영상의 에지 특징정보를 이용한 주석기반 및 내용기반 영상 검색 시스템의 구현

(Implementation of Annotation-Based and Content-Based Image Retrieval System using Edge Feature Information of Image)

이 태 동 † 김 민 구 ††
(Tae-Dong Lee) (Min-Koo Kim)

요 약 영상은 대용량적인 특성과 비정형적인 특성을 가지고 있으므로 신속하고 효율적으로 영상을 검색하기 위해 영상의 정확한 특징정보를 추출하여 검색 시스템을 구축하여야 한다. 영상 검색 시스템은 텍스트 기반의 전통 데이타베이스와는 다른 모델링 방법과 검색방법을 사용한다. 따라서, 영상 검색 시스템에서의 검색속도와 정확도를 향상시키기 위해서는 새로운 영상 데이타베이스 생성기법과 효율적인 검색기법이 필요하다. 본 논문에서는 입력 영상으로부터 검색에 사용되는 에지 특징정보 추출을 위해 라플라시안 마스크와 입력 영상을 컨벌루션하여 에지의 외곽선 데이타를 추출하였으며, 그리고 추출한 에지 특징정보와 메타데이타로 영상 데이타베이스를 생성하여 신속하고 효율적으로 영상을 검색할 수 있도록 주석기반 및 내용기반 영상 검색 시스템을 구현하였다. 주석기반 및 내용기반 영상 검색 시스템은 영상의 하위레벨에 표현된 내용기반 에지 특징정보와 특징정보 추출이 어려운 상위레벨에 표현된 주석기반 에지 특징정보를 영상의 색인으로 구성하여 사용하기 때문에 영상 컨텐츠 검색의 성능을 향상시킬 수 있다. 마지막으로 본 논문에서 제시한 영상 검색 시스템은 메타데이타에 의해 영상 데이타베이스를 구축하므로 정확한 영상 컨텐츠 정보의 축적관리와 영상의 정보공유 및 재이용이 가능하다.

Abstract Image retrieval system should be construct for searching fast, efficient image be extract the accurate feature information of image with more massive and more complex characteristics. Image retrieval system are essential differences between image databases and traditional databases. These differences lead to interesting new issues in searching of image, data modeling. So, cause us to consider new generation method of database, efficient retrieval method of image. In this paper, To extract feature information of edge using in searching from input image, we was performed to extract the edge by convolution Laplacian mask and input image, and we implemented the annotation-based and content-based image retrieval system for searching fast, efficient image by generation image database from extracting feature information of edge and metadata. We can improve the performance of the image contents retrieval, because the annotation-based and content-based image retrieval system is using image index which is made up of the content-based edge feature extract information represented in the low level of image and annotation-based edge feature information represented in the high level of image. As a conclusion, image retrieval system proposed in this paper is possible the accurate management of the accumulated information for the image contents and the information sharing and reuse of image because the proposed method do construct the image database by metadata.

† 정 회 원 : 경문대학 컴퓨터응용과 교수
LTD@kmc.ac.kr
†† 종신회원 : 아주대학교 정보및컴퓨터공학부 교수
minkoo@madang.ajou.ac.kr
논문접수 : 2001년 3월 13일
심사완료 : 2001년 8월 13일

1. 서 론

오늘날 통신 및 컴퓨터 기술의 발달로 미디어의 새로운 혁명인 멀티미디어 시대에 도래하게 되었다. 멀티미디어 정보는 인터넷의 폭발적인 보급과 웹의 출현에 힘

입어 수요가 급증하고 있다. 최근 웹(World Wide Web)의 급속한 발전에 힘입어 홈쇼핑, 전자상거래, VOD (Video On Demand), TV중계, 원격수업, 뉴스전송, 원격진료, 인터넷 방송 등과 같은 다양한 응용 분야에서 영상 검색 서비스가 필수적으로 요구되고 있다. 그러나 영상 정보는 기존의 텍스트 기반의 데이타에 비하여 대용량적인 특성과 비정형적인 특성을 가지고 있어서 신속하고 효율적인 검색에 많은 어려움이 따른다. 따라서 방대한 영상 정보를 검색하기 위한 대용량의 데이타베이스 생성과 적합한 검색기법에 대한 연구가 필요하다[1]. 이에 따라 영상 정보를 효과적으로 저장하고 검색할 수 있는 연구가 활발히 진행되고 있다[2]. 본 논문의 연구 배경 및 목적은 멀티미디어 기술의 발전과 웹 컨텐츠의 증가로 인한 다양한 사용자의 요구에 부응하고 개인용 컴퓨터에서 영상 정보를 손쉽게 관리하고, 신속하고 효율적으로 검색할 수 있도록 하는데 그 목적이 있다.

따라서 본 논문에서는 영상의 내용기반 에지 특징정보와 주석기반 에지 특징정보를 영상 데이타베이스의 색인으로 구성하여 신속하고 효율적으로 원하는 영상을 정확하게 검색할 수 있도록 하기 위해 주석기반질의, 내용기반질의, 주석기반 및 내용기반 통합질의와 같은 다양한 검색방법을 제공하는 영상 검색 시스템을 구현하여 제시한다. 그리고, 이를 위해 영상의 하위레벨에 표현된 내용기반 에지 특징정보와 영상의 특징정보 추출이 어려운 상위레벨에 표현된 주석기반 에지 특징정보, 즉 객체, 객체 유형, 객체 위치, 객체 이벤트, 객체 내용 등을 추출하는 방법을 제시한다. 본 논문에서 제안하는 영상 검색 시스템은 영상의 하위레벨에 표현된 내용기반 에지 특징정보와 상위레벨에 표현된 주석기반 에지 특징정보를 이용하여 다양한 검색방법을 제공함으로서 신속하고 효율적으로 원하는 영상을 정확하게 검색할 수 있다. 그리고 이들의 에지 특징정보를 영상의 색인으로 구성하기 때문에 영상 컨텐츠 검색의 성능을 향상시킬 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구를 통하여 영상 검색방법에 관하여 살펴본다. 3장에서는 영상의 에지 특징정보를 이용한 주석기반 및 내용기반 영상 검색 시스템의 구축 방법과 생성기법을 제시한다. 그리고, 4장에서는 본 논문에서 구현한 영상 검색 시스템에 대한 다양한 실험결과를 기존의 사례연구와 비교 분석하고, 마지막으로 5장에서는 결론 및 향후 연구과제를 제시한다.

2. 영상 검색방법 관련 연구

이 장에서는 관련 연구를 통하여 영상 검색방법에 관하여 살펴본다. 영상 검색방법으로는 주석기반 검색과 내용기반 검색, 그리고 주석기반 검색과 내용기반 검색을 통합한 검색 방법이 있다. 주석기반 영상 검색방법은 제한된 범위 내에서는 효율적인 검색 결과를 제공한다. 반면에 대용량일 경우에는 주석 개발의 어려움과 영상 내용에 대한 주관적 해석의 차이, 그리고 주석 할당의 일관성 결여와 색인 과정의 많은 시간과 고비용 등의 문제점을 가지고 있다. 이러한 단점을 극복하기 위하여 영상 정보로부터 기본적인 특징을 추출하여 이를 기반으로 영상을 검색하는 내용기반 영상 검색(Content-Based Image Retrieval) 방법을 많이 연구하고 있다[3]. 영상 정보로부터 추출할 수 있는 영상의 특징정보는 패턴의 변화, 색상(Color), 질감(Texture), 모양(Shape), 영상 내 개체간의 공간상 위치관계 등을 말한다. 이러한 영상내의 특징기반 영상 검색방법은 단일 특징만을 이용할 경우 정확한 검색이 이루어질 수 없으며 대용량의 데이타베이스에 적용하면 검색 효율이 떨어질 수 있다.

