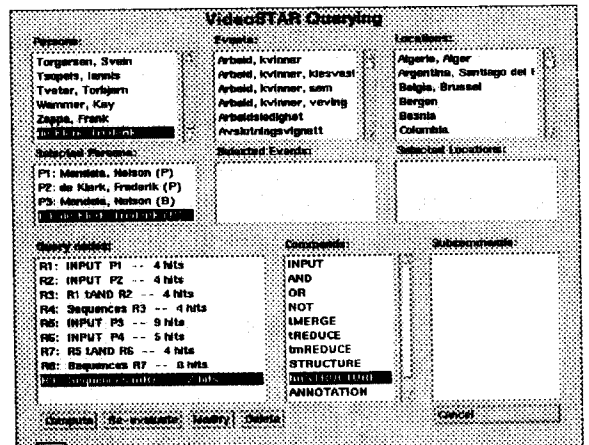
4.1 뉴스 데이터베이스의 예

영국 Olivetti[1]연구소의 뉴스 검색 시스템은 BBC1의 9시 뉴스를 저장하여 검색하고 실시간으로 재생해주는 실험 시스템을 구축하였다. 여기에서는 장면 전환이 일어나는 곳을 수동으로 찾아내는 기능을 지원하고 있다. 뉴스에서 전달하는 텍스트의 내용은 텔레텍스트로 전달되는데 이것을 수신하여 해당 뉴스 비디오 구간과 연관시켜 놓는다. 이것은 비디오 구간에 주석을 붙이는 것과 같은 효과를 낸다. 이 텍스트, 즉 주석 색인을 바탕으로 키워드 검색을 통하여 해당 비디오 구간을 찾도록 하

였다.

노르웨이 공대의 Hjelvold[2]라고 하는 사람은 비디오데이터베이스에 관련된 시스템 구조, 비디오 데이터의 저장, 전송, 비디오 데이터 모델, 색인, 검색 등에 대하여 연구하였는데 여기에서 실험적으로 사용한 데이터 역시 뉴스 비디오 데이터이다. 실험적으로 개발한 시스템의 사용자 인터페이스는 다음 그림과 같다. 여기에서는 비디오 데이터에 키

(그림 3) VideoSTAR의 사용자 인터페이스



워드로 색인을 해 놓고 사람, 사건, 장소 등의 분류에서 인스턴스를 선택하여 필요한 자료를 검색하도록 하였다.

미국의 디지털 라이브러리 관련 사업의 일환으로 CMU에서 개발한 Infomedia[3] 역시 TV 뉴스를 실험 데이터로 사용하였다. 음성으로 질의를 하게 되면 이 내용이 텍스트로 표시된다. 이후에 관련 비디오 클립이 나타나게 되고 여기에서 원하는 클립을 선택하면 이 데이터가 재생된다. 이렇게 재생이 되면 이 비디오 데이터와 관련된 텍스트도 아울러 재생되도록 하였다. 이 시스템은 기존의 다른 시스템의 기능에 음성 인식 기능을 추가하여 제공하고 있는 것이 특징이다.

Smoliar[4]는 비디오 파싱을 하고 장면마다 조각을 낸 후에 검색하는 주제에 관한 연구를 수행하였는데 이의 실험은 뉴스 비디오를 대상으로 수행하였다.

다음 절에서는 영상 데이터베이스 구축 예와 구축 도구 등에 관하여 기술한다.

4.2 기타 영상데이터베이스 구축 예 및 구축 도구

(1) QBIC[5]

IBM에서 개발한 정지영상 및 동영상 검색 엔진으로써 특징 기반 검색 기능이 추가 되고, 주석 기반 검색 기능이 추가된 시스템이다. 해가 거듭됨에 따라 그 기능이 향상되고 있다. 여기에 최근에는 장면 전환 검색 기능과 스토리보딩에 의한 비디오 요약 출력 기능도 첨가하였다.

(2) Virage[6]

Virage사에서 개발한 API가 갖추어진 텍스트, 정지영상 및 동영상 검색 엔진이다. ORACLE, Informix 등이 자체 DBMS에 멀티미디어 확장 모듈로 채택해서 판매하고 있을 정도로 시스템의 안정성이 뛰어나고 범용 기능을 갖고 있다.

