

멀티미디어 정보 검색 기술 동향

유 성 준

한국전자통신연구원 정보검색연구팀

I. 서 론

이 논문에서는 멀티미디어 정보 검색 기술의 최근 연구 및 개발 동향에 대하여 언급한다. 이 분야에 포함되어 있는 기술 분야는 매우 광범위하나 지면상 각각에 대한 세부 내용은 언급을 피하고 개략적인 내용만을 소개한다. 따라서 각 세부 기술에 대하여 보다 자세한 사항은 이 학회지의 관련 논문을 참조하거나 참고 문헌을 참조할 수 있다. 2절에서는 멀티미디어 정보 검색 시스템을 설계할 때 고려해야 할 기술적 고려 사항에 대하여 소개하며 3절부터 10절까지는 각 고려 사항에 대하여 보다 자세히 언급한다. 1절에서는 먼저 멀티미디어 정보 검색 시스템이 우리 주변에서 이용되는 예를 살펴보고 멀티미디어 정보 검색 시스템의 기능에 대하여 알아본다.

청소년들은 자기가 좋아하는 가수나 배우의 사진을 찾기 위해 인터넷에서 키워드를 이용하여 검색을 시행할 수 있다. 이 방식은 멀티미디어 정보 검색의 입장에서는 초보적이면서도 보편적으로 이용하는 방법이다.

게임 제작자들은 게임을 제작하는 데에 필요한 3차원 그래픽 모델, 효과 음이나 배경음악 등을 찾을 필요성을 항상 느끼고 있지만 원하는 자료를 쉽게 찾지 못하고 있다. 교육용 콘텐츠 제작자 역시 게임 제작자들과 비슷하게 그래픽 모델, 효과음, 동영상, 정지영상 등을 찾게 된다. 영화 제작자는 영화를 제작하는 데에 필요한 기존 영화 장면, 3차원 그래픽 모델, 효과음 등 외에도 어느 배우의 연기 장면, 촬영할 장소를 미리 볼 필요성을 느끼

고 있지만 이러한 자료들을 찾기가 쉽지 않다.

방송물 제작자는 기존에 방송된 자료, 촬영한 자료 화면, 녹음된 자료 등의 검색을 하게 된다. 현재는 대부분이 키워드 기반 검색을 통하여 자료 테이프 번호를 찾게 되고 이를 바탕으로 테이프를 재생해보게 된다. 그러나 이 때 테이프를 전달한다든지, 원하는 자료를 찾지 못했을 때 다시 검색 요청을 해서 원하는 자료를 찾기까지 걸리는 시간은 경우에 따라 꽤 길어서 뉴스 등을 편집할 때에 자료 화면을 정확하고 빠르게 찾기 위해서는 보다 효율적인 영상 자료 검색 시스템의 구축이 필요하다.

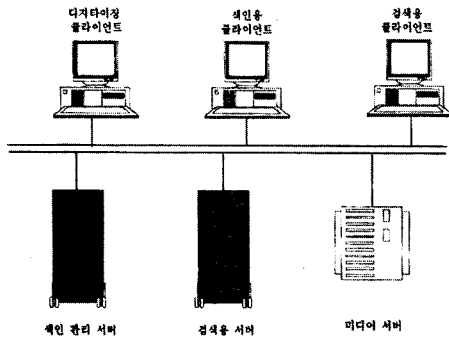
교사나 강사는 학생 교육용 자료를 만들기 위해 녹화된 멀티미디어 자료에 대한 검색이 필요하다. 이와 아울러 의사들의 교육에 수술 장면을 녹화해 놓았다가 보여 주기 위하여 검색이 필요할 수가 있다. 복잡한 기계를 고치는 전문가를 훈련시키는 데에도 이 방식을 이용할 수 있다.

