

내용 기반 영상 정보 검색 기술의 현황

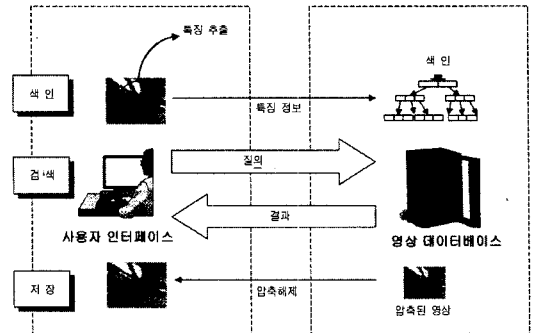
노형기, 황본우, 문종섭, 이성환
고려대학교

I. 서론

최근 디지털 영상과 비디오 데이터의 양이 늘어 감에 따라 이를 색인하고 검색할 수 있는 방법이 필수적으로 요구되고 있다. 특히 디지털 방송 관리, 주문형 비디오, 디지털 라이브러리, 의료 진단 시스템, 홈쇼핑 등 멀티미디어 데이터를 다루는 대부분의 서비스에서 영상 정보 검색에 대한 수요가 나날이 증가하고 있다^[1,2].

영상 정보 검색은 영상을 분석하여 특징을 추출한 다음, 이를 색인하고 유사한 특징을 가지는 영상을 검색하는 기술이다. 기본적으로 영상간의 유사도 판단을 위해서는 영상이 가진 내용을 파악하여 그 정보들을 서로 비교 가능한 데이터로 변환해야만 한다. 따라서 영상 정보의 검색 기술은 궁극적으로 영상 분석(image analysis), 영상 이해(image understanding), 또는 객체 인식(object recognition) 기술과 같은 기반 기술과 함께 발전해 나가야 한다. 다만 현재로서는 이러한 기술들의 수준이 영상 검색에 사용할 만큼 성능을 보이지 못하고 있으므로 영상을 대표할 수 있는 특징 정보를 영상에서 추출하여 그 정보를 이용하는 방법이 주로 사용되고 있다.

영상 정보 검색 시스템은 크게 영상에서 특징 정보를 추출하는 부분, 추출된 특징 정보를 색인하는 부분, 그리고 사용자에게 질의를 수행할 수 있도록 하는 인터페이스 부분과 데이터베이스와의 연동을 위한 기술로 나눌 수 있다. 영상 정보 검색은 텍스트를 이용한 질의 이외에도 영상의 특징 정보를 이용한 질의를 위한 다양한 인터페이스를



(그림 1) 영상 검색 시스템 개요

갖추어야 실용적인 시스템이 될 수 있다. 그림 1은 일반적인 영상 검색 시스템을 보여 주고 있다.

최근 web이 각광을 받기 시작하면서 인터넷상의 영상 검색 엔진의 개발에 대한 시도도 이루어지고 있는데 일반 사용자들이 쉽게 접근하여 이용할 수 있다는 장점으로 인해 앞으로 이 분야에 있어서 많은 발전이 있을 것으로 기대된다^[3]. 또한 동영상 데이터의 경우 컷 검출 기법과 대표 프레임 추출 기술을 사용하여 장면 단위로 분할이 가능한데, 각 대표 프레임 별로 이러한 영상 정보 검색 기술을 적용하면 비디오 정보 검색 또한 현실화될 수 있을 것이다^[4].

II. 본론

영상 검색 기술은 영상에서 특징 정보를 추출하기 위한 영상 분석 기술, 추출된 정보를 색인하는 영상 색인 기술, 영상의 저장과 질의 수행을 위한

데이터베이스와의 연동 기술, 사용자의 편리한 질의를 위한 인터페이스 기술로 나누어 볼 수 있다. 이 중에서 영상 분석 기술이 핵심적인 기술로 정확한 영상 분석을 통해 영상 고유의 특징 정보를 추출할 수 있어야 한다.

영상 분석 기술은 검색 방법에 따라서 텍스트 기반 검색과 내용 기반 검색으로 분류할 수 있다. 텍스트 기반 검색은 영상을 여러 개의 키워드나 주석으로 표현하여 일반적인 텍스트 데이터 검색 방법을 적용시킨 것이고 내용 기반 검색은 영상 분석을 통해 추출된 정보를 이용하여 영상간의 유사도를 측정하는 방법을 사용한다.

영상 분석 기술을 추출되는 특징 정보의 종류에 의해서 색상과 질감, 형태 정보 및 영상을 구성하고 있는 객체들의 공간적인 배치 정보를 이용한 검색으로 나눌 수 있다. 색상 정보는 영상의 전체적인 색상 분포를, 질감은 영상에 포함된 객체나 배경의 표면의 특징 정보를, 형태 정보는 영상에 포함된 객체의 대략적인 형태를 알려 준다. 공간적인 배치 정보는 여러개의 객체가 포함된 경우 이들의 관계를 알려 주어 검색에 이용하게 된다.