(3) Columbia 대학의 VisualSEEK[7]과 WebSEEK[8]

영상 정보 검색 시스템에 관한 연구 개발 기관으로서 빼 놓을 수 없는 곳이 미국의 Columbia 대학이다. 여기에서는 독립적으로 수행되는 검색 엔진인 VisualSEEK을 개발한 바 있고 이를 발전시켜 인터넷에서 영상 정보를 검색할 수 있는 사용자 인터페이스를 제공하고 있다. 초기 화면에는 저장하고 있는 데이터들을 분류한 메뉴 화면이 뜨고 이를 통해 주제별로 검색을 할 수 있도록 해 놓았다. 이와 함께 제공되고 있는 검색 지원 기능은 한 단어를 집어 넣어 검색할 수 있는 키워드 검색 기능이 있고, 특징을 기반으로 검색할 수 있는 SAFE라는 엔진을 포함하고 있다. 여기에서는 이미지나 비디오 상에 있는 개체들의 공간 관계나 그들의 특징을 기반으로 검색할 수 있는 기능을 지원한다. 1996년도에 조사된 바에 의하면 전체 검색 건수 828,802건 중에 분류 메뉴 기반 검색이 354,512건으로 42.8%를 차지하고 있고 내용을 보면서 브라우징한 것은 189,512건으로 22.9%를 차지하고 있다. 이외에 주제 기반 질의가 175,359건으로 전체의 21.2%, 텍스트 기반 질의가 120,832건으로 14.6%를 차지한다. 의외로 특징 기반 검색은 전체의 4% 정도에 그쳤다.

(4) 기타

이외에 버클리 공대에서 개발한 POSTGRES를 기반으로 해서 개발한 Chabot[9]이라는 시스템이 있는데 이 역시 특징 기반 검색 기능이 추가 되는 시스템이다. 이것은 캘리포니아 수자원관리 조직에서 소장하고 있는 이미지를 저장하고 검색하는 것을 위주로 시스템을 개발했는데 하나의 이미지를 고해상도의 이미지부터 다섯 단계의 해상도를 갖도록 하여 서비스하도록 하였다. 이에 따라 방대한 저장 용량이 필요하였고 관련 연구도 진행되었다. 질의 항목은 일반적인 서지 정보 외에 '지역', '색', '개념' 등을 입력할 수 있다. MIT 공대에서 개발한 이미지 검색 시스템으로 Photobook[10]이라는 것이 있다. 모양, 외형, 질감 특징을 기반으로 검색할 수 있다.

5. 영상 데이터베이스의 설계

영상 데이터베이스는 다음의 몇 가지 사항을 만족하여야 한다.

- * 적용 영역에서 다루는 여러 가지 정보를 효율적으로 저장하고 검색할 수 있도록 해야 한다. 이에 따라 데이터에 붙여진 주석 정보를 검색하고 다룰 수 있는 도구를 제공해야 한다.

- * 데이터간의 연관성을 부여할 수 있도록, 여러 데이터에 한 설명을 적용할 수 있어야 한다.

- * 하나의 데이터에 여러 가지 설명을 붙일 수 있어야 한다.

- * 하나의 비디오 데이터 내의 여러 구간에 대해서 중첩된 색인을 할 수 있어야 한다.

- * 관련 정보들간의 효율적인 연계와 통합을 이루어야 한다.

- * 주석 정보를 쉽게 추가, 삭제, 갱신할 수 있어야 한다.

- * 특징 기반 색인을 할 수 있도록 해야 한다.

5.1 검색 정보의 종류

- * 일반 정보(General Characteristics): 이는 영화나 뉴스 기사 등의 비디오 데이터 전체에 해당하는 일반적인 정보로서 제목, 생성일, 제작자, 감독, 비디오 편집자, 수상 기록, 희극, 비극 등이 이에 해당된다.

- * 내용 정보(Content): 어떤 장면이 담고 있는 객관적 내용에 대한 정보이다. 예를 들어서 스티브 맥퀸이 절벽에서 떨어지는 장면과 같은 것들이 될 수 있다. 내용을 자연어나 키워드 형태로 입력할 수도 있고, 실제로 이미지나 비디오를 주고 그와 유사하거나 같은 데이터가 있는 장면을 검색할 수도 있다.