영상 정보를 표현하고 검색하기 위한 첫 번째 방식은 주석기반 영상 검색방법으로 각각의 영상 정보에 자동 인식이 어려운 의미정보를 수동으로 먼저 파악한 후, 이를 사용자가 직접 텍스트로 주석을 부여하여 저장한 후 검색 시 미리 부여된 주석을 이용하여 영상을 검색하는 방법이다[4,5]. 이 방법은 사용자가 직접 영상을 보면서 내용을 주석으로 처리할 수 있기 때문에 영상의 내용을 정확하게 모델링하고 검색할 수 있다는 장점이 있다. 그러나, 각각의 영상에 대해 문자로 사용자가 일일이 주석을 부여하여야 하므로 많은 시간과 노력을 필요로 하며 주석의 양이 방대하게 증가 할 수 있다는 단점이 있다. 또한 영상 내용을 사용자의 관점에 따라 주관적으로 주석을 부여할 수 있으므로 검색 시 영상의 내용과 다르게 질의를 할 수 있다는 단점이 있다. 그리고, 영상 정보의 시각적인 의미정보를 정확하게 문자로 표현하기가 어렵다.

두 번째 방식은 내용기반 영상 검색방법으로 영상 정보자체로부터 그 정보가 가지고 있는 특징정보들을 자동 또는 반자동적으로 추출하여 검색하는 방법이다[6, 7,8]. 내용기반 검색방법은 영상 자체의 시각적 특징, 즉 영상 특징정보를 자동 또는 반자동적으로 추출하여 검색함으로서 기존의 주석기반 검색방법의 단점을 해결할 수 있다. 그러나, 영상의 특징정보를 정확하게 추출하기 위한 알고리즘 개발과 정확한 분류가 어렵다는 단점이 있다.

세 번째 방식은 주석기반 영상 검색방법과 내용기반 영상 검색방법을 통합하여 사용하는 방법이다[4,5]. 즉,

영상의 각 객체들 사이의 시간관계, 공간관계 등의 하위레벨 특징정보들은 내용기반 영상 검색방법을 이용하여 자동 또는 반자동적으로 특징정보를 추출해 내고, 그리고 자동 또는 반자동으로 추출이 불가능한 사건, 장소 등의 상위레벨 의미정보들은 주석기반 영상 검색방법을 이용하여 사용자가 직접 의미정보를 기술해 주는 방법이다.

관련 연구에서 고찰한 바와 같이 영상내의 특징기반 영상 검색방법은 단일 특징만을 이용할 경우 정확한 검색이 이루어질 수 없으며 대용량의 데이타베이스에 적용하면 검색 효율이 떨어질 수 있다. 이들의 단점을 극복하기 위해 검색방법들간의 상호 보완적인 역할이 필요하다. 따라서 이들 검색방법들간의 상호 보완적인 역할을 갖도록 하기 위해서는 영상 검색 시스템을 구현함에 있어서 다양한 검색방법을 제공하여야만 상호 보완적인 역할이 가능하다. 이와 같이 다양한 검색방법을 제공하기 위해서는 정확한 영상의 하위레벨에 표현된 내용기반 특징정보와 상위레벨에 표현된 주석기반 특징정보를 반드시 추출하여야 한다. 이를 위해 본 논문에서는 영상의 하위레벨에 표현된 내용기반 에지 특징정보와 상위레벨에 표현된 주석기반 에지 특징정보를 추출하는 방법을 제시한다. 그리고 영상의 내용기반 에지 특징정보와 주석기반 에지 특징정보를 영상 데이타베이스의 색인으로 구성하여 신속하고 효율적으로 원하는 영상을 정확하게 검색할 수 있도록 하기 위해 주석기반질의, 내용기반질의, 주석기반 및 내용기반 통합질의와 같은 다양한 검색방법을 제공하는 영상 검색 시스템을 구현하여 제시한다.

3. 주석 및 내용기반 영상 검색 시스템

이 장에서는 영상의 내용기반 에지 특징정보와 주석기반 에지 특징정보를 영상 데이타베이스의 색인으로 구성하여 영상 검색 시스템을 구축하는 방법을 기술한다. 이를 위해 영상의 하위레벨에 표현된 내용기반 에지 특징정보와 영상의 상위레벨에 표현된 주석기반 에지 특징정보, 즉 객체, 객체 유형, 객체 위치, 객체 이벤트, 객체 내용 등을 추출하는 방법과 이들의 에지 특징정보를 메타데이타로 제공하여 영상 데이타베이스의 색인을 구성하는 방법을 기술한다. 그리고 영상 데이타베이스의 색인을 이용하여 신속하고 효율적으로 원하는 영상을 검색할 수 있도록 하기 위한 방법으로 주석기반질의, 내용기반질의, 주석기반 및 내용기반 통합질의와 같은 다양한 영상 검색방법을 모델링하여 기술한다.

3.1 영상 검색 시스템의 전체적인 구조

본 논문에서 제안하는 영상 검색 시스템의 전체적인 구조는 그림 1과 같이 크게 영상 인덱싱 처리 부분과 영상 데이타베이스 관리 부분, 그리고 사용자 인터페이스 부분으로 구성한다. 영상 인덱싱 처리 부분은 영상의 에지 특징정보를 추출하여 영상 데이타베이스의 인덱싱을 생성한다. 그리고, 영상 데이타베이스 관리 부분은 영상의 에지 특징정보 메타데이타와 영상 데이타를 생성하고 관리한다. 그리고 사용자 인터페이스 부분은 Image Processing, Image Browser, Image Query Processing 기능의 사용자 인터페이스를 제공한다.

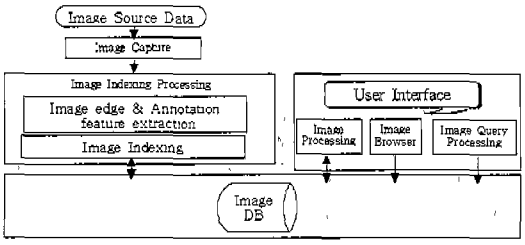


그림 1 영상 검색 시스템의 전체적인 구조

3.2 메타데이타 및 속성(Metadata and Attribute) 데이타 모델링

메타데이타는 영상 정보의 내용을 구조적으로 체계화한 데이타이다. 본 절에서는 영상 제작과정의 일반적인 특징 정보와 시각적인 특징(Visual Feature) 정보, 그리고 영상의 주석기반 에지 특징정보와 내용기반 에지 특징정보에 대한 Keyword, Object Name, Object Type, Object Position, Object Event, Object Content, Object Feature 등의 Annotation & Content Edge 특징정보를 기반으로 메타데이타를 분류한 후 이를 영상 데이타베이스의 메타데이타 클래스와 속성으로 정의한다. 영상 데이타베이스에서 모든 컨텐츠의 정보는 메타데이타를 통하여 표현되고, 메타데이타를 통하여 실제 데이타에 접근할 수 있다. 메타데이타는 사용자와 영상 데이타를 상호 연결하는 매개자 역할을 수행한다.

3.3 영상의 에지 특징정보 추출과정

(1) Visual Image class

Class	Attribute	Content
Visual_Image_class	Visual_Image_to_Image_Feature	Image Feature class pointer(Image_Feature_id)
	Visual_Image	Visual Image data

(2) Image Feature class

Class	Attribute	Content
Image_Feature_class	Image_Feature_id	Image Feature class(Unique identifier)
	Image_Feature_Object_name	Image Feature extraction(Object name)
	Image_Feature_Object_type	Image Feature extraction(Object type)
	Image_Feature_Object_position	Image Feature extraction(Object position)
	Image_Feature_Object_event	Image Feature extraction(Object event)
	Image_Feature_Object_content	Image Feature extraction(Object content)
	Image_Edge_Feature_extraction	Image Feature extraction(Edge extraction)

영상의 에지는 객체의 위치, 객체의 모양과 크기, 표면의 무늬 등에 대한 정보를 알려준다. 에지는 영상의 픽셀크기가 크게 변화는 위치에 존재한다. 영상의 에지 검출(Edge Detection) 연산자는 라플라시안 연산자(Laplace Operator), 프리윗 연산자(Prewitt Operator), 소벨 연산자(Sobel Operator), 로버츠 연산자(Roberts Operator), 커쉬 연산자(Kirsch Operator), 로빈슨 연산자(Robinson Operator) 등이 있다[9,10]. 본 논문에서는 영상의 에지 검출 시 잡음에는 민감하지만 다른 기법들보다 상대적으로 해당 에지 검출 성향이 높고 모든 방향의 에지 검출이 가능한 라플라시안 마스크(Laplacian Mask)를 이용한다[9,10]. 영상의 에지 특징정보 추출과정은 먼저 입력 영상으로부터 내용기반 영상 검색에 사용되는 에지 특징정보를 추출하기 위하여 라플라시안 마스크와 입력 영상을 컨벌루션(Convolution)하여 에지 외곽선 데이타를 추출하고, 이 에지 외곽선 데이타에 대하여 세 채널을 갖는 완전 컬러 영상의 각 채널별 RGB 픽셀 값을 추출한다. 이러한 결과에 의해 생성된 에지 외곽선 데이타의 RGB 픽셀 값을 각 채널별로 평균값을 구한 후, 이렇게 구해진 각 채널별 RGB의 평균값을 최종적으로 내용기반 영상 검색에 사용되는 메타데이타로 제공하여 영상 검색 시스템의 인덱스로 구성한다. 그림 2는 에지 추출에 일반적으로 많이 사용하는 라플라시안 마스크의 예이다[9,10].