최근 들어 영상 데이터의 질이나 크기가 급격하게 증가하여 JPEG, GIF 등의 다양한 파일 압축 방식이 나타나게 되었다. 이에 따라서 영상 분석 기술을 대상 데이터에 의해 구분하면 압축된 영상에 대한 분석과 비압축 영상에 대한 분석으로 분류가 가능하다. 압축된 영상에서는 압축할 때 사용하는 영상처리 기법의 특성을 이용하여 검색에 필요한 특징 정보를 추출한다.

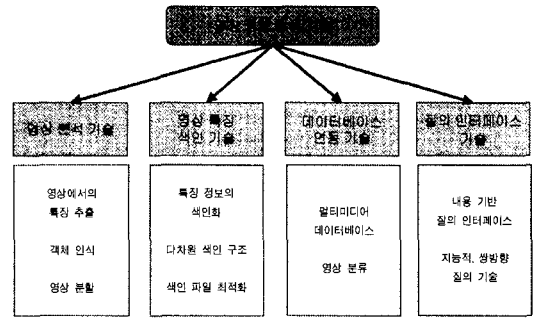
영상 색인 기술은 추출된 특징 정보간의 유사도 측정을 통해서 영상을 정렬하고 색인하는 방법에 대한 것이다. 기존의 텍스트 기반 정보 검색에서는 역파일 기법이나 B+트리 구조 등을 사용하여 이를 구현했으나 영상의 내용 정보의 경우에는 특징 정보의 종류도 다양하고, 보통 다차원의 특징 정보를 갖기 때문에 이에 적합한 여러 가지 방법들이 제안되고 있다.

데이터베이스와의 연동 기술은 구조적이고 효율적인 영상 정보 검색을 위해 필수적인 영상 데이터에 적합한 개념, 데이터 모델, 질의 언어를 포함

한 총체적인 기술로 다른 멀티미디어 데이터인 음성이나 동영상을 포함한 멀티미디어 데이터베이스로 발전해 나가야 할 것이다.

질의 인터페이스 기술은 사용자가 내용 정보에 의한 검색을 수행하기 위해 질의를 생성하는 부분으로 영상이 가진 다양한 내용 정보를 충분히 이용할 수 있어야 한다. 또한 web 상에서 동작하는 web 영상 검색 엔진을 위해서는 데이터 수집을 위한 로봇 에이전트 기술 및 CGI나 Java와 같은 web 인터페이스 기술 역시 필요하다.

본 고에서는 앞서 언급한 여러 가지 영상 정보 검색의 기법과 현재의 영상 검색 기술의 현황 및 앞으로 연구되어야 할 방향을 알아본다.



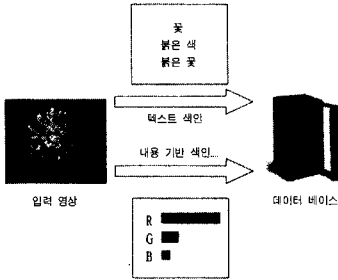
〈그림 2〉 영상 정보 검색의 요소 기술

1. 영상 분석 기술

(1) 텍스트 기반 영상 검색

텍스트 기반 영상 검색 방법은 영상을 대표하는 키워드나 주석과 같이 부가적인 정보를 영상과 함께 저장하고, 사용자가 입력한 질의어와 유사도 비교를 통해서 검색을 하는 방법으로 기존의 텍스트 문서 검색 방법과 동일한 방법을 사용한다. 영상에서 자동으로 키워드를 생성하는 것은 영상에서의 객체 추출과 객체 인식을 비롯하여 지식 기반 인공 지능 시스템 등이 필요한데 이러한 기술에는 아직 이르지 못했기 때문에 텍스트 키워드와 주석은 사람에 의해서 수행되고 있다. 이 방법은 영상이 특정한 주제를 가진 경우 또는 하나의 객체를 포함한 경우(예를 들면 자동차 사진) 우수한 성능을 보일 수 있지만 복잡한 영상이나 영상이 지닌

주제가 명확하지 않은 경우에 대해서는 텍스트 색인을 하는 사람의 주관에 개입될 수밖에 없다. 또한 이러한 색인 작업은 시간과 인력을 많이 필요로 한다. 이러한 단점을 대체하기 위해 80년대 말부터 영상을 텍스트 주석과 같은 부가정보가 아닌 그 영상 자체의 내용으로 검색하려는 방법들이 시도되기 시작하였다.