병원에서는 환자와 관련된 X-ray, CT 자료, 내시경 자료 등을 검색할 수 있다. 노래방에서, 부르 고자 하는 노래가 대부분 노래 제목 순으로 나열 되어 있는데 이를 중간에 있는 노래 가사 및 음정 등을 이용하여 찾는 방법을 고려해 볼 수 있다. 방송국에서 음악 검색에서도 위에 적은 방법 이외에 노래의 분위기에 따라, 예를 들어서 슬픈 분위기, 발랄한 분위기 등으로 찾을 수도 있을 것이다. 극장에 가고자 할 때 보고자 하는 영화의 내용을 미리 예고편을 통해 찾아보고 검색할 수 있다.

이상에서 나열하였듯이 다양한 분야에서 멀티미디어 정보 검색의 필요성은 제기되고 있으며 오늘날 초보적인 기능의 시스템이 운용되고 있다.

II. 멀티미디어 정보 검색 시스템의 구성 예

위와 같은 기능을 수행하기 위한 멀티미디어 정보 검색 시스템의 한 예로 다음 그림과 같은 클라이언트/서버 구조를 생각해 볼 수 있다. 이 구조는 크게 입력과 출력 부분으로 나누어진다. 입력 부분은 데이터를 디지털화 하는 작업을 담당하는 클라이언트, 색인을 관리하는 서버, 색인을 담당하는 클라이언트 등으로 이루어지고 검색 부분은 검색을 담당하는 클라이언트, 검색을 담당하는 서버, 미디어를 관리하는 서버 등으로 나뉜다.



(그림 1) 멀티미디어 정보 검색 시스템의 구성예

III. 멀티미디어 정보 검색 시스템 설계시 고려 사항 및 관련 기술

멀티미디어 정보 검색 시스템을 설계하는 데에는 대상 미디어의 종류, 검색 정보의 종류, 웹 기반인가 사내망 기반인가, 범용 시스템인가, 특정 분야만 다루는가 등을 먼저 고려해야 한다. 세부적으로 검토할 것은 자료의 입력이나 수집, 색인, 저장, 검색, 출력 등과 같은 항목에 대한 것들이다. 검색 질의 형태를 메뉴, 텍스트, 멀티미디어 중 어떤 것으로 하겠는가, 텍스트 질의 검색이면 키워드로 할 것인가, 키워드로 할 것인가, 멀티미디어로 한다면 멀티모달 인터페이스를 허용할 것인가 등

이다. 또한 검색의 단위를 어느 정도로 할 것인가, 어떤 압축 방법을 택할 것인가. 주석 기반 검색에서는 단순히 텍스트 정보만 출력할 것인가, 텍스트와 멀티미디어 정보를 함께 출력할 것인가. 멀티미디어 정보를 출력한다면 데이터 전체를 한꺼번에 보여줄 것인가 아니면 데이터의 요약본을 먼저 보여준 뒤 필요시 데이터의 원래 값을 보여줄 것인지 등이 결정해야 할 요소들이다. Thumbnail, Preview, 요약 방법들을 어떻게 조합할 것인가, Relevance Feedback은 어떻게 할 것인가, 어떠한 색인 기술을 사용할 것인가, 데이터 모델링은 어떻게 할 것인가 등이 또한 검토해야 할 요소들이다. 이외에도 데이터 수집의 자동성, 주석 색인의 유연성, 하위정보와 상위 정보간의 연계, 정보의 자동 분류 기능, 압축 영역에서의 미디어 처리 여부 등이 모두 결정해야 할 요소들이다. 이 글에서는 이중 중요한 몇 가지에 대한 최근 기술 동향을 개략적으로 살펴 본다.

사용자가 멀티미디어 정보를 찾을 수 있는 곳은 인터넷과 같은 공개된 장소이거나 방송국이나 신문사내에서 자체적으로 운영하는 시스템과 같이 한정된 공간으로 나눌 수 있다. 검색 시스템이 다루는 정보의 범위도 특정 분야에 국한하지 않을 수도 있고 반대로 특정 분야에 국한해서 접근할 수도 있다.