1	-2	1
-2	4	-2
1	-2	1

(a) Mask 1

0	-1	0
-1	4	-1
0	-1	0

(b) Mask 2

-1	-1	-1
-1	8	-1
-1	-1	-1

(c) Mask 3

그림 2 라플라시안 마스크 예

그림 2와 같이 라플라시안 마스크는 인접 픽셀간의 값 차이가 별로 없는 영역에 마스크가 있을 경우에는 거의 0에 가까운 값이 출력되며 인접 픽셀간의 값 차이가 큰 부분에 마스크가 있을 경우에는 큰 값이 출력된다. 그리고, 방향성이 없는 에지를 추출하기 위한 라플라시안 마스크의 기본 요건은 마스크의 가로 세로 크기가 같고 홀수이어야 하며, 모든 가중치의 합은 0이 되어야 한다. 그리고 중심점의 가중치는 양수이어야 하며, 중심점 외의 가중치는 음수가 들어가야 한다. 또한 에지의 방향성을 고려하지 않을 경우 마스크의 가중치는 중심점을 기준으로 상하좌우가 대칭이어야 한다. 그러나 마스크의 가중치는 에지 검출 연산자들의 특성에 따라 다른 값을 갖게 될 수 있으며, 사용 목적에 따라 마스크의 크기와 모양이 다를 수 있다.

3.3.1 영상 처리의 컨벌루션

영상 처리의 컨벌루션은 영상의 에지 추출을 위한 도구이며, 입력 픽셀의 이웃에 있는 픽셀들의 가중치 합이다. 가중치를 곱하여 합을 구하는 계산은 상수 값에 의해 곱해진 픽셀 밝기들의 합과 동일하기 때문에 선형처리라고 한다[9,10]. 여기서 상수 값은 컨벌루션 마스크의 가중치이다. 일반적으로 가중치는 픽셀의 평균 처리에 의해 구해지며, 컨벌루션 마스크는 입력 영상의 각 영상을 중심으로 이동하면서 새로운 출력 픽셀을 생성한다. 그림 3은 입력 영상의 중심 픽셀에 3×3 라플라시안 마스크를 대응시켜 컨벌루션을 수행한 예이다.

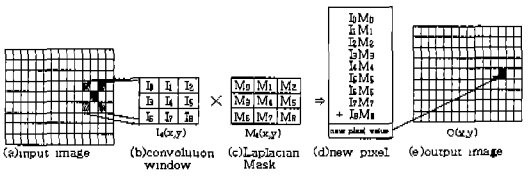


그림 3 3×3 라플라시안 마스크의 컨벌루션

그림 3에서와 같이 영상 처리의 컨벌루션은 입력 영상의 중심 픽셀에 에지 검출 마스크를 대응시킨 뒤 마스크의 인수와 마스크에 대응하는 영상의 픽셀 값을 곱하여 이를 합산하는 것이다. 여기서 출력되는 결과 값은 출력 영상 내에 입력 영상의 중심 픽셀 위치와 동일한 위치에 생성된다.

3.3.2 에지 외곽선 데이타 추출

영상의 에지는 입력 영상에 포함된 픽셀 값이 급격하게 변하는 부분이다. 에지 외곽선 데이타 추출은 라플라시안 마스크에 의해 얻을 수 있다[9,10]. 에지 외곽선 데이타 추출 연산은 픽셀 밝기의 기울기를 기초로 한다[9,10]. 본 논문에서는 다른 기법들보다 상대적으로 해당 에지 검출 성향이 높고 모든 방향의 에지 검출이 가능한 대표적인 2차 미분 연산자인 그림 2의 라플라시안 마스크를[9,10] 적용하여 에지 외곽선 데이타를 추출한다. 그림 2의 라플라시안 마스크는 모든 가중치의 합이 0이 되므로 입력 영상을 고려하지 않고 인접 픽셀과의 차이만을 고려하여 에지 외곽선 데이타를 추출할 수 있다. 에지 외곽선 데이타는 입력 영상과 라플라시안 마스크를 컨벌루션하여 식 (1)의 공간 컨벌루션 처리 수식에 의해 입력 영상의 모든 픽셀을 대상으로 에지 외곽선 데이타 출력 값 $E_0(x,y)$를 구한다.

그리고, 본 연구에서는 입력 영상과 라플라시안 마스

크를 컨벌루션하여 에지 외곽선 데이타를 추출할 때 식 (2)와 같이 에지 외곽선 데이타 출력 값 $E_v(x,y)$의 크기를 비교하여 만약에 $E_v(x,y)$의 값이 음수일 경우에는 $E_v(x,y)$의 값을 0으로 치환하고, 그리고 $E_v(x,y)$의 값이 255이상일 경우에는 $E_v(x,y)$의 값을 255로 치환하여 영교차 처리한다. 에지 외곽선 데이타 출력 값 $E_v(x,y)$의 크기가 0인 것은 무시해도 될 만큼의 작은 변화를 가진 경우로서 보다 선명한 에지를 추출할 수 있다. 그리고 컨벌루션 함수를 구현할 때 영상의 가장자리 부분에 대한 처리는 여러 방법들이 있으나 본 연구에서는 컨벌루션 윈도우가 영상과 중첩되지 않는 (1,1) 위치의 픽셀에서 컨벌루션이 시작되도록 하여 (719,479) 위치의 픽셀에서 컨벌루션이 끝나도록 영상의 가장자리 부분을 처리한다. 그리고 출력 영상에서 컨벌루션 된 영상의 경계 부분 픽셀들은 입력 영상과 같은 해상도를 가진 영상으로 생성하기 위해 가장자리 부분을 복사하여 처리한다.

$$\begin{aligned} E_v(x,y) = & (I_0(x-1,y-1)\times M_0)+(I_1(x,y-1)\times M_1) \\ & +(I_2(x+1,y-1)\times M_2)+(I_3(x-1,y)\times M_3) \\ & +(I_4(x,y)\times M_4)+(I_5(x+1,y)\times M_5) \\ & +(I_6(x-1,y+1)\times M_6)+(I_7(x,y+1)\times M_7) \\ & +(I_8(x+1,y+1)\times M_8) \end{aligned} \quad (1)$$

$$E_v(x,y)=\begin{cases} 255, & \text{if } |E_v(x,y)| > 255 \\ 0, & \text{if } |E_v(x,y)| < 0 \end{cases} \quad (2)$$