- * 느낌 정보(Connotation): 영화의 한 장면과 같은 비디오 데이터의 한 시퀀스에 대하여 주관적으로 느낄 수 있는 느낌에 대한 정보이다. 예를 들어서 어떤 장면은 우스운 내용이다, 슬픈 내용이다,

긴장된 내용이다 등과 같은 것이 이에 해당된다. 주관적인 정보이다.

- * 구조정보: 비디오 데이터 내에 존재하는 논리 구조, 공간 구조, 시간 구조, 의미 구조에 대한 정보이다. 하나의 영화는 여러 개의 시퀀스로 이루어져 있고 시퀀스는 여러개의 shot으로 이루어져 있는데 이러한 논리적인 계층을 논리적 구조(Logical Structure)라 한다. 공간적 구조(Spatial Layout Structure)라 함은 자막이 화면 상에서 어떠한 위치에 있으며, 뉴스 프로그램에서의 사건 화면의 위치가 어떻게 되는가와 같은 내용으로 이루어진다. 시간적 구조라 함은 Shot이나 Sequence 내의 구성 요소들 즉, 비디오, 그래픽, 자막, 음향 등의 시간적 배열 구조를 말한다. 다음 그림에 구조 정보의 예를 보인다. 의미 구조는 비디오 데이터의 임의 구간이 어떤 의미를 갖는지와 또한 이들간의 관계를 나타낸다.

- * 촬영 및 제작 기법 관련 정보: 비디오 촬영 및 제작 기법과 관련된 정보이다. 조명, 카메라 각도, Zoom-in, Zoom-out, Filter, 전환 효과 등이 이에 속한다.

- * 저장 정보: 비디오 데이터 자체에 관한 정보이다. 1초당 Frame 수, 압축율, 데이터의 크기, 화면 크기 등에 대한 내용이 이에 속한다.

5.2 정보의 효율적인 갱신 및 추가

영상 데이터의 내용은 보는 사람의 시각에 따라서 또는 보는 시간에 따라서 달라질 수 있다. 따라서 내용 정보를 주석으로 달아 놓기 위해서는 이 주석 정보를 유연하게 추가하거나 갱신할 수 있도록 해야 한다. 스키마를 고정시켜 놓으면 정보를 유연하게 추가하기가 어렵다. 이러한 기능을 가능하도록 하게 하기 위하여 OVID[11]라는 시스템에서는 고정된 스키마를 사용하지 않고 개체 식별자와 속성 이름, 속성 값을 지원할 수 있는 방법을 제안하였다. 이 방법을 이용하면 속성을 자유로이 정의할 수 있으며, 그에 따른 값은 속성 값 난에

집어 넣을 수 있으므로 다양하게 정보를 색인할 수 있다. 자연어 처리나 텍스트 정보 검색의 입장에서는 자유롭게 주석을 달 수 있도록 한 뒤 여기에서 키워드를 추출하여 검색에 사용하는 방법을 사용하고 있다. 최근에는 어떤 개체와 그에 따른 속성을 하나로 묶어 생각하는 키펙트에 관한 연구도 진행되고 있다. 이러한 시도는 개체의 속성을 이용하여 검색해야 하는 영상 정보 검색을 수행하는 데에 매우 적합한 시도로 본다.

5.3 상위정보와 하위정보의 효율적인 통합

영상 처리 알고리즘을 이용하여 추출한 특징이 빨간색 의자 모양의 물체라고 한다면 이것은 특징 표현 방법으로 내부적으로 표현될 것이고 이는 사용자가 입력한 키워드 '의자'와 연관을 갖는다. 이는 다시 의자의 상위 개념인 '가구'의 한 인스턴스로 존재하도록 관계를 갖게 된다. 이러한 일련의 관계를 설정해주는 작업이 영상 정보 검색 시스템 내부에서 필요하다. 김기병과 김형주가 제안한 주석과 특징 정보의 통합 방법이 한 예이다. Kashyap과 Shah 그리고 Amit Sheth가 제안하고 있는 통합의 개념도 참고할 만하다.