〈그림 3〉 텍스트 색인 방법과 내용 기반 색인 방법의 비교

(2) 내용 기반 영상 검색

1) 색상 정보를 이용한 영상 검색

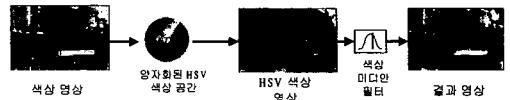
색상은 영상을 표현하는 주된 특징으로 내용 기반 영상 검색의 초기부터 사용되어 온 방법이다. 색상 정보는 영상내의 조명 변화나 관측 위치, 크기 변화 등에 어느 정도 무관하게 적용 가능하다는 장점을 가지고 있으며 방법이 간단하여 내용 기반 영상 검색에서 가장 많이 사용되는 방법 중의 하나이다.

색상 특징 정보를 추출하는 대표적인 방법이 색상 히스토그램을 이용하는 것이다. 색상 히스토그램을 특징 정보로 하면 색상의 수 만큼에 해당하는 크기의 공간이 필요하므로 보통 유사한 색상별로 군집화하는 색상 집합을 사용하게 된다. 초기에는 히스토그램을 균등한 비율로 나누어 히스토그램 구간(bin)을 결정하는 방식으로 색상을 군집화했지만, 효율적인 검색 시스템을 위해서 유사한 색상끼리 군집화하는 방법들이 제안되었다^[5].

색상 정보를 사용한 검색 방법의 경우에는 색상 공간상에서의 거리 측정을 정확하게 하기 위하여 일반적인 색상 공간인 RGB 모델을 사용하는 대

신, HSV, YIQ, YUV 등과 같은 모델을 사용하고 있다. 이중 HSV 색상 모델은 인간의 시각 능력과 유사한 특성을 갖는 것으로 알려져 있는데, 각각 색상(hue), 명도(value), 채도(saturation)의 세 가지 컴포넌트로 구성되어 유사한 색상 별로 쉽게 군집화가 가능하다는 장점이 있다.

색상 히스토그램이나 색상 집합을 이용하면 영상 내 객체들의 배치 즉 색상 분포의 공간적인 정보를 전혀 이용할 수 없다는 단점이 있다. 따라서 전체 영역을 몇 개의 부분 영역으로 분할하여 부분 영역에 대한 색상 히스토그램을 사용하여 공간 정보의 이용이 가능하다. 부분 영역으로 분할하는 방법들에는 균일한 크기로 분할하거나 영상의 중앙에 주요 객체가 존재하는 경우가 많기 때문에 중심 영역과 바깥 영역으로 분할하는 방법, 그리고 질감이나 외곽선과 같은 다른 특징 정보를 이용하여 부분 영역을 분할하는 방법들이 적용 가능하다^[6].



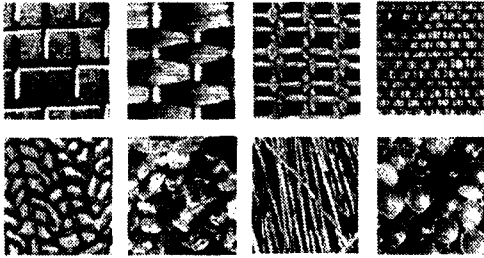
〈그림 4〉 HSV 색상 집합을 이용하여 특징 추출

2) 질감 정보를 이용한 영상 검색

질감은 영상 내부의 동일한 패턴을 가지는 영역에 대한 특징으로 하나의 색상 정보로만은 표현할 수 없다. 질감은 영상내의 객체의 표면이 가지는 특성을 표현하고 있기 때문에 질감을 가지는 영상에게는 매우 중요한 특징이 된다. 특히 항공 사진을 분석하는 경우에 질감 정보가 많이 사용되어 왔는데, 질감을 통해서 사진의 각 영역이 물의 표면, 경작지, 산 또는 도시 중 어디에 속하는지를 분석할 수 있다.

질감의 특성을 표현하는 요소에는 질감의 미세 정도를 나타내는 거칠기(coarseness), 명도의 높낮이 분포를 나타내는 대비(contrast), 질감이 가지고 있는 특정한 방향을 나타내는 방향성(directionality) 등이 있다.