3.3.3 영상의 에지 특징정보추출

영상 정보로부터 추출할 수 있는 영상의 특징정보는 패턴의 변화, 색상(Color), 질감(Texture), 모양(Shape), 영상 내 개체간의 공간상 위치관계 등을 말한다. 이런 영상 정보로부터 추출 가능한 특징정보는 영상의 모양, 색상, 질감 등과 같은 기본적인 특징(Primitive Feature) 정보와 영상의 다양한 의미정보(Semantic Information)를 문자 형태로 표현하는 논리적인 특징(Logical Feature) 정보로 나눌 수 있다[11]. 이들 특징정보 중에서 색상 정보를 가장 보편적으로 많이 사용되고 있다[12]. 그리고 색상 정보를 이용한 질의 검색에는 평균 RGB를 이용한 질의 검색이 있다[13]. 평균 RGB를 이용한 검색은 평균 색상으로 영상을 검색하므로 원하는 색상을 가진 영상이 아님에도 검색이 될 수 있다는 단점이 있다. 이러한 영상내의 특징기반 영상 검색방법은 단일 특징만을 이용할 경우 정확한 검색이 이루어질 수 없으며 대용량의 데이타베이스에 적용하면 검색효율이 떨어질 수 있다. 이들의 단점을 극복하기 위해서는 검색방법들간의 상호 보완적인 역할이 필요하다. 따라서 검색방법들간의 상호 보완적인 역할을 갖도록 하기 위해서는 영상 검색 시스템을 구현함에 있어서 다양한 검색방법을 제공하여야만 상호 보완적인 역할이 가능하다. 이와 같이 다양한 검색 방법을 제공하기 위해서는 정확한 영상의 하위레벨에 표현된 내용기반 특징정보와 상위레벨에 표현된 주석기반 특징정보를 반드시 추출하여야 한다. 이들 영상의 하위레벨에 표현된 내용기반 특징정보와 상위레벨에 표현된 주석기반 특징정보를 이용하여 영상 검색 시스템에 다양한 검색방법을 제공함으로서 신속하고 효율적으로 원하는 영상을 정확하게 검색할 수 있다. 따라서 본 절에서는 내용기반 특징정보인 영상의 하위레벨에 표현된 에지의 특징정보를 추출하는 방법을 기술한다. 본 연구에서는 이를 위해 영상의 특징정보 중에서 보편적으로 가장 많이 사용되고 있는 색상 특징정보를 내용기반 특징정보로 이용한다. QBIC[4,5]과 같은 시스템에서는 식 (3)과 같이 입력 영상의 전체 픽셀에 평균 RGB 방법을 이용하여 고차원 컬러 히스토그램을 3차원 컬러 히스토그램으로 변환한 색상 정보를 이용하고 있다.

$$\begin{aligned} R_{avg} &= \left(\frac{1}{N}\right)\sum_{p=1}^{N} R(p) \\ G_{avg} &= \left(\frac{1}{N}\right)\sum_{p=1}^{N} G(p) \\ B_{avg} &= \left(\frac{1}{N}\right)\sum_{p=1}^{N} B(p) \end{aligned} \quad (3)$$

p:픽셀, N:전체픽셀수

이 방법은 입력 영상에 대한 색상의 분포를 특징으로 이용하기 때문에 입력 영상에 유사한 색상 분포를 지니는 특징일 경우에는 후보 영상이 선택되어질 수 있다는 문제점이 있다. 또한 고차원 컬러 히스토그램의 급격한 저차원 컬러 히스토그램 변환 과정으로 인한 정보 손실이 존재할 수 있다. [14]에서는 이런 문제점들을 해결하기 위해 전체 영상을 같은 크기의 타일로 분할하여 각 타일에 대한 평균 RGB 값을 구한 후 영상의 타일 평균 RGB 값을 영상의 인덱스로 구성하는 방법을 제시하고 있다.

본 논문에서는 QBIC[4,5]과 같은 시스템의 문제점들을 극복하기 위해 영상의 에지 외곽선 데이타를 추출하여 이 에지 외곽선 데이타에 대한 각 채널별 RGB의 평균값을 구한 후 에지 외곽선 데이타의 각 채널별 RGB 평균값을 영상의 인덱스로 구성하는 방법을 제시한다. 영상의 에지 특징정보 추출 방법은 먼저 입력 영상으로부터 내용기반 영상 검색에 사용되는 에지 특징정보를 추출하기 위하여 3.3.2절에서 기술한 방법에 의해 에지 외곽선 데이타를 추출한다. 그리고, 이 에지 외곽선 데이타에 대하여 세 채널을 갖는 완전 컬러 영상에 대한 각 채널별 RGB의 픽셀 값을 추출한다. 에지 외곽선 데이타에 대하여 RGB 픽셀 값을 추출하기 위해선 raw 데이타로 변환해야 한다. raw 데이타는 영상 정보와 RGB 데이타 부분으로 구분된다. raw 데이타의 정보 부

분은 영상의 폭, 영상의 높이, 영상의 색상 정보를 Space(0x20)나 Line Feed(0x0a)로 구분하여 나타낸다. 나머지 RGB 데이타 부분은 각 채널별 3byte씩 RGB 값을 나타낸다. 이러한 변환 결과에 의해 생성된 에지 외곽선 raw 데이타로부터 추출한 각 채널별 RGB 픽셀 값을 식 (3)과 같이 평균 RGB 방법을 이용하여 각각 더한 후 에지 외곽선 데이타의 전체 픽셀 값으로 나누어 각 채널별 RGB의 평균값을 구한다. 이러한 결과로 구해진 각 채널별 RGB의 평균값을 최종적으로 내용기반 영상검색에 사용되는 메타데이타로 제공하여 영상 검색 시스템의 인덱스로 구성한다. 알고리즘 1은 완전 컬러 영상으로부터 에지 특징정보를 추출하는 과정이다.

알고리즘 1 에지 특징정보 추출과정

- 입력 : 완전 컬러 영상(Full Color Image)
- 출력 : 추출한 에지의 각 채널별 평균 RGB 값

1. 입력 영상과 라플라시안 마스크를 컨벌루션하여 에지 외곽선 데이타를 추출한다.
   1.1 컨벌루션을 처리할 때 결과 값의 픽셀을 영교차 처리한다.
   1.2 컨벌루션 함수를 구현할 때 영상의 가장자리 부분을 처리한다.
2. 1의 과정에 의해 추출된 에지 외곽선 데이타에 대하여 각 채널별로 RGB의 픽셀 값을 각각 추출한다.
3. 2의 과정에 의해 추출된 각 채널별 RGB의 픽셀 값을 식 (3)과 같이 평균 RGB 방법을 이용하여 각각 더한 후 에지 외곽선 데이타의 전체 픽셀 값으로 나누어 각 채널별 RGB의 평균값을 구한다.
4. 3의 과정에 의해 구해진 각 채널별 RGB의 평균값을 에지의 특징 데이타로 생성한다.
5. 4의 과정을 통해 생성된 에지 특징 데이타를 최종적으로 검색에 사용되는 에지 특징정보의 메타데이타로 제공하여 영상 검색 시스템의 인덱스로 구성한다.

QBIC[4,5]과 같은 시스템에서는 영상의 전체 픽셀에 대한 평균 RGB 값을 구하여 고차원 컬러 히스토그램을 3차원 컬러 히스토그램으로 변환한 색상의 분포를 특징으로 이용하기 때문에 영상에 유사한 색상 분포를 지닌 특징일 경우에는 후보 영상이 선택되어질 수 있다는 단점이 있다. [14]에서는 이런 단점들을 해결하기 위해 전체 영상을 같은 크기의 타일로 분할하여 각 타일에 대한 평균 RGB 값을 구한 후 영상의 타일 평균 RGB 값을 영상의 인덱스로 구성하는 방법을 제시하였다. 반면에 본 논문에서는 알고리즘 1에서와 같은 과정을 통하여 에지 외곽선 데이타에 대한 각 채널별 RGB의 평균값을 영상 검색 시스템의 인덱싱으로 구성함으로서 대규모의 영상 데이타베이스에서 효율적이고 신속하게 원하는 영상 검색이 가능하다. 또한, QBIC[4,5]과 같은 시스템에서는 고차원 컬러 히스토그램의 급격한 저차원 컬러 히스토그램 변환과정으로 인한 정보 손실이 존재할 수 있다는 단점이 있다. 따라서 질의 영상의 전체적인 모양이 데이타베이스에 저장된 영상과 유사하더라도 고차원 컬러 히스토그램의 급격한 저차원 컬러 히스토그램 변환과정으로 인한 정보 손실에 의해 색상 구성이 다르게 될 수 있으므로 원하는 영상을 제대로 찾지 못할 수 있다. 본 논문에서 제안하는 방법은 영상의 에지 특징정보를 표현하고 있는 에지 외곽선 데이타에 대하여 각 채널별 RGB의 평균값을 에지 특징정보의 메타데이타로 제공하므로 사용자의 질의 영상과 에지 모양이 유사한 영상을 일관성 있게 검색할 수 있다.