5.4 영상 데이터의 색인

영상 데이터의 색인은 영상 데이터가 어떻게 표현되는가, 사용자가 어떤 검색을 하고자 하는가에 좌우된다. 영상 데이터에 대해서 종래 사용해오던 방법은 크게 주석 기반 색인과 특징 기반 색인이 있다. 색인시 기본적으로 고려해야 할 사항은 하나의 영상 내 여러 구간에 대해서 중첩된 색인을 할 수 있어야 한다는 것도 중요하다.

5.4.1 주석 기반 색인

주석 기반 색인은 어떤 데이터에 관한 정보를 개발자가 원하는대로 표현할 수 있는 장점이 있다. 또한 주석은 색깔이나 모양과 같은 외연적 특징보

다는 대상이 가지고 있는 개념이나 의미를 표현할 수 있는 장점이 있다. 이러한 개념이나 의미를 대상 데이터로부터 완전히 자동으로 추출해내는 것은 현재나 가까운 미래의 기술로는 달성하기가 어렵다. 따라서 영상 데이터에 의미나 개념 정보를 담은 것은 앞으로도 꽤 오랫동안 시스템 개발자나 영화 제작자 또는 주석 전문가가 수작업을 하는 수밖에 없을 것이다. 위에 기술한 주석 기반 색인 방법의 장점에 반해 이 기술은 다음과 같은 단점을 갖고 있다.

(1) 이 작업은 많은 시간을 필요로 한다. 따라서 적은 양의 영상 데이터에 대한 색인에는 어느 정도 적합하지만 많은 양의 데이터에 대해서는 적합하지가 않다.

(2) 주석 색인은 주관적인 관점을 담기가 쉽다. 따라서 이것이 범용성을 갖기에는 부적합하다. 이러한 이유로 해서 기존의 주석 기반 색인 기술은 일차로 색인어나 키워드 등을 선정하는 데에 치중해 있던지 아니면 사용자가 주석을 달 수 있도록 하는 사용자 인터페이스를 설계하는 데에 편중되어 있다.

주석으로 많이 사용되고 있는 키워드는 다음과 같은 단점이 있다.

(1) 영상 데이터의 시공간적인 관계나 또 다른 정보를 충분히 표현하기가 어렵다.

(2) 키워드로서는 의미정보를 충분히 표현할 수 없을뿐더러 기술하고 있는 용어간의 상속 관계, 유사성, 관계에 의한 추론 등을 쉽게 할 수 없다. 예를 들어서 '개'라는 키워드로 색인된 데이터를 검색하고자 할 때 '세퍼드'로 색인된 데이터는 검색하지 못한다는 것이다. 물론 최근에는 이러한 문제를 해결하기 위해서 개체간의 관계를 고려한 검색을 시도하고 있기도 하다.

(3) 기술하고 있는 용어들간의 관계를 나타내기 힘들다. 예를 들어서 '사람', '물다', '개'라는 키워드를 이용하여 검색을 하다 보면 '개가 사람을 문' 장면과 '사람이 개를 문' 장면이 모두 출력될 수 있다. 즉, 키워드로서는 기술하고 있는 용어간의

관계를 충분히 표현하기가 어렵다.

다음에는 최근, 영상 처리 기법을 이용하여 특징을 추출해서 이를 검색에 이용하는 특징 기반 색인에 대하여 기술한다. 이들은 주로 영상 데이터를 대상으로 하고 있다.

5.4.2 특징 기반 색인

특징 기반 색인은 주석 기반 색인과는 달리 자동으로 이루어지는 특성이 있다. 비디오 데이터를 자르고, 대표 프레임을 뽑아내고, 특징을 추출해 내는 영상처리 기술이 근간을 이룬다. 여기서 한 비디오의 특징은 색깔, 질감, 모양, 움직임 등이 될 수 있다. 이들의 장점이 색인을 완전 자동화 할 수 있음에도 불구하고 데이터가 갖고 있는 의미는 끌어내지 못한다는 단점이 있다. 따라서 이 방법은 이러한 기능을 주로 요구하는 분야에서만 사용할 것이고 일반적으로는 주석 기반 색인과 함께 사용되어야만 이런 문제점을 극복할 수 있다.