현실적으로 인터넷은 일반인들이 가장 쉽게 정보를 찾을 수 있는 곳이다. 인터넷 정보 검색 시스템은 불특정 다수를 대상으로 하면서 전체적으로 보면 어떤 특정 정보만 찾을 수 있는 것이 아니고 거의 모든 분야의 정보를 대상으로 해야 한다. 이와 반면에 방송국이나 신문사, 병원 등에서는 자기 조직 내에서만 유통이 가능한 정보 시스템을 운용하고 있다. 이러한 정보 검색 시스템에서 찾을 수 있는 정보의 분야는 특화되어 있으면서 자료의 양이 방대한 것이 특징이다. 이러한 자료들은 가끔 인터넷에 공개되어 일반 사람들이 이용할 수도 있도록 하는 것도 있지만 매우 제한적이며 때로는 유료로 운용되기도 한다.

IV. 검색 정보의 종류

사용자가 검색하고자 하는 정보의 종류에는 어떠한 것들이 있는가 하는 것이 시스템을 설계하기 위해 사전에 분석해야 할 요소 중 하나이다. 검색할 정보의 종류에 대한 분류 방법은 여러 가지가 제안되어 왔다. 그러나 여기에서는 이들을 단순화하여 내용 및 서지 정보, 그리고 시간 구조나 공간 구조, 논리 구조 등과 같은 구조 정보로 나누어 생각하기로 한다.

서지 정보라 함은 기존의 도서관에서 자료를 찾기 위해 붙여 놓았던 여러 가지 정보들과 같은 것으로서 영화를 예로 들면, 언제 만들어졌는가, 누가 만들었는가, 주인공 배우는 누구인가, 길이는 얼마나 되는가, 디지털화 된 자료의 저장 용량은 얼마이고 압축 방법은 무엇을 사용했는가 등이 이에 해당된다. 이외에 비디오나 이미지가 도입되면서 촬영 거리, 촬영 높이 등 데이터의 외형적인 정보도 여기에 포함된다고 볼 수 있다.

내용 정보는 자료 속에 들어 있는 개념이나 사실, 개체 등을 대상으로 한다. 영상자료를 예로 들면 “박세리가 US 여자 Open에서 우승하고 아버지와 함께 기뻐하는 장면”과 같은 것이 될 수가 있다. 또한 “빨간색 구두”, “녹색 정원에 피어 있는 빨간 장미가 있는 사진” 등이 이러한 범주에 속한다. 이러한 방법으로 정보를 찾기 위해서는 주석 기반 검색과 특징 기반 검색과 같은 접근 방법이 사용되고 있는데 이는 7 절에서 상세하게 기술한다.

구조 정보라 함은 주로 구조화된 멀티미디어 문서에서 흔히 얘기되는 개념으로서 논리적 구조라 함은 기존 책의 장, 절과 같은 것을 말하며 영상 자료에서는 Shot, Scene, Sequence, Story, Program 등의 계위로 나타낼 수가 있다. 시간적 구조라 함은 멀티미디어 자료를 구성하고 있는 여러 구성 요소들의 시간적인 배열 관계에 주목하는 개념으로서 흔히 멀티미디어 프리젠테이션에서의 시간적인 전후 관계를 따질 때에 언급되는 개념이다. 마지막으로 공간적 구조라 함은 하나의 쪽 내

에 배열된 요소들의 배치 구조 등을 말한다. 여기서 사용되는 ‘쪽’이라는 개념은 정지 영상의 한 화면에 해당이 될 수도 있고 동영상의 한 프레임에 해당이 될 수도 있다.