3.4 영상의 주석기반 에지 특징정보 추출

본 논문에서는 영상의 특징정보 추출이 어려운 상위 레벨에 표현된 객체, 객체 유형, 객체 위치, 객체 이벤트, 객체 내용 등의 주석기반 에지 특징정보를 메타데이타로 제공하여 영상 데이타베이스의 색인으로 구성한다. 본 논문에서 추출되어지는 에지 외곽선 데이타에 대한 에지 추출은 에지에 대한 특징정보를 표현하고 있으며, 또한 에지에 대한 시각적이고 공간적인 특징정보를 나타내고 있다. 따라서, 본 연구에서는 주석기반 에지 특징정보를 추출할 때 이들의 에지 외곽선 데이타에 대한 에지의 시각적인 특징정보를 이용하여 객체, 객체 유형, 객체 위치, 객체 이벤트, 객체 내용 등의 주석기반 에지 특징정보를 추출한다. 주석기반 에지 특징정보를 추출하는 방법은 먼저 Image Feature Extraction 사용자 인터페이스를 통하여 해당 입력 영상에 대한 에지 외곽선 데이타의 시각적인 특징정보를 추출한다. 이러한 결과에 의해 추출된 에지 외곽선 데이타의 시각적인 특징정보를 Image Feature Extraction 입력 영상과 직접 비교하여 입력 영상내의 에지 외곽선 모양에 해당하는 객체에 한정하여 사용자가 직접 주석기반 에지 특징정보를 수동으로 추출하여 모델링한다. 이러한 모델링 결과에 의해 추출된 객체, 객체 유형, 객체 위치, 객체 이벤트, 객체 내용의 에지 의미정보를 메타데이타로 제공하여 영상 데이타베이스의 색인으로 구성한다. 이와 같이 본 논문에서 제시하는 주석기반 에지 특징정보 추출 방법은 추출된 에지 외곽선 데이타에 대한 시각적인 특징정보를 입력영상에 비교하여 주석기반 에지 특징정보를 직접 사용자가 추출하기 때문에 주석기반 에지 특징정보를 쉽게 추출할 수 있으며, 또한 주석기반 에지 특징정보를 정확하게 모델링할 수 있다.

3.5 주석기반 및 내용기반 영상 인덱싱

대규모 영상 데이타베이스에서 영상을 신속하고 효율

적인 질의와 검색을 지원하기 위해서는 영상 색인기법이 필수적이다. 영상 검색의 궁극적인 목적은 사용자가 원하는 조건을 만족하는 영상을 신속하고 정확하게 찾아내는데 있다. 이러한 관점에서 본 논문에서는 영상의 에지 특징정보 추출과정에서 추출한 영상의 주석기반 에지 특징정보와 내용기반 에지 특징정보의 메타데이타를 영상의 색인으로 구성하여 주석기반질의, 내용기반질의, 그리고 주석기반 및 내용기반 통합질의와 같은 다양한 검색방법을 제공한다. 이와 같이 주석기반 에지 특징정보와 내용기반 에지 특징정보를 영상의 색인으로 구성하여 영상 검색 시스템을 구현함에 있어서 다양한 검색방법을 제공할 수 있다. 본 논문에서 제안하는 영상 검색 시스템은 영상의 하위레벨에 표현된 내용기반 에지 특징정보의 메타데이타 유형과 영상의 특징정보 추출이 어려운 상위레벨에 표현된 주석기반 에지 특징정보의 메타데이타 유형을 영상의 색인으로 활용하기 때문에 영상 검색의 성능을 향상시킬 수 있다.

3.6 영상 질의 및 사용자 인터페이스

영상 데이타는 질의 방식에 따라 영상 데이타베이스에 저장된 데이타와 동일한 유형의 데이타를 질의 조건 값으로 받아들여 질의를 수행하는 직접 질의(Direct Query) 방식과 영상 데이타베이스에 저장된 데이타와 간접적인 유형의 데이타를 질의 조건 값으로 받아들여 질의를 수행하는 간접 질의(Indirect Query) 방식으로 구분될 수 있다. 간접 질의를 수행하는 방법에는 주석을 직접 부여하여 처리하는 방법과 추출된 특징정보들의 값을 이용하여 처리하는 방법이 있다. 본 논문에서 제안하는 영상 검색 시스템은 메타데이타를 이용하여 간접 질의 방식으로 질의를 할 수 있다. 또한 사용자 인터페이스에 의해 사용자가 편리하고 쉽게 영상 데이타를 질의할 수 있도록 하는 질의 검색 기능과 영상 데이타를 스토리보드 별로 원하는 데이타를 검색할 수 있도록 하는 영상 브라우저 검색 기능을 이용할 수 있다.

4. 실험 및 결과

본 논문은 Windows 98 환경에서 Microsoft Visual C++ 6.0, Microsoft Access 2000 DBMS, Microsoft Visual InterDev, Firebird 1.12를 사용하여 구현하였다. 본 실험에 사용한 영상데이타는 Firebird 1.12를 사용하여 일반용 캠코더로 촬영한 비디오 컨텐츠 데이타를 스냅 형태로 캡쳐한 영상 데이타를 사용하였다.

4.1 구현한 영상 검색 시스템의 메인 구조

본 논문에서 구현한 영상 검색 시스템의 메인 구조는 그림 4와 같다.

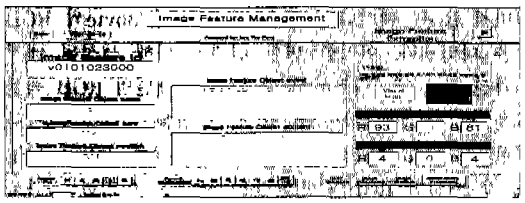


그림 4 영상 검색 시스템의 메인 구조

그림 4에서와 같이 영상 검색 시스템은 Image Feature Extraction 기능, Image Processing 기능, Image Browser 기능, Image Query Processing 기능, Image Metadata Management 기능의 5 가지 기능으로 구현하였다. 그림 4에서 "Image Color Feature Extraction Value", "Image Edge Feature Extraction Value" 항목은 내용기반 에지 특징정보 즉, 영상의 하위레벨에 표현된 객체의 에지 특징정보를 추출한 값이며, "Image Feature Id", "Image Feature Object Name", "Image Feature Object Type", "Image Feature Object Position", "Image Feature Object Event", "Image Feature Object Content" 항목은 영상의 특징정보 추출이 어려운 상위레벨에 표현된 주석기반 에지 특징정보를 추출하여 입력한 값이다. 그리고 "Visual Image" 항목은 영상 검색 시스템의 원시 영상이다.

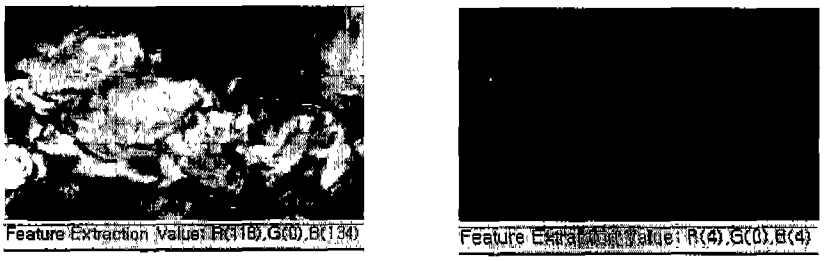
4.2 영상의 에지 특징정보추출

본 논문에서 사용한 실험 영상 데이타는 영상 해상도가 720×480인 24비트 True Color Mode를 사용하였다.

그림 5는 본 논문에서 구현하여 실험한 내용기반 에지 특징정보 추출 결과 영상 및 값을 나타낸 예이다. 그림 5(a)는 초기 입력 영상에 대하여 각 채널별로 RGB 평균값을 내용기반 특징정보로 추출한 결과이며, 그림 5(b)는 그림 5(a)의 입력 영상에 대한 에지를 추출하여 각 채널별로 RGB의 평균값을 내용기반 에지 특징정보로 추출한 결과이다.

4.3 영상 검색 시스템의 사용자 인터페이스

본 논문에서 제안하는 영상 검색 시스템의 사용자 인터페이스는 영상 처리를 위한 Image Processing 기능과 영상의 특정 부분을 내용 브라우징(Content Browsing)하기 위한 Image Browser 기능, 그리고 주석기반 질의, 내용기반 질의, 주석기반 및 내용기반 통합질의와 같은 다양한 검색방법을 제공하여 상호 보완적인 역할을 갖도록 하기 위한 Image Query



(a) 초기 입력 영상의 특징정보추출결과: ⇒ R(118), G(0), B(134)

(b) 입력 영상의 에지 특징정보추출결과: ⇒ R(4), G(0) , B(4)

그림 5 영상의 에지 특징정보 추출 결과

Processing 기능으로 구분하여 구현하였다. 본 절에서는 영상 검색 시스템의 사용자 인터페이스 중 영상 검색을 지원하기 위한 방법으로 구현된 영상 브라우저와 영상 질의 프로세싱 기능에 관하여 설명한다.