그림 5에서 보는 것과 같이 자연 영상에 나타나는 질감들의 특성은 매우 다양하고, 영상내의 한



(그림 5) 질감을 포함한 영상의 예

점이 아닌 어떤 일정한 범위의 영역에서 질감 특징을 추출해야 하는 점이 질감 정보의 추출을 어렵게 만들고 있지만 질감 분석을 위해 많은 연구가 진행되어 오고 있다. 초기에는 통계적인 분석이나 구조적인 분석 또는 이들의 조합을 통해서 추출하거나 Markov 모델이나 프랙탈 모델과 같은 모델 기반 분석 방법을 통해 추출하는 질감 정보를 추출 하였는데 특히 프랙탈 모델을 이용한 질감 분석 방법은 구름이나 나무와 같은 자연 현상의 질감 정보를 분석하는데 적합하다^[7]. 최근에는 Gabor 필터, wavelet 변환 등을 통해서 질감 정보를 주파수 분석 방법이 많이 연구되고 있는데 이러한 방법은 인공적인 질감 즉 건물의 벽돌 무늬나 의류에서 나타나는 무늬 등의 분석에 큰 효과를 보이고 있다^[8].

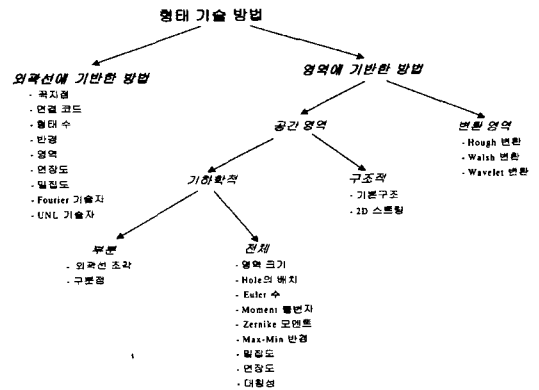
3) 형태 정보를 이용한 영상 검색

영상의 내용을 파악하기 위해서는 무엇보다도 영상내의 객체의 형태에 관심을 가질 필요가 있다. 자연 영상에서 객체의 형태를 배경으로부터 완벽하게 추출해 내는 일은 현재의 기술로는 불가능하지만 형태 분석에 대한 기술은 의료 사진에서 여러 가지 기관이나 세포조직의 분석 또는 기계 부품 분류 등의 응용 분야에서 사용되어 왔고, 더불어 많은 연구가 이루어지고 있다^[9].

형태 분석 및 표현 방법은 크게 형태의 외곽선 정보를 이용한 것과 영역 내부를 이용한 것으로 분류가 가능하다. 외곽선 정보를 이용하는 방법은 영상에서 외곽선을 추출하여 이를 Fourier 기술자와 같이 외곽선의 분석을 통해 그 변화를 특징 정보로 삼는 것이다. 영역 내부 정보를 이용하는

것은 영상 전체의 화소의 분포를 분석하는 것으로 영상에 대한 모멘트를 구하는 방법과 같이 형태의 분포 정도를 특징 정보로 삼는 것이다. 또한 분석 결과에 따라서 연산 결과가 일련의 수치 데이터로 나오는 경우와 영상으로 결과가 나오는 경우로 분류해 볼 수 있다. 전자를 스칼라 변환 방법이라고 하고, 후자를 공간상의 방법이라고 한다. 스칼라 변환에는 Fourier 기술자 등이 있고, 공간상의 방법에는 중앙 축 변환(medial axis transform) 등이 있다.

형태 특징 정보를 사용하는데 있어서 영상간의 효율적인 유사도 측정을 위해서는 형태를 표현하는 특징 수치가 서로 비교하는데 적합해야 한다. 형태는 어떤 특정한 도형 또는 객체를 표현하는 것이므로 원래 도형이 특정한 기하 변환(예를 들면 회전 변환, 크기 변환, 위치 변환)에 의해 변형되어도 표현 수치는 변경되지 않아야 한다는 특징을 가져야 한다.



(그림 6) 형태 특징을 표현하는 여러 가지 방법^[10]

4) 위치 정보를 사용한 검색 방법

영상내 객체들의 위치 정보를 사용한 검색 방법은 색상이나 형태, 질감 정보를 사용한 검색 방법에 비하여 색인에 필요한 처리 과정이 복잡하기 때문에 현재까지는 많은 연구가 이루어지고 있지 못하다.

영상의 위치 정보는 객체간의 상관관계를 표현하는 것으로 객체들간의 공간적, 위상적 관계에 의해 객체의 인접, 겹침, 포함 등의 정보를 기술하는

것이다. 이러한 방법을 위해서는 정확한 객체 분할과 인식이 가능해야 하고 가변적인 크기의 색인을 가져야 하는 여러 가지 문제점을 가지고 있다.

또 다른 방법으로는 객체와 무관하게 영상을 분할하여 “영상의 위쪽에는 노란색 물체가, 아래쪽에는 푸른색 물체가 존재한다.”와 같이 영상을 표현한다. 이 방법은 객체 분할이 필요없지만 객체에 대한 정보를 표현하지 않기 때문에 정확한 검색은 불가능하다.