V. 자료의 입력 및 수집

멀티미디어 데이터는 기존의 텍스트 데이터와 비교하여 그 특성이 다르기 때문에 입력에는 스캐너나 카메라, 인코딩 장비 등을 사용할 필요가 있다. 특히 비디오 데이터를 입력하는 데에는 많은 시간을 필요로 하며 의미있는 구간별로 절단하는 데에도 노력이 많이 필요하다. 여기에서 비디오 데이터를 장면 단위로 찾아내는 작업이 필요한데 이것을 보다 쉽게 하기 위하여 장면 변화 지점을 자동으로 찾아주는 기술이 있다. 이러한 기술은 비디오 데이터의 압축 상태를 그대로 유지하면서 수행하게 되면 보다 빨리 찾을 수 있고 압축한 것을 풀 뒤에 처리하면 속도가 늦어지게 된다. 최근에는 압축 영역에서 이러한 작업을 수행하는 연구가 행해지고 있고 관련 시스템들이 등장하고 있다.^[1] 특히 일본의 NTT 같은 방송국에서는 장면 전환이 일어나는 구간을 자동으로 찾아주는 시스템을 실제로 이용하고 있어서 자료 입력 과정의 생산성을 높이고 있다.

자료의 입력은 위와 같이 수동이나 반 자동으로 수행할 수 있다. 이와 아울러 한 곳에 멀티미디어 자료를 모아 놓고 서비스 하는 시스템과는 달리 색인만을 저장해놓고 해당 멀티미디어 데이터를 찾아갈 수 있도록 하는 웹 기반 멀티미디어 검색 엔진도 최근 등장하고 있다^[2]. 이 시스템은 텍스트를 찾은 뒤 관련 멀티미디어 자료가 있으면 부가적으로 볼 수 있도록 한 것이 아니고 파일 타입 등을 이용하여 멀티미디어 자료를 직접 찾아내는 방식을 택하고 있다.

자료의 입력 시에 또 하나 문제가 되는 것이 자동 분류(Categorization) 기술이다. 즉, 멀티미디어 데이터가 들어오게 되면 이의 특징이나 주석

정보를 분석하여 그들이 어느 부류에 속하는지를 자동으로 가려내어 분류 기반 검색에 이용하도록 할 수 있다.

VI. 데이터 모델링

사용자의 요구 사항에 상응하는 검색 서비스를 제공하기 위해서는 데이터를 저장하는 효율적인 모델을 정해야 한다. 데이터 모델은 몇 가지 조건을 만족해야 한다. 첫째로, 적용 영역에서 다루는 여러 가지 정보를 효율적으로 저장하고 검색할 수 있도록 해야 한다. 또 하나는 유연성을 제공해야 한다는 것이다. 일반적으로 하나의 멀티미디어 데이터에 대해서 사람들마다 또는 시간에 따라 부여하는 의미가 여러 가지로 다를 수 있는데 이를 지원해 주어야 한다는 것이다. 세 번째로는 관련 정보들간의 효율적인 연계와 통합을 이루어야 한다. 이 밖에 기존의 멀티미디어 데이터를 편집, 조작하여 새로운 멀티미디어 데이터를 만들고자 할 때 또는 사용자가 보고자 하는 관점으로 출력하고자 할 때 이에 관련된 연산을 정의해야 하고 또한 해당 색인을 조정하는 기능을 지원해야 한다. 여기에서는 두 번째, 세 번째 항목에 대하여 기술한다.

1. 정보의 효율적인 갱신 및 추가

멀티미디어 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. 예를 들어 비디오 데이터는 사람에 따라 또는 시간에 따라 여러가지 의미를 부여할 수 있다. 이러한 요구 사항을 고려하여 유연하게 정보를 갱신하고 추가할 수 있도록 하는 기능을 지원하고자 한다면 이에 걸맞는 방법을 사용하여야 한다.

일반적으로 스키마를 고정시켜 놓으면 정보를 유연하게 추가하기가 어렵다. 이러한 기능을 가능하도록 하게 하기 위하여 OVID[2]라는 시스템에

서는 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.

이렇게 데이터베이스 입장에서 스키마를 정의하도록 하는 반면에 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 킵트에 관한 연구도 진행되고 있다. 이러한 시도는 개체의 속성을 이용하여 검색해야 하는 멀티미디어 정보 검색을 수행하는 데에 적합한 시도로 본다.