QBIC[5], VisualSEEK[7], Chabot[9], Photobook[10], CORE[12] 등이 특징 기반 검색 엔진으로서 부분적으로 키워드 및 메뉴 기반 검색 기능을 제공하고 있다.

5.4.3 영상 색인 정보의 색인

위에서 기술한 색인 정보는 시스템 내부에서 가공되고 내부 표현 형식으로 변환하여 저장한다. 주석 기반 색인 정보는 내부에서 형태소 분석기 등을 통하여 키워드나 키펙트를 추출한다. 이들은 기존의 정보 검색 시스템에서와 같이 역 색인 파일(Inverted Index File)에 저장된다.

특징 기반 색인 정보를 저장하고 빠르게 검색하기 위하여 그동안 여러가지 방법이 제안되어 왔다. 그 예를 보면 초창기 QBIC 시스템에서 사용되었던 R-Tree의 변환 방식 R*-Tree등이 있고, KD-Tree, KDB-Tree, Quad Tree 등 수많은 방법이 제안되어 왔다. 그러나 기존의 방법들은 차원이 높아짐에 따라 검색 성능이 지수적으로 떨어지는 문제점이 있다. 이에 따라 최근에는 다차원 색인 정보를 효율

적으로 색인할 수 있는 방법들이 연구되고 있다. TV-트리, SS-트리, VAM-KD 트리, VAM Split R-트리, X-트리, CIR 트리 등이 그 예이다.

5.5 영상 정보의 검색

기존의 정보 검색에는 주로 분류된 주제로부터 찾는 방법, 메뉴 기반 검색 방식, 키워드 검색 방식 등이 사용되어 왔다. 영상 정보 검색에서도 이러한 방식의 접근 방법을 고려하지 않을 수 없다. 최근에 영상의 색깔이나 질감 등을 지정하거나 모양 등을 그리든지, 아니면 찾고자 하는 장면이 있는 사진을 Query로 입력하여 질의하는 방법이 시도되고 있다. 자연어를 이용하여 사용자가 원하는 내용을 질의하는 접근 방법도 시도되고 있다. 이외에 영상 자료와 연관되어 있는 텍스트 정보를 음성으로 질의하여 관련 영상을 찾는 방법이 있을 수 있고 배경 음악의 일부분을 입력하여 해당 영상을 찾을 수도 있다.

5.5.1 영상 데이터의 질의 유형

영상 데이터에 대한 질의 방법은 내용에 의한 분류, 정확성에 의한 분류, 검색 단위에 의한 분류, 질의를 해나가는 절차에 의한 분류, 질의 내용의 정의 방법에 의한 분류, 질의 미디어에 의한 분류 등으로 나눌 수 있다. 보다 자세한 내용은 [13]를 참조하도록 한다.

5.5.2 질의의 처리

질의의 처리 과정은 파싱, 계산, 색인 검색, 결과 출력 등으로 이루어진다. 질의의 파싱 단계에서 질의의 내용은 보통 최소 단위로 분해된다. 계산 과정에서 질의로 입력된 데이터의 특징은 영상 처리 알고리즘을 이용하여 계산한다. 이후 색인을 이용하여 관련된 데이터를 찾아 주어진 값과 일치하는지 아니면 유사도가 만족한 범위에 들어오는지 판단한다. 이렇게 해서 얻어진 결과를 사용자가 보기 편리한 형태로 출력한다.

5.6 검색 결과의 출력

일반적으로 기존의 텍스트 검색에서는 제목이나 문서 번호, 문서의 내용 요약을 출력하는데 때로는 전문을 출력하기도 한다. 그러나 영상 정보 검색에서는 이러한 텍스트 자료 외에 영상 자료까지도 출력해 주어야 하는데 일반적으로 영상 자료가 크기 때문에 적은 양을 출력하면서 어떻게 전체 내용을 빨리 파악해 볼 수 있도록 하는가가 관건이다.