(3) 압축 영상에서의 검색

영상의 압축은 영상의 질과 크기가 증가함에 따라서 필수적인 기술이 되고 있는데 초기에는 일반 데이터와 마찬가지로 런 길이 부호화(run length encoding)나 사진을 사용한 압축 방법이 사용되어 왔는데 그다지 효율이 높지 못했다. 이러한 한계를 극복하기 위해서 영상 정보가 갖는 고유한 특성을 이용하기 시작했는데 이러한 예로는 VQ, wavelet, JPEG, 프랙탈 압축 방식 등이 있다. 이 중 가장 널리 쓰이는 JPEG은 DCT를 이용하여 영상을 고주파 성분과 저주파 성분으로 나누고 인간의 눈에 민감하지 않은 고주파 성분을 제거함으로써 압축 효과를 증대시키는 방법을 사용하고 있다^[11].

DCT의 특징 중의 하나가 영상을 주파수 성분으로 분리하는 것이기 때문에 한 DCT 블록의 계수값으로 블록의 평균 색상 값을 얻어낼 수 있다. 따라서 압축을 모두 해제하지 않고 허프만 복호화만



(그림 7) 압축 영상에서의 특징 정보 추출

하여 DCT 영역에서 DC 계수로부터 색상 특징을 추출해 낼 수 있다. 이 때 사용하는 색상 정보 추출 방법은 전 영상에 비해 크기만 1/64로 줄어들었으므로 일반적인 비압축 영상에서의 추출 방법과 동일한 방법을 사용할 수 있다. 또한 AC 계수들을 통해서 주파수 특성을 알 수 있기 때문에 질감 특징을 직접 추출할 수 있다. 이러한 방법은 압축을 모두 해제한 뒤에 특징을 추출하는 방법보다

시간이나 공간 비용이 적게 든다.

2. 영상 색인 기술

영상이 포함한 특징 정보는 일반 데이터와는 달리 다차원의 특성을 가지고 있다. 쉬운 예로 색상 특징의 경우 RGB 색상에서 하나의 색은 R, G, B로 이루어진 3차원 공간상의 한 점으로 볼 수 있다. 이러한 정보를 쉽게 색인하고 검색하기 위해서는 다차원 정보를 색인 하는 방법을 개발하거나 다차원 정보를 1차원의 선형 정보로 변환하는 방법이 필요하게 된다.

일반적인 트리 구조를 사용하여 각 차원별로 색인을 하게 되면 차원이 증가함에 따라서 색인에 필요한 공간이 기하급수적으로 늘어나게 되는데 이러한 현상을 보통 “dimensionality curse”라고 표현한다. 이러한 현상을 막기 위해서 다차원 정보를 쉽게 다루기 위한 자료 구조들이 연구되고 있다.

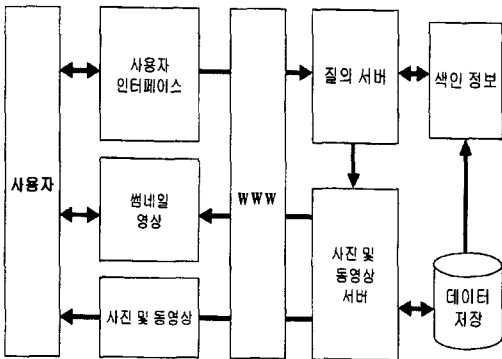
다차원 데이터를 다루기 위한 일반적인 방법인 R+ -트리 R* -트리 및 R-트리 등은 최소 경계 사각형(MBR: minimum bounding rectangle)을 사용하는데 이는 각 노드의 최대, 최소 범위를 결정하여 검색의 범위를 줄여주지만 차원이 증가할수록 최소 경계 사각형을 표현하기 위한 공간 역시 증가하므로 dimensionality curse 문제에 부딪히게 된다. TV-트리(telescopic vector tree)의 경우에는 특징 정보의 차원을 확장 또는 축소시킴으로써 색인과 검색 공간을 효율적으로 사용할 수 있다. 다차원의 특징 정보는 필요에 따라 고차원으로 확장되거나 축소되는데 주요한 특징 차원을 우선적으로 확장하기 위해서 특징 차원의 순서를 정하는 과정이 필요하다. 이 때 일반적으로 KL (Karhunen Loeve) 변환이나 DCT 변환 등이 사용되는데 KL 변환의 경우 데이터들이 가지는 주요한 축을 찾을 수 있어 이 축에 가까운 벡터부터 높은 순위를 주면 색인의 효과를 높일 수 있다. DCT 변환은 중요한 저주파 성분과 상대적으로 덜 중요한 고주파 성분으로 분리하여 순서를 결정하게 되므로 일반적으로 DCT 변환도 큰 효과를 가져온다^[12].