2. 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 멀티미디어 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.

먼저 김기병과 김형주가 제안한 방법을 보면 비디오 관련 정보를 네개로 나누고 있다. 이들은 최하위에 특징정보, 최상위에 개념정보가 있고 그 사이에 특징 시퀀스 정보, 부개념 정보를 체계적으로 연관시켜서 검색하기 위하여 다섯 개의 계층을 도입하였고 이들 사이의 연산을 정의하였다.

Kashyap 등이 기술한 Ontology와 메타 데이터, 그리고 원시 데이터와의 관계를 설정하고자 한 개념은 위에서 기술한 계층 구조와 같은데 여기에서 강조하고 있는 것은 미디어 타입별로 Ontology나 메타 데이터가 다른 것이 있기 때문에 이들을 분리해서 고려해야 한다는 것이다. 미디어 형태에 관

련된 Ontology가 있을 수 있고 아니면 그것과 관계가 없는 Ontology가 있다. 이들은 메타 데이터들과 연관되어 있고 이 메타 데이터들은 다시 이미지, 음향, 비디오 등의 특징 데이터와 연관되어 있다.

VII. 멀티미디어 데이터의 색인

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

1. 주석 기반 색인

주석 기반 색인은 어떤 데이터에 관한 정보를 개발자가 원하는대로 표현할 수 있는 장점이 있다. 또한 주석은 색깔이나 모양과 같은 외연적 특징보다는 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 멀티미디어 데이터에 의미나 개념 정보를 담는 것은 앞으로도 꽤 오랫동안 시스템 개발자나 영화 제작자 또는 주석 전문가가 수작업을 하는 수밖에 없을 것이다. 최근에 '내용 기반 멀티미디어 정보 검색'이라고 번역이 돼서 언급하고 있는 "Content Based Multimedia Retrieval" 기술은 사실은 다음에 기술할 특징기반검색으로서 어떤 데이터 내에 있는 내용에서 개념이나 사건을 색인하거나 검색하는 데에는 불충분한 기술이며, 텍스트 정보 검색 기술을 포함하는 기존의 정보 검색 기술에의 의존을 배제할 수가 없다. 물론 이것은 응용 분야에 따라 다르기는 하다. 즉, 아래 기술하는 영상 처리 기법 등을 이용한 특징 기반

색인 기술을 주로 요구하는 분야도 있을 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 멀티미디어 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) 멀티미디어 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.

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 사용하여야만 이런 문제점을 극복할 수 있다.

QBIC[3,4], VisualSEEK[5], Chabot[6], Photobook[7], CORE[8] 등이 특징 기반 검색 엔진으로서 부분적으로 키워드 및 메뉴 기반 검색 기능을 제공하고 있다.

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

4. 멀티미디어 색인 정보의 표현

MPEG-7 그룹에서는 멀티미디어 색인 정보 표현 방법의 표준을 어떻게 정할 것인지에 대하여 논의 중에 있다. 색인 정보 표현 방법을 표준화한다면 데이터를 주고 받을 때 이들 간의 색인 정보도 함께 받을 수 있고 이를 이용해 서로 떨어져 있

는 사용자나 다른 시스템을 운용하고 있는 상황에서도 표준화된 데이터 표현 방식을 해석하여 데이터의 내용을 알 수 있을 것이다. 보다 자세한 내용은 이 학회지의 관련 논문을 참조할 수 있다.

VIII. 멀티미디어 자료의 저장

주지하는 바와 같이 멀티미디어 자료는 크기가 기존 텍스트 자료보다 월등히 큰 특징을 갖고 있다. 그 중에서도 비디오 데이터는 저장 장치를 구입하는 데에 경제적으로 큰 부담을 가져오기 때문에 이에 대한 검토가 필요하다. 실제로 방송국에서 매일 쏟아지는 비디오 데이터 1년 분을 MPEG-2 (6Mbps)로 하드 디스크에 저장한다고 하면 소요되는 비용은 수 천억원에 달한다. 따라서 디지털 리니어 테이프(DLT)와 같은 3차 저장 장치를 이용하는 것을 고려해야 한다.