이에 따라 사용하는 방식을 보면 영상 정보와 같은 경우 작은 크기의 영상 대표 정보(Thumbnail)를 출력하는 방식이 많이 쓰인다. 이 중에 원하는 정보가 있으면 미리보기(Preview)를 할 수도 있다. 이러한 자료를 훑어 본 사용자는 원하는 영상 자료를 선택하여 보다 자세한 정보를 볼 수 있다. 특히 동영상인 경우는 이렇게 한 프레임에 대한 정보만을 가지고 전체 내용을 파악하기가 어려울 수 있다. 이를 위해서 예고편과 같이 전체 영상 자료를 단 몇 초나 몇 분 분량으로 요약해 놓을 수 있다. 동영상 자료를 요약해 놓는 방법은 여러 가지가 있을 수 있는데 임의의 프레임을 선택하여 몇 개만 보여주는 방법이나 장면 전환이 일어나는 첫번째 프레임들만 모아 놓았다 보여 주는 원시적인 방법

부터, 동영상 전체의 줄거리를 잘 요약해서 보여 줄 수 있는 방법(Video Abstracting)이나 장면과 장면의 흐름을 나타내도록 하는 방법(Shot Transition Graph Technique)과 같은 고급 기술을 사용할 수 있다. 현재 이러한 분야에 대한 연구가 활발히 진행되고 있다.

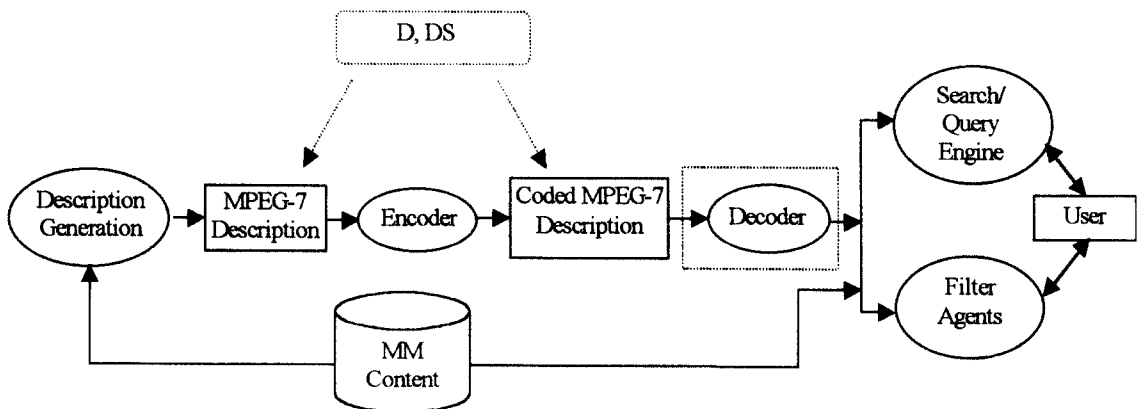
사용자가 검색하는 단위도 출력 방법을 결정하는 데에 변수가 되는데 뉴스를 예로 들면 뉴스 전체를 출력하도록 하는가 아니면 하나의 기사에 대해서만 출력하는가 하는 것도 결정되어야 한다. 이와 아울러 전송 속도 문제 등을 해결하기 위하여 저 해상도의 비디오나 이미지 데이터를 먼저 보여 주고 최종 선택 시에 원래 데이터를 내 보낼 수도 있다. 이상에서 영상 정보 검색 시스템의 일반적인 특징을 기술하였다.

다음 절에서는 영상 정보의 기술에 대한 표준안을 심의하고 있는 MPEG-7 그룹에 대한 간단한 소개와 현재까지 도출된 요구 사항을 요약한다.