이외에도 SS-tree, X-tree를 비롯하여 다양한 다차원 데이터에 대한 색인 방법이 계속적으로 제안되고 있다^[13,14].

3. Web 기반 영상 검색 엔진의 구성

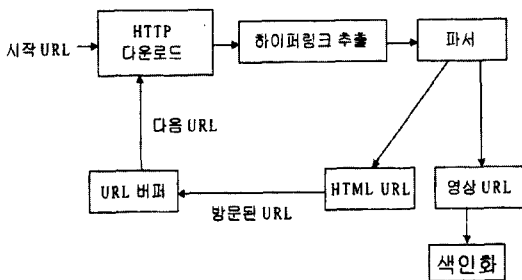
Web이 대중화되어 다양한 정보를 가진 사이트들이 생겨남에 따라서 사용자가 원하는 정보를 지닌 사이트를 찾아주는 검색 엔진이 등장하게 되었다. Web의 특징은 텍스트 정보 이외에도 다양한 영상과 멀티미디어를 포함할 수 있다는 것인데 수없이 산재한 web 상의 영상을 자동으로 색인하고 검색하는 시스템이 web 기반 영상 검색 엔진이다.

Web 검색 엔진은 사용자 인터페이스가 web 브



(그림 8) web 기반 영상 검색 엔진의 구조도

라우저 상에 구현되는 것으로 시스템 구성은 기존의 것과 크게 다르지 않다. Web 검색 엔진에서 중요한 것은 다양하게 퍼져있는 사이트를 자동적으



(그림 9) web에서 영상을 수집하는 로봇 에이전트의 구성^[15]

로 찾아다니면서 영상을 얻고, 그에 대한 자료를 수집하는 메커니즘이다.

이렇게 web에서 데이터를 수집하는 일을 담당하는 로봇 에이전트는 HTML 문서에 포함된 하이퍼링크를 통해서 재귀적으로 탐색하여 영상과 그에 대한 정보를 수집한다. HTML에 포함된 영상에는 텍스트 설명문이 포함되는 경우가 많기 때문에 이를 색인 정보로 사용하면 자동으로 텍스트 색인이 가능하기 때문에 검색의 정확도를 높일 수 있다.

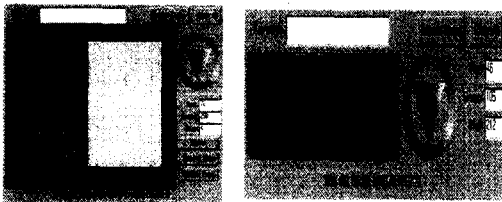
4. 데이터베이스와의 연동

ER-모델과 SQL로 대표되는 관계형 데이터베이스들은 정형적인 데이터들에 대해서 쉽게 관리하고 검색하는 능력을 보이지만 비정형적인 데이터들 즉 일반 텍스트 문서나 영상, 음성, 동영상들에 대해서는 저장 기능 외에는 검색이나 색인 기법을 제공하지 못하고 있다. 그러나 영상 데이터들이 증가함에 따라서 구조적이고 안전한 관리가 가능한 영상 데이터베이스의 필요성이 대두되었다. 앞서 설명한 다차원 색인 트리 기법이 영상 데이터의 색인 문제를 해결해 주기는 했지만 데이터베이스 관리 시스템이 가지는 구조화된 질의와 복구 기능과 같은 기능을 제공하는 통합적인 영상 데이터베이스 관리 시스템이 요구되기 시작한 것이다. 기존의 방법들은 관계형 데이터베이스에 객체 지향 요소를 포함시킨 객체 관계형 데이터베이스를 사용하거나, 기존의 데이터베이스를 영상 데이터에 맞게 확장하는 형식으로 이러한 문제들을 해결해 왔다^[16].

그러나 보다 효율적인 영상 정보 검색 시스템을 마련하기 위해서는 객체에 기반한 객체 지향 데이터베이스가 필수적이며 이에 알맞은 질의 최적화, 트랜잭션, 데이터 타입 정의, 풍부한 질의언어를 제공할 수 있어야 한다. 또한 내용 기반 정보에 적합한 유연한 질의 개념이 포함된다면 텍스트 기반 질의와 함께 뛰어난 성능을 보일 수 있을 것이다.