시중에 나와 있는 시스템으로서 DLT 라이브러리 가격은 소프트웨어 비용이 만만치 않기 때문에 아직은 저렴하다고 볼 수 없다. 따라서 적은 비용으로 여러 개의 DLT를 묶어서 3차 저장 장치로 이용하고자 한다면 구동 소프트웨어 기술을 개발해야 한다. MPEG-2로 저장하는 것이 부담이 되기 때문에 그 중 일부만 MPEG-2로 저장하고 나머지는 저해상도의 작은 크기 데이터로 저장하는 방법이 있다. 즉, Thumbnail 수준이나 단순히 Preview만 해볼 수 있을 정도의 데이터를 저장할 수도 있다. Preview하는 데에도 자료 전체를 모두 보거나 듣지 않고 비디오 데이터 같은 경우는 대표되는 몇 개의 프레임만 저장해서 보여 줄 수도 있다.

IX. 멀티미디어 정보의 검색

기존의 정보 검색에는 주로 분류된 주제로부터 찾는 방법, 메뉴 기반 검색 방식, 키워드 검색 방

식 등이 사용되어 왔다. 멀티미디어 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. 주석으로부터 키워드를 찾아서 관련된 멀티미디어 자료를 찾는 방법도 있다. 이와 아울러 웹 문서에 포함되어 있는 화일의 타입(jpg, bmp, wav, avi)으로 멀티미디어 자료를 찾을 수도 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. 또한 찾고자 하는 노래나 음악의 일부분만을 알고 있을 때 그 음악을 허밍으로 입력하여 전체 곡을 찾는 방법도 개발 시도 중인 기술 중의 하나이다.

1. 멀티미디어 데이터의 질의 유형

멀티미디어 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 [9]를 참조하도록 한다.

2. 질의의 지정과 처리

(1) 멀티미디어 데이터베이스 검색을 위한 질의 언어

SQL과 같은 텍스트 형태의 질의어는 멀티미디어 정보 검색하기 위한 질의어를 지정하는 데에 한계가 있다. 이는 텍스트 형태의 질의어로는 시공간적인 정보나 시각적인 정보를 충분히 표현하기가 쉽지 않기 때문이다. 시각적인 정보의 정의는 주로 Query by example 형태로 할 수 있다. 시공간적인 정보에 관한 질의를 위해서는 몇가지 언어가 제안되어 있는데 여기에는 TSQL, STL, VQL 등이 있다.

이외에 GUI도 사용자의 질의를 보다 더 명확하

게 표현하도록 하는 데에 사용할 수 있는 도구이다. OVID 시스템에서의 SELECT-FROM-WHERE 구문을 집어 넣도록 해서 검색하는 데이터를 선택할 수 있도록 한다든지 하는 것은 이러한 한 예이다. 특히 데이터베이스 내의 내용을 위의 구문을 집어 넣는 곳에서 메뉴로 선택할 수 있도록 한 것이 특징이다. 이외에 Video Query System이라는 것이 있는데 이것은 다음과 같은 연산 기능을 갖고 있다.

- 두 개의 비디오 구간이 겹치는지 결정.
- 두 개의 비디오에 대한 합집합, 교집합, 여집합 등을 계산.
- 집합에 대해서 반복 작업을 수행.

이 외에 질의어에 대한 보다 자세한 사항은 [9]를 참조하도록 한다.

(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.

X. 검색 결과의 출력

일반적으로 기존의 텍스트 검색에서는 제목이나 문서 번호, 문서의 내용 요약을 출력하는데 때로는 전문을 출력하기도 한다. 그러나 멀티미디어 정보 검색에서는 이러한 텍스트 자료 외에 멀티미디어 자료까지도 출력해 주어야 하는데 일반적으로 멀티미디어 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 방법부터, 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 멀티미디어 정보 검색 시스템의 일반적인 특징을 기술하였다. 다음 장에서는 멀티미디어 정보 검색 시스템을 설계하기 위하여 고려해야 할 사항 및 관련 기술에 대하여 설명한다.