6. MPEG-7의 범위[14,15]

비디오 데이터의 저장 및 검색과 관련성이 가장 깊은 표준은 MPEG-7이다. 현재 요구 사항에 대해

(그림 4) MPEG-7의 범위와 주변 기능과의 관계



일부가 정리되어 있는데 이후 2000년 초까지 표준화 작업이 계속될 예정이다. MPEG-7은 음향 및 비디오 데이터를 중심으로 한 여러 가지 멀티미디어 데이터와 관련된 정보의 표현 방법을 어떻게 표준화 하여 정보 검색 과정에 도움을 줄 수 있을 것인가 하는 관점에서부터 논의가 시작되었다. MPEG-7 그룹에서는 이것을 “멀티미디어 내용 기술 인터페이스의 표준(Multimedia Content Description Interface)”이라는 말로 정의하였다. MPEG-7을 이해하고자 할 때 특히 유의해야 할 것은 기존의 MPEG 표준화 활동이 주로 압축에 의한 데이터 양을 줄이고자 하는 데에 있었지만 이것은 압축과는 다른 표현에 대한 문제를 다루고 있다는 점이다. 특히 특징을 추출해내는 알고리즘이나 데이터의 검색을 수행하는 검색 엔진에 대한 내용은 MPEG-7의 범위 밖이라는 것이다. 다음 그림은 MPEG-7 표준화 그룹에서 보고한 회의록에 있는 것으로 앞에 기술한 MPEG-7의 범위를 잘 나타내고 있다.

멀티미디어 데이터의 내용(Multimedia Content)은 표준화된 방법을 이용하여 내용을 기술하게 되

고(Description Generation) 이는 부호화기에 의해 부호화될 수 있다. 이 정보는 수신측에 전달되거나 검색측에서 검색하고자 할 때에 복호기를 이용하여 해석한 다음 내용을 검색할 수 있다.

이 그림에서 ‘D’와 ‘DS’는 각각 기술자(Descriptor)와 기술 구조(Description Scheme)를 나타낸다. 기술자는 멀티미디어 데이터가 갖고 있는 정보(Feature)를 표현하기 위해 사용하는 단위나 값의 표현 형식이다. 기술 구조는 기술자나 기술 구조들간의 관계를 나타내는 방식을 말한다. 이와 함께 또 하나의 중요한 개념은 기술에 사용되는 언어(Description Definition Language)이다. 이는 DS를 정의할 때 사용한다. 아래 표 1에 멀티미디어 데이터에 관련된 여러 가지 유형의 정보와 그들을 나타내기 위한 기술자(Descriptor)들을 예시한다.

7. 결어

이상에서 방송용 영상데이터베이스 관련 기술에 대하여 간략히 소개하였다. 방송용 영상 데이터베이스를 구축할때 가장 큰 문제는 저장 장치의 비용 문제이다. 이 문제가 해결되지 않고는 모든 데이터

(표 1) MPEG-7에서 정의한 멀티미디어 정보와 기술자(Descriptors)

정보의 유형	정보	기술자(Descriptors)
N차원 시공간 구조	정보음악의 길이(duration of music segments)	타임코드, 기타
	개체의 궤적(trajectory)	체인코드, 기타
통계적 정보	색깔(colour)	칼라 히스토그램, 기타
	음향 주파수(audio frequency content)	주파수 성분의 평균, 기타
객관적 정보	개체의 색깔(color of an object)	칼라 히스토그램, 텍스트, 기타
	모양(shape of an object)	다각형 좌표의 집합, 모멘트의 집합, 기타
주관적 정보	질감(texture of an object)	wavelet 계수, 콘트라스트의 집합, 저침 및 방향성, 기타
	감정(행복, 분노, 슬픔 등)	eigenface 패러미터의 집합, 텍스트, 기타
	스타일(style)	텍스트, 기타
제작 관련 정보	극작가(author)	텍스트, 기타
	제작자(producer)	텍스트, 기타
	감독(director)	텍스트, 기타
합성 관련 정보	장면 구성	트리 그래프, 기타
개념 정보	사건(event)	텍스트, 기타
	활동(activity)	텍스트, 기타

를 디지털화하여 저장한다는 것은 거의 불가능하다. 따라서 저장 장치의 값이 충분히 저렴해질 때까지는 부분적인 영상 자료만 데이터베이스화해서 사용할 수 있다. 또는 전체 자료에 대해서도 해상도가 낮은 영상 자료 즉, Thumbnail과 같은 자료만이라도 보관해서 텍스트로만 검색하는 것보다는 효율이 높은 검색을 시행할 수 있도록 할 수 있다. 경우에 따라서는 미리 보기 기능도 집어 넣을 수 있을 것이다. 특히 방송 편집에 직접 사용할 수 있도록 하려면 MPEG-2 품질의 영상 데이터가 필요한데 이러한 요구 사항을 만족시키기 위해서는 저장 장치에 투자하는 비용이 매우 클 것이므로 상황에 따라 선택적으로 구축 계획을 세울 필요가 있다.