5. 사용자 인터페이스 기술

내용 기반 영상 검색은 다양한 영상의 특징 정보를 이용하여 검색을 수행하기 때문에 사용자 인터페이스의 구성도 검색 시스템의 매우 중요한 요소가 된다. 주로 사용하는 방법은 데이터베이스내의 영상 중에 임의의 예제를 사용자에게 보여주고 그와 유사한 영상을 출력하는 예제에 의한 검색(Query by Example)이다. 그러나 이런 방법으로는 사용자가 원하는 영상을 신속하고 정확하게 검색하기 어렵기 때문에 색상 비율을 설정하거나 색상 영역의 레이아웃을 그려주거나, 스케치 형태와 같은 형태를 그려주는 인터페이스가 사용되고 있다. 그림 10은 이와 같은 인터페이스의 예를 보여주고 있다.



〈그림 10〉 다양한 사용자 인터페이스의 예(IBM의 QBIC)

III. 결 론

지금까지 영상 정보 검색, 특히 내용 기반 영상 검색에 필요한 기술들의 현황을 알아보았다. 현재의 영상 검색 시스템들은 대부분 텍스트에 의한 검색만을 지원하고 있으며 내용 기반 검색 방법은 실험적인 시스템만이 존재한다. 이는 사용자들이 그 동안 텍스트를 이용한 검색 방법에 너무 익숙해져 있기 때문으로도 볼 수 있지만 아직 내용 기반 검색 기술이 실용적으로 쓰일 만큼 높은 성능을 보이고 있지 않기 때문이다. 따라서 좀더 효율적이고 정확한 검색 기술의 개발이 필수적인데 이를 위해서는 영상에서의 특징 추출과 다차원 특징

정보의 색인화 기술, 데이터베이스 연동 기술이 서로 유기적으로 연구되어야 한다.

최근의 연구 동향을 보면 색상과 질감 정보를 동시에 사용하여 질감 정보를 이용하여 영상에서 영역을 분리한 뒤 각 영역별로 특징을 추출하고 이를 윤곽이 분명하지 않은 작은 형태(blob)로 표현하는 방법이 제시되었는데, 이는 영상의 내용을 단순한 색상이나 질감 또는 형태 요소만으로 나타내지 않고 객체로 판단될 수 있는 후보 영역별로 특징을 추출했다는 점에서 매우 주목할 만하다^[17]. 기존의 색상 히스토그램이나 wavelet을 이용한 질감 분석, 모멘트와 같은 형태 특징 추출법 모두 저수준의 화소에서 전체 영상의 특징을 이끌어내는 방법이었는데 비해, 위와 같은 영역 분할에 의한 색인 및 검색 방법들은 화소 단위의 정보를 윤곽이 분명하지 않은 작은 형태를 이용하여 보다 의미 있는 정보를 추출할 수 있고, 각 영역의 공간적인 분포를 쉽게 알아낼 수 있기 때문에 공간 정보를 검색에 쉽게 사용가능 하다는 장점도 있다. 영역 분할이 객체의 영역과 유사한 정도까지 되지 않고 그 이상으로 더 많이 분할되거나 덜 분할되는 경우에는 올바른 검색이 되지 않기 때문에 정확한 영역 분할을 위한 알고리즘의 개발이 필요하다. 즉, 기존의 색상, 질감, 형태와 같은 저수준의 분석 방법에서 탈피하여 객체 인식이나 영상 이해와 같은 고수준의 영상 분석 기술에 대한 꾸준한 연구가 계속되어야 한다.

내용 기반 영상 검색 기술은 데이터들이 디지털화 되고 영상 정보의 양이 증가함에 따라서 중요한 기술의 하나로 자리잡아 가고 있다. 영상을 일반 텍스트와 동일한 인터페이스로 색인하고 검색하는 기술은 영상 미디어의 발전 뿐만 아니라 사회 전반, 특히 산업 및 경제분야 등에 필수적인 기술로 사용될 것이다.

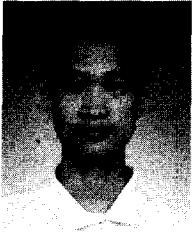
감사의 말씀

본 연구는 정보통신부 대학기초연구지원사업의 연구비 지원을 받았음.