XI. 국내의 시스템 개발 동향

이미 연구 개발했거나 상용으로 쓰이는 멀티미디어 정보 검색 시스템은 크게 인터넷을 기반으로 동작하는 것이거나, 아니면 어느 조직 내에서 움직이는 두 가지로 나누어 생각할 수 있다. 이와 아울러 범용의 멀티미디어 정보 검색 엔진이나 아니면 실험 시스템, 또는 콘텐츠가 들어 있는 실용 시스

템 등으로 나눌 수 있다. 여기에서는 몇 가지 대표적인 시스템에 대해서 그 특징을 살펴보기로 한다.

(1) 인터넷 문서 검색 및 디지털 라이브러리

광의로 보면 인터넷 문서도 멀티미디어 문서이다. 이와 아울러 멀티미디어를 포함한 미래의 디지털 라이브러리도 멀티미디어 정보 검색 시스템으로 볼 수 있다. 특히 미국의 디지털 라이브러리 프로젝트 중에서 CMU의 Infomedia 시스템은 멀티미디어 정보 검색 중에서 멀티 모달 인터페이스를 제공하고 있는 시스템이다. 이 시스템에서는 이미지 처리 뿐만 아니라, 음성 인식, 자연어 처리 등을 이용하여 데이터의 저장 및 검색에 이용하고자 하는 의욕을 보였다. 특히 News-on-Demand 시스템은 자연어를 집어 넣었을 때 관련되는 뉴스 화면을 검색할 수 있도록 하는 기능을 구현하여 시험하기도 하였다.

(2) 인터넷을 이용한 카펫 구매 시스템

www.floorspecs.com에 있는 카펫 구매 시스템을 보면 메뉴 방식으로 사용자가 원하는 사양- 제조사, 카펫 타입, 무게 등 - 을 집어 넣은 뒤 그에 해당되는 카펫 리스트와 모양을 출력해주는 시스템이다. 여기에서는 실제로 구매가 이루어지고 있고 향후 질의 이미지로부터 특징 기반 검색을 거쳐 해당 카펫을 찾을 수 있도록 기능을 추가할 예정이다.

(3) 인터넷 상의 3차원 그래픽 모델 라이브러리

Viewpoint, Zygote, Avalon이 대표적인 라이브러리이다. 한 때 Avalon은 가장 규모가 크고 역사가 오래된 공공 3D 아카이브를 제공했었지만 현재는 Viewpoint사에 흡수 합병되었다. Viewpoint사는 라이브러리 판매와 주문 제작 서비스를 해주고 있다. 모델 라이브러리의 검색은 분류 메뉴를 찾아 들어가면 찾을 수 있도록 하였다. 검색된 모델들은 사용자에게 보여지고 이들 중에 원하는 것을 선택하여 더 자세한 내용을 알 수 있도록 하였다.

(4) QBIC

IBM에서 개발한 정지영상 및 동영상 검색 엔진으로써 특징 기반 검색 기능이 추가 되고, 주석 기반 검색 기능이 추가된 시스템이다. 해가 거듭 됨에 따라 그 기능이 향상되고 있다. 여기에 최근에

는 장면 전환 검색 기능과 스토리보딩에 의한 비디오 요약 출력 기능도 첨가하였다.

(5) Virage

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.

(6) Columbia 대학의 VisualSEEk과 Web-SEEk^[5,11]

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.