위에서 기술한 내용 외에도 비디오 데이터베이스 시스템을 구축하고자 할 때에는 네트워크 문제, 미디어 서버 문제 등도 해결해야 할 사항들이다. 이들에 대한 자세한 사항은 이 논문에서는 지면상 다루지 못했으므로 관련 문헌을 참조하도록 한다.

[참고 문헌]

1. M.G. Brown et al., "Automatic Content-Based Retrieval of Broadcast News," *Multimedia* 95, 1995
2. Rune Hjelsvold et al., "Modeling and Querying Video Data," the 20th Conference on Very Large Databases, September 1994.
3. M. Christel, T. Kanade, M. Mauldin, R. Reddy, M. Sirbu, S. Stevens and H. Wactlar, "Informedia Digital Video Library," *Communications of the ACM*, Pages 57-58, April 1995.
4. H.J. Zhang, C.Y.Low, S.W. Smoliar and J.H. Wu, "Video Parsing, Retrieval and Browsing: An Integrated and Content-Based Solution," *Multimedia* 95, 1995.
5. W. Niblack, Xiaoming Zhu, James Lee Hafner, Tom Bruel, Dulce Ponceleon, Dragutin Petkovic, Myron Flickner, Eli Upfal, Sigfredo I. Nin, Sanghoon Sull, Byron Dom, Boon-Lock Yeo, Savitha Srinivasan, Dan Zivkovic, and Mike Penner, "Updates to the QBIC System," Vol. 3312, Pages 150-161, SPIE, February 1997
6. Jeffrey R. Bach, Charles Fuller, Amarnath Gupta, Arun Hampapur, Bradley Horowitz, Rich Humphrey, Ramesh Jain, Chiao-fe Shu, "The Virage Image Search Engine: An open framework for image management," Vol. 2670, Pages 76-87, SPIE, February 1996.
7. John R. Smith and Shi-Fu Chang, "VisualSEEK: a fully automated content-based image query system," *ACM Multimedia* 96, November 20, 1996.
8. Shih-Fu Chang, John R. Smith, Mandis Beigi, and Ana Benitez, "Visual Information Retrieval from Large Distributed Online Repositories," *Communications of the ACM*, December 1997.
9. Virginia E. Ogle and Michael Stonebraker, "Chabot: Retrieval from a Relational Database of Images," *Communications of the ACM*, Vol. 28, No.9, September 1995.
10. A. Pentland, R.W. Picard, and S. Sclaroff, "Photobook: Tools for content-based manipulation of image databases," *Storage and Retrieval for Image and Video Databases II*, Pages 34-47, San Jose, CA, SPIE, February 7-8 1994
11. Eitesu Oomoto and Katsumi Tanaka, "OVID: Design and Implementation of a Video-Object Database System," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 5, No. 4, August 1993.
12. J.K. Wu, A. Desai Narashimhalu, B.M. Mehtre, C.P. Lam, Y.J. Gao, "CORE: A Content-Based Retrieval Engine for Multimedia Information Systems," *Multimedia Systems*, 1995(3)
13. Ahmed K. Elmagarmid, Haitao Jiang, Abdelslam A. Heldel, Anupam Joshi and Magdy Ahmed, *Video Database Systems - Issues, Products*

and Applications, Kluwer Academic Publishers, 1997.

14. ISO/IEC JTC1/SC29/WG11, "MPEG-7 Requirements Document V.6," ISO/IEC JTC1/SC29/WG11/ N2327, July1998/Dublin

15. ISO/IEC JTC1/SC29/WG11, "MPEG-7 Context and Objectives V.9," ISO/IEC JTC1/SC29/WG11/ N2327, July1998/Dublin



유성준

- 1982년 고려대학교 전자공학과 졸업
- 1990년 고려대학교 전자공학과 대학원졸업 정보처리 전공 석사
- 1996년 시리큐스대학교 전기전산학과 졸업 전산학 박사
- E-mail : ETRI. Re. Kr.