4.3.1 영상 브라우저

영상 브라우저 기능의 목적은 영상 검색 시스템의 전체 영상 컨텐츠의 내용을 사용자가 빠르고 정확하게 검색할 수 있도록 하는 기능이다. 본 논문에서 제시하는 영상 브라우저는 영상 데이타를 스토리보드별로 브라우저에서 원하는 데이타를 자체적으로 검색할 수 있도록 하는 기능과 영상의 특정 부분을 내용 브라우징 할 수 있도록 하는 기능을 함께 제공하여 구현하였다. 본 시스템의 영상 브라우저 기능은 각 영상에 대한 내용을 쉽게 파악할 수 있도록 영상의 객체, 객체 타입, 객체 위치, 객체 이벤트, 객체 내용 등의 에지 특징정보를 내용 브라우징하는 기능이다. 그리고 영상 브라우저에는 각자 별도의 자체검색 기능을 제공하고 있다. 그림 6은 본 논문에서 구현한 영상 검색 시스템의 영상 브라우저 기능 중 Image Object의 "꽃"을 자체검색 브라우징하여 실험한 결과 예이다.

그림 6에서 "Image Color Chanel", "Image Edge Color Chanel" 항목은 내용기반 에지 특징정보를 검색한 브라우저 결과 값이며, "Image ID", "Image Object Name", "Image Object Type", "Image Object Position", "Object Event", "Object Content" 항목은 주석기반 에지 특징정보를 검색한 브라우저 결과 값이다. 그리고 "Visual Image" 항목은 검색한 브라우저의 결과 영상이다.

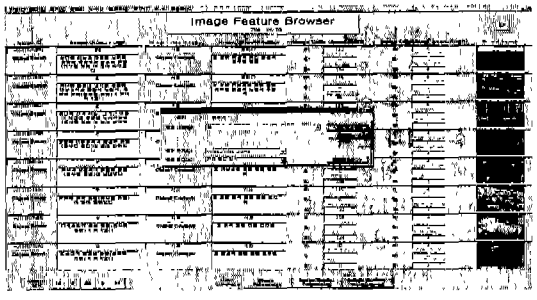


그림 6 영상 브라우징 예

4.3.2 영상 질의 프로세싱

본 논문에서 구현한 영상 검색 시스템의 영상 질의 프로세싱 기능은 영상 데이타베이스에 저장된 메타데이타와 간접적인 유형의 메타데이타를 질의 조건 값으로 받아들여 질의를 수행하는 간접 질의 방식을 이용하였다. 영상 질의 처리방식은 영상 검색 시스템의 메타데이타와 영상의 상호관계를 기반으로 사용자가 영상 질의 프로세싱 인터페이스 화면을 통하여 쉽게 질의할 수 있도록 구현하였다. 영상 질의 방법은 검색방법들간의 상호 보완적 역할을 갖도록 하여 신속하고 효율적으로 원하는 영상을 질의할 수 있도록 주석기반질의, 내용기반질의, 그리고 주석기반 및 내용기반 통합질의 방식의 세 가지 질의방식을 제공하였다. 그리고, 사용자가 질의어에 대한 지식이 없더라도 쉽고 정확하게 영상 데이타를 질의하고 검색할 수 있도록 다양한 질의 연산 방법을 제공하여 질의 및 검색에 편리성을 제공하였다. 사용자 인터페이스에 제공하는 영상 질의 연산방법은 표1과 같다.

본 절에서는 본 논문에서 구현한 영상 질의 프로세싱의 주석기반질의, 내용기반질의, 그리고 주석기반 및 내용기반 통합질의의 세 가지 영상 질의 방법을 표 1에서 제시한 질의 연산방법으로 실험하였다. 그리고 이 세 가지 영상 질의 방법을 실험한 영상검색 결과를 통하여 검색방법들간의 상호 보완적인 역할을 보여준다. 그림 7은 표 1에서 제시하는 질의 연산방법을 이용하여 주석기반질의, 내용기반질의, 그리고 주석기반 및 내용기반 통합질의 방식으로 질의한 내용을 보여준 예이다.

표 1 사용자 인터페이스의 영상 질의 연산방법

질의 방법	표현식	연산방법
주석기반질의	[객체명]	단일항목연산
	[객체명] OR [객체타입 | 객체위치 | 객체이벤트 | 객체내용]	OR 연산
	[객체명] AND [객체타입 | 객체위치 | 객체이벤트 | 객체내용]	AND 연산
내용기반질의	[기본색상특징정보 | 에지특징정보]	단일항목연산
통합질의 (주석기반 및 내용기반)	[기본색상특징정보 | 에지특징정보] AND [객체명]	단일항목연산
	기본색상특징정보 | 에지특징정보] OR [객체명] OR [객체타입 | 객체위치 | 객체이벤트 | 객체내용]	OR 연산
	기본색상특징정보 | 에지특징정보] AND [객체명] AND [객체타입 | 객체위치 | 객체이벤트 | 객체내용]	AND 연산

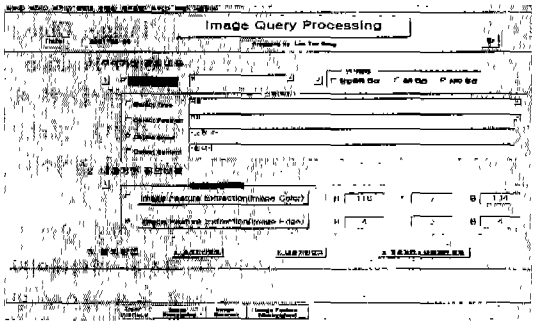
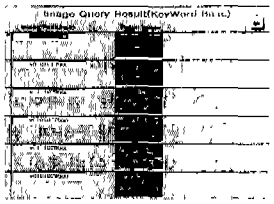


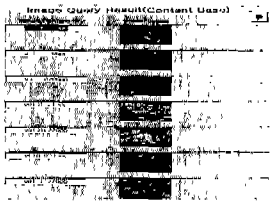
그림 7 영상 질의 내용 예

그림 8은 그림 7에서 영상의 주석기반 에지 특징정보 추출 값이 노랑꽃인 영상을 주석기반 질의방식으로 질의하여 실험한 영상 검색 결과를 보여준 예이다.



(a) 영상 질의 방법: 주석기반 질의,
(b) 질의 연산 방법: AND 연산
(c) 표현식: [객체명] AND [객체타입 | 객체위치 | 객체이벤트 | 객체내용]
(d) 주석기반 에지 특징정보: [Object: "꽃", Object Content: "*노랑색*"]

그림 8 주석기반 질의 검색 결과 예

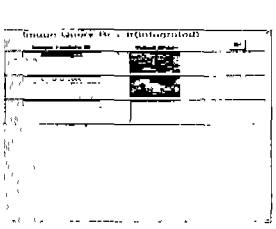


(a) 영상 질의 방법: 내용기반 질의,
(b) 질의 연산 방법: 단일항목연산
(c) 표현식: [기본색상특징정보 | 에지특징정보]
(d) 내용기반 에지 특징정보: [R:4, G:0, B:4]

그림 9 내용기반 질의 검색 결과 예

그리고, 그림 9는 그림 7에서 영상의 내용기반 에지 특징정보 추출 값이 R:4, G:0, B:4 인 영상을 내용기반 질의방식으로 질의하여 실험한 영상 검색 결과를 보여준 예이다.

그리고, 그림 10은 그림 7에서 영상의 내용기반 에지 특징정보 추출 값이 R:4, G:0, B:4 이고 주석기반 에지 특징정보 추출 값이 흰 꽃 영상을 주석기반 및 내용기반 통합질의방식으로 질의하여 실험한 영상 검색 결과를 보여준 예이다.