(7) 국내의 시스템 개발 동향

국내에서는 이미 문화 방송이나 KBS에서 조선일보사 등과 함께 자체 조직 내의 멀티미디어 정보 검색 시스템을 구축하여 부분적으로 사용하고 있다. 이들은 주로 주석 기반 검색 시스템으로 알려져 있다. 이외에 현재 멀티미디어 콘텐츠진흥센터에서 국내의 멀티미디어 콘텐츠 제작업자를 위한 자료를 저장하고 검색 기능을 지원하기 위한

시스템을 구축하고 있다. 한국전자통신연구원의 정보검색연구팀에서는 비디오 및 이미지 검색 시스템 구축 솔루션을 개발하고 있다. 대한 뉴스 및 수천장의 이미지를 저장하여 검색하고자 하는 시스템을 개발하면서 내부 구성 요소인 키워드 기반 텍스트 정보 검색 기능, 형태소 분석 기반 색인 기능, 사전 구축, 특징 기반 상표 검색 기능, MPEG의 장면전환 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 계획이다.

(8) 기타

VideoSTAR는 노르웨이 공대에서 90년대 중반에 발표한 DBMS 기반 시스템이다. 데이터베이스에 저장되어 있는 자료를 메뉴로 제공하여 검색을 할 수 있도록 한 비디오 데이터베이스의 초창기 시스템이다. 이외에 버클리 공대에서 개발한 POSTGRES를 기반으로 해서 개발한 Chabot이라는 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 Photobook이라는 것이 있다. 모양, 외형, 질감 특징을 기반으로 검색할 수 있다.

XII. 결 어

이상에서 멀티미디어 정보 검색 시스템의 이용 예를 통한 기능의 설명과 최근 기술 연구 및 개발 동향을 살펴 보았다. 멀티미디어 정보 검색 시스템은 멀티미디어 기술의 발달 및 전송 속도의 발달

에 따라 그 이용 범위가 점점 늘어날 것이고 이에 따라 멀티미디어 정보를 보관하고 검색하고자 하는 필요성도 증가하고 이 기술이 보다 널리 유용하게 사용되리라 본다. 그러나 실제로 검색 시스템을 구축하고자 할 때에는 위에 기술한 내용과 같이 기술적으로 결정해야 할 것이 많고, 비용면에서도 저장 장치의 구입을 고려해야 하기 때문에 쉽게 접근할 수 있는 분야는 아니다.

참 고 문 헌

- [1] Chee Sun Won, Dong Kwon Park and Seong-Joon Yoo, "Extracting Image Features from MPEG-2 Compressed Stream," Storage and Retrieval for Image and Video Databases VI, vol. 3312, February 1998.
- [2] 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.
- [3] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, and P. Yanker, "The QBIC Project: Querying Images by Content Using Color, Texture, and Shape," Storage and Retrieval for Image and Video Databases, Vol. 1908, Pages 173-187, SPIE, February 1993.
- [4] 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.
- [5] John R. Smith and Shi-Fu Chang, "VisualSEEK: a fully automated content-based image query system," ACM Multimedia 96, November 20, 1996.
- [6] Virginia E. Ogle and Michael Stonebraker, "Chabot: Retrieval from a Relational Database of Images," Communications of the ACM, Vol. 28, No.9, September 1995.
- [7] 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.
- [8] 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).
- [9] Ahmed K. Elmagarmid, Haitao Jiang, Abdelslam A. Heldel, Anupam Joshi and Magdy Ahmed, Video Database Systems - Issues, Products and Applications, Kluwer Academic Publishers, 1997.
- [10] 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.
- [11] 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.

저자 소개

**兪 成 濬**

1959년 10월 18일생, 1982년 고려대학교 전자공학과 졸업, 1990년 고려대학교 전자공학과 대학원 졸업 정보처리 전공 석사, 1996년 시라큐스대학교 전기 및 전산학과 졸업 전산학 박사, 1982년 3월~

현재 한국전자통신연구원 정보검색연구팀장, 책임연구원,
 〈주관심 분야: 멀티미디어 정보 검색, 구조화 멀티미디어 문서 정보 검색, 멀티미디어 데이터베이스, MPEG-7 지식관리시스템, 전자정부, 디지털라이브러리, 전자상거래〉