참 고 문 헌

- [1] 이성환, “내용 기반 영상 및 비디오 검색 기술,” 한국정보과학회 인공지능연구회/뉴로컴퓨팅연구회/컴퓨터비전 및 패턴인식연구회 공동주최 ‘98 지능기술 튜토리얼 발표자료집, 숙명여자대학교, pp. 57-98, 1998년 5월.
- [2] 이 미숙, 황 분우, 이 성환, “내용 기반 영상 및 비디오 검색 기술의 연구 현황,” 한국정보과학회지, 제 15권 제 9호, pp. 10-19, 1997년 9월.
- [3] 황 분우, 노 형기, 이 성환, “색상 및 형태 정보를 이용한 내용 기반 영상 검색 시스템의 Web 상에서의 구현,” 한국정보과학회 봄 학술발표논문집, 제 25권 제 1호, pp. 607-609, 서울, 1998년 4월.
- [4] 김 영민, 최 송하, 황 분우, 양 윤모, 이 성환, “내용 기반 비디오 검색 기술,” 한국정보과학회지, 제 16권 제 8호, 1998년 8월.
- [5] Y. Gong, C. H. Chuan and G. Xiaoyi, “Image indexing and retrieval based on color histograms,” *Multimedia Tools and Applications*, Vol. 3, No. 3, pp. 133-156, 1996.
- [6] J. R. Smith and S.-F. Chang. “Tools and techniques for color image retrieval,” *Proc. SPIE Conf. on Electronic Imaging: Storage & Retrieval for Image and Video Databases IV*, Vol. 2670, pp. 426-437, San Jose, CA, 1996.
- [7] L. M. Kaplan, R. Murenzi and K. R. Namuduri, “Fast texture database retrieval using extended fractal features,” *Proc. SPIE Conf. on Electronic Imaging: Storage & Retrieval for Image and Video Databases IV*, Vol. 3312, pp. 162-173, San Jose, CA, 1997.
- [8] J. R. Smith. “Integrated spatial and feature image systems: retrieval, compression and analysis,” Ph.D. Thesis, Graduate School of Arts and Sciences, Columbia University, February, 1997.
- [9] S. Loncaric, “A survey of shape analysis techniques,” *Pattern Recognition*, Vol. 31, No. 8, pp. 983-1001, 1998.
- [10] B. M. Mehtre, M. S. Kankanhalli and W. F. Lee, “Shape measure for content based image retrieval: a comparison,” *Information Processing & Management*, Vol. 33, pp. 319-337, 1997.
- [11] W. B. Pennebaker and J. L. Mitchell, *JPEG still image data compression standard*, Van Nostrand Reinhold, 1993.
- [12] K. I. Lin, H. V. Jagadish and C. Faloutsos, “The TV-tree: an index structure for high-dimensional data,” *VLDB Journal*, Vol. 3, No. 4, pp. 517-542, 1994.
- [13] D. A. White and R. Jain, “Similarity indexing with the SS-tree,” *Proc. 12th IEEE Conf. on Data Engineering*, pp. 516-523, New Orleans, Louisiana, February 1996.
- [14] S. Berchtold, D. A. Keim and H. Kriegel, “The X-tree: an index structure for high-dimensional data,” *Proc. of the 22th VLDB Conf.*, pp. 28-39, 1996.
- [15] J. R. Smith and S.-F. Chang, “An image and video search engine for the world-wide web”, *Proc. SPIE Conf. on Electronic Imaging: Storage & Retrieval for Image and Video Databases IV*, Vol. 3312, pp. 84-95, San Jose, CA, 1997.
- [16] V. E. Ogle and M. Stonebraker, “Chabot: retrieval from a relational database of images,” *IEEE Computer*, pp. 40-48, 1995.
- [17] Berkeley Digital Library Project, “<http://elib.cs.berkeley.edu/vision.html>,”

저자 소개



盧炯基

1998년 고려대학교 컴퓨터학과 학사, 1998년~현재 고려대학교 컴퓨터학과 석사과정 재학중, <주관심분야: 내용 기반 영상 검색, 영상 데이터베이스 등>



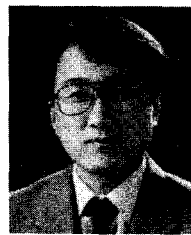
文鍾雙

1981년 서울대학교 계산통계학과 학사, 1983년 서울대학교 계산통계학과 석사, 1991년 Illinois Institute of Technology 전산학과 박사, 1983년~1985년 금성통신 연구소 연구원, 1993년~현재 고려대학교 전자 및 정보공학부 부교수, <주관심분야: 패턴 인식, 영상처리, 신경망 등>



黃本祐

1995년 성균관대학교 전자공학과 학사, 1997년 성균관대학교 전자공학과 석사, 1997년~현재 고려대학교 컴퓨터학과 박사과정 재학중, <주관심분야: 내용 기반 영상 검색, 컴퓨터 시각 등>



李晟煥

1984년 서울대학교 계산통계학과 학사, 1986년 한국과학기술원 전산학과 석사, 1989년 한국과학기술원 전산학과 박사, 1989년~1994년 충북대학교 컴퓨터학과과 조교수, 1995년~현재 고려대학교 컴퓨터학과 부교수, 1997년~현재 한국정보과학회 뉴로컴퓨팅연구회 위원장, 1998년~현재 고려대학교 인공시각연구센터 소장, <주관심분야: 패턴 인식, 컴퓨터 시각, 신경망 등>