(a) 영상 질의 방법: 주석기반 및 내용기반 통합질의,
(b) 질의 연산 방법: AND 연산
(c) 표현식: [기본색상특징정보 | 에지특징정보] AND [객체명] AND [객체타입 | 객체위치 | 객체이벤트 | 객체내용]
(d) 내용기반 에지 특징정보: [R:4, G:0, B:4)]
(e) 주석기반 에지 특징정보: [Object : "꽃", Object Content : "*흰색*"]

그림 10 주석기반 및 내용기반 통합질의 검색 결과 예

위 실험에서와 같이 영상의 하위레벨에 표현된 내용기반 에지 특징정보와 상위레벨에 표현된 주석기반 에지 특징정보를 이용하여 다양한 검색방법으로 질의함으로서 검색방법들간의 상호 보완적인 역할에 의해 신속하고 효율적으로 원하는 영상을 정확하게 검색할 수 있다는 것을 알 수 있다.

4.4 실험결과 분석

본 절에서는 영상 검색 시스템의 검색 적합성(Retrieval Effectiveness)[15]에 대한 실험 결과를 설명한다. 영상 검색 시스템의 완전 일치 검색을 기반으로 하는 검색의 적합성은 재현율(Recall)과 정확율(Precision)로 검색의 성능을 측정하고 있다. 반면에 완전 일치 검색이 아닌

표 2 질의 영상에 대한 적합성 실험 결과

영상구분	검출대상영상수	검출영상적합수	검출영상부적합수	검출영상전체수	Recall(%)	Precision(%)
자막영상	13	13	0	13	26	100
구름영상	5	5	0	5	10	100
꽃(튜립) 영상	14	14	1	15	28	93.3
노을영상	1	1	0	1	2	100
동물(말) 영상	1	1	0	1	2	100
불꽃영상	10	10	1	11	20	90.9
야생화영상	2	2	1	3	4	66.7
컬러바영상	1	1	0	1	2	100
태양영상	3	3	1	4	6	75
계	50	50	4	54	11.1	91.77

표 3 기존 영상 검색 시스템들과의 비교

시스템 / 항목	QBIC[4,5]	QVE[6]	Jacobs[7]	VisualSEEK[8]	제안 시스템
질의기능	- 제한적 주석기반질의 - 내용기반질의	- 내용기반질의	- 내용기반질의	- 내용기반질의	- 주석기반질의 - 내용기반질의 - 주석기반 및 내용기반 통합질의
질의형태	- 시각적인 이미지 특징정보를 통한 시각적인 내용기반질의형태(직접질의)	- 시각적 질의툴 이용한 내용기반질의형태(예시질의)	- 특징벡터 값 추출에 의한 내용기반질의형태(직접질의)	- 시각적 특징정보 및 공간관계 특징정보 추출에 의한 내용기반질의형태	- Object, Object Type, Object Position, Object Event, Object Content, 에지 특징정보 등의 추출에 의한 주석기반 및 내용기반 질의형태(간접질의)
질의기반	- Image Annotation - Image Feature	- Image Feature	- Image Feature	- Image Feature	- Image Annotation Metadata - Image Feature Metadata
질의방법	- 예제영상을 통해 예제영상과 유사성이 높은 영상들을 브라우징 - 시스템에서 제공하는 시각적인 구성 그림도구를 기반으로 사용자가 직접 그린 구도정보를 통한 질의	- 화면에 나타나는 예제에 의한 시각적 질의	- 특징정보추출 과정에서 색상영상에 직접 웨이브릿 변환을 적용하여 추출한 계수를 특징벡터로 사용하여 질의	- 웹상에서 사용자 인터페이스기반의 분류메뉴를 통해 주제별로 질의	- Image Browser와 Image Query Processing 화면 등의 사용자 질의 인터페이스 기반과 사용자 브라우징 인터페이스 기반에 의한 질의
특징	- 영상 및 비디오 데이타베이스를 기반으로 키워드, 색상, 모양, 질감, 스케치 등의 시각적인 영상특징정보를 통한 다양한 형태의 시각적 복합질의와 색상 구성에 의한 질의기능 - 시각적인 구성 그림도구를 기반으로 사용자가 직접그린 구도정보를 통한 질의기능 - 예제영상질의를 통해 예제영상과 유사성이 높은 영상 브라우징 기능 - 질의 인터페이스가 복잡하며, 질의 인터페이스가 부분적으로 통합되지 않아 이용하기가 용이하지 않음 - 주석기반 영상선의가 의미정보에 의해 제한적임	- 사용자의 다양한 질의로부터 색상 히스토그램, 질감분석 등의 형태속성을 얻거나 이들 속성들은 합성하는 방법에 의한 시각적 질의기능 - 영상으로부터 추출한 외곽선 데이타를 그대로 특징 데이타로 사용하므로 질의처리 및 검색과정에서 유사성 검사 시 데이타베이스에 저장된 각각의 영상에 대한 기하학적인 변환, 즉 이동, 스케일링, 회전 등을 일일이 고려해야 하는 단점이 있음	- 영상특징정보 추출과정에서 컬러 영상에 직접 웨이브릿 변환을 적용하여 추출한 계수(Coefficient)들을 특징벡터로 사용하여 컬러 영상의 각 RGB 채널마다 따로 특징 벡터를 구성하므로 각 특징 벡터의 차원이 너무 커서 영상 색인으로 부적합하다는 단점이 있음	- 웹기반 영상정보검색 사용자 인터페이스 지원 - 웹기반 초기 사용자 인터페이스 화면의 분류메뉴에 의한 주제별 질의기능 - 영상개체들의 시각적 특징정보와 공간적 특징정보를 이용하여 동시에 내용기반질의와 공간질의가 가능하다는 장점이 있음 - 각 색상영역간의 공간적인 관계를 추론함으로써 추론과정에 많은 시간이 소요될 수 있다는 단점이 있음	- 주석기반질의, 내용기반질의, 주석 및 내용기반통합질의 등의 다양한 질의기능 - 주석기반 및 내용기반 에지 특징정보의 메타데이타에 의한 영상 색인 구성이 용이 - 에지 외곽선 데이타에 대한 시각적인 특징정보에 의한 주석기반 특징정보 추출이 용이 - 영상 질의 및 브라우저 검색기능의 다양한 질의연산방법에 의한 다양한 형태의 복합질의기능 - 사용자 질의 인터페이스 기반과 사용자 브라우징 인터페이스 기반에 의한 질의기능 - 주석기반 에지 특징정보와 내용기반 에지 특징정보의 메타데이타에 의한 정확한 영상 컨텐츠 정보의 축적관리와 웹 컨텐츠 제작시 정보공유 및 재이용 가능

유사 검색(Similarity Retrieval) 시스템은 표준 정확율과 재현율(Normalized Precision and Recall)을 이용하여 평가한다[16]. 본 논문에서 구현한 영상 검색 시스템은 완전 일치 검색을 기반으로 하는 영상 검색 시스템이므로 재현율과 정확율을 이용하여 실험결과에 대한 검색 적합성을 설명한다. 어떤 질의에 대하여, 질의에 적합한 전체 영상 수를 T, 적합하게 검색된 영상 수를 Rr, 검색된 전체 영상 수를 Tr이라 하면, 정확율은 Rr/Tr로 정의되며, 재현율은 Rr/T로 정의된다[16]. 본 영상 검색 시스템의 실험은 총 50개의 영상 데이타를 영상 종류별로 구분하여 각 질의 영상에 대한 검색결과를 실험해 보았다. 표 2는 질의 영상에 대한 적합성 실험 결과를 보여주고 있다.

표 2와 같이 영상 데이타베이스에 저장하여 사용하는 영상 데이타에 따라 재현율과 정확율은 다소 차이가 있겠으나 본 실험에서는 대체로 높은 정확율을 얻을 수가 있었다.

4.5 기존 영상 검색 시스템들과의 비교 분석

기존의 개발된 영상 검색 시스템의 사례로는 IBM Almaden 연구소에서 개발한 주석 및 내용기반 영상 검색 시스템인 QBIC(Query by Image and Video Content)[4,5]과 Hirata와 Kato가 만든 내용기반 영상

검색 시스템인 QVE(Query by Visual Example)[6]가 있다. 그리고 Jacobs[7] 등이 제안한 내용기반 영상 검색 시스템과 Columbia 대학에서 개발한 내용기반 영상 검색 시스템인 VisualSEEK[8] 등이 있다. 본 절에서는 본 논문에서 제시하여 구현한 영상 검색 시스템에 대한 특징과 영상질의처리, 그리고 영상브라우징 기능 등을 기존의 개발된 영상 검색 시스템들과 비교 분석하였다.

QBIC은 입력 영상의 전체 픽셀에 평균 RGB 방법을 이용하여 고차원 컬러 히스토그램을 3차원 컬러 히스토그램으로 변환한 색상 특징정보를 이용하고 있다. 이 시스템은 현재 영상이나 비디오를 내용기반으로 검색할 수 있도록 지원하는 상용화된 영상 검색 시스템이다. QVE는 사용자의 다양한 질의로부터 색상 히스토그램, 질감분석 등의 형태속성을 얻거나 이들 속성들을 합성하는 방법으로 시각적 질의를 제공하는 내용기반 영상 검색 시스템이다. 그리고 Jacobs 등이 제안한 영상 검색 시스템은 영상의 특징정보 추출과정에서 최초로 웨이브릿 변환을 사용한 내용기반 영상 검색 시스템이며, VisualSEEK는 웹을 이용한 내용기반 영상 검색 시스템이다.

본 논문에서 제안하는 영상 검색 시스템은 영상의 하위레벨에 표현된 내용기반 에지 특징정보의 메타데이타 유형과 영상의 상위레벨에 표현된 주석기반 에지 특징정보의 메타데이타 유형을 영상의 색인으로 구성하는 주석기반 및 내용기반 영상 검색 시스템을 제시하여 구현하였다. 기존의 개발된 영상 검색 시스템들과 본 논문에서 제안하여 구현한 영상 검색 시스템을 비교 분석한 내용을 표 3에 요약하였다.

5. 결론 및 향후 연구과제

본 논문에서는 영상의 에지 특징정보를 추출하는 방법과 주석기반 에지 특징정보와 내용기반 에지 특징정보를 메타데이타로 제공하여 영상 데이타베이스를 생성하는 기법을 제시하였다. 그리고 이를 위해 영상의 에지 특징정보를 이용한 주석기반 및 내용기반 영상 검색 시스템을 구현하였으며 그 성능을 분석하였다. 영상내의 특징기반 영상 검색방법은 단일 특징만을 이용할 경우 정확한 검색이 이루어질 수 없으며 대용량의 데이타베이스에 적용하면 검색 효율이 떨어질 수 있다. 따라서 본 논문에서는 영상의 하위레벨에 표현된 내용기반 에지 특징정보와 영상의 특징정보 추출이 어려운 상위레벨에 표현된 주석기반 에지 특징정보를 영상의 색인으로 구성하여 다양한 검색방법을 제공하였다. 본 논문에서 제안하는 영상 검색 시스템은 영상의 하위레벨에 표현된 내용기반 에지 특징정보와 상위레벨에 표현된 주석기반 에지 특징정보를 이용하여 다양한 검색방법을 제공함으로서 신속하고 효율적으로 정확하게 원하는 영상을 검색할 수 있어 영상 컨텐츠 검색의 성능을 향상시킬 수 있다. 그리고, 사용자가 질의어에 대한 지식이 없더라도 쉽고 정확하게 영상 데이타를 질의하고 검색할 수 있도록 질의 기능과 브라우저 기능의 사용자 인터페이스를 제공하여 질의 및 검색에 편리성을 제공하였다. 본 논문에서는 영상 검색 시스템의 메타데이타 모델을 제시하여 정확한 영상 컨텐츠 정보의 축적관리와 구축작업의 효율화가 가능하도록 하였으며, 또한 웹 컨텐츠 제작 시 정보공유 및 재이용이 가능하도록 하여 웹 컨텐츠의 다양한 응용서비스에 효율성을 높일 수 있도록 하였다. 그리고 본 논문에서 구현한 영상 검색 시스템의 질의 기능과 브라우저 기능의 특징은 다음과 같다. 첫째, 영상의 주석기반 에지 특징정보와 내용기반 에지 특징정보의 메타데이타를 영상의 색인으로 구성하여 다양한 검색방법을 제공함으로서 사용자 질의를 매우 다양하게 지원할 수 있다. 둘째, 주석이나 영상 특징에 의해 내용 브라우징 기능을 지원하여 사용자가 원하는 정보에 보다 정확하고 효율적으로 접근할 수 있다. 셋째, 사용자가 질의어에 대한 지식이 없더라도 쉽고 정확하게 영상 데이타를 질의하고 검색할 수 있도록 질의 기능과 브라우저 기능에 다양한 질의 연산방법을 제공하여 질의 및 검색에 편리성을 제공하였다. 넷째, 질의 및 브라우저 인터페이스에 자체 검색 기능과 브라우저 기능을 제공하여 질의 및 브라우저에 편리성을 제공하였다.

향후 연구과제로는 본 논문에서 제안한 영상 검색 시스템을 다양한 멀티미디어 컨텐츠에 적용될 수 있도록 확장하는 것이다. 그리고 좀 더 효율적으로 영상의 특징정보를 추출하고, 영상의 다양한 특징정보를 신속하고 정확하게 추출하여 영상 검색 시스템을 구현할 수 있도록 하는 것이 연구과제이다.

참 고 문 헌

[1] Yossi Rubner, Carlo Tomasi and Leonidas J. Guibas, "Adaptive Color-Image Embeddings for Database Navigation," ACCV'98 Computer Vision-Volume I, pp. 104-119, Jan. 1998.

[2] Yong Rui, Thomas S. Huang, and Shih-Fu Chang, "Image Retrieval: Current Techniques, Promising Direction and Open Issues," Journal of Visual Communication and Image Representation Vol. 10, 39-62, March, 1999.

[3] A. Pentland, R. W. Picard, and S. Sclaroff,

"Photobook: Content-Based Manipulation of Image Database," International Journal of Computer Vision, fall 1995.
[4] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Pekovic, P. Yanker, C. Faloutsos and G. Taubin, "The QBIC project: Querying images by Content using color, texture, and shape," Proc. of SPIE Storage and Retrieval for Image and Video Databases, pp.173-187, February 1993.
[5] Myron Flickner and et. al, "Query by Image and Video Content: The QBIC system," IEEE Computer, 28(9), pp.23-32, 1995.
[6] K. Hirata and T. Kato, "Query by visual example: Content based image retrieval," Advances in Database Technology, pp.56-61, 1992.
[7] C. E. Jacobs, A. Finkelstein, and D. H. Salesin, "Fast Multiresolution Image Query," Proc. ACM SIGGRAPH, New York, 1995.
[8] J. R. Smith and S.-F. Chang, "VisualSEEK: A Fully Automated Content-based Image Query System," ACM Multimedia 96, Boston, MA, 1996.
[9] M. Sonka, V. Hlavac, and R. Boyle, "Image Processing Analysis, and Machine Vision," Brooks/Cole Publishing, Second Edition, 1999.
[10] Randy Crane, "A Simplified approach to Image Processing," Prentice-Hall, 1997.
[11] V. N. Gudivada and V. V. Raghavan, "Content-Based Image Retrieval Systems," IEEE Computer, 28(9), 1995.
[12] V. V. Vinod, H. Murase and C. Hashizume, "Focussed Color Intersection with Efficient Searching for Object Detection and Image Retrieval," Proceedings of the IEEE International Conference on Multimedia Computing and System, pp.229-233, 1996.
[13] 홍성용,나연묵 "혼합형 이미지 메타데이타를 이용한 지능적 이미지 검색 시스템 설계 및 구현", 멀티미디어학회논문지, 3권 3호, 한국멀티미디어학회, pp.209-223, Jun. 2000.
[14] 한정운, 김병곤, 이재호, 임해철, "이미지 내용 기반 검색을 위한 이미지 타일 평균 RGB 방법", 한국정보과학회 제26회 추계학술발표회, 1999.
[15] D. Shasha and T-L. Wang, "New techniques for best-match retrieval," ACM TOIS, 8(2) pp.140-158, April 1990.
[16] G. Salton and M. J. McGill, "Introduction to Modern Information Retrieval," McGraw-Hill, 1983.

이 태 동
1987년 서울산업대학교 전자계산학과(공학사). 1990년 광운대학교 전자계산학과(이학석사). 1995년 ~ 현재 아주대학교 컴퓨터공학과(박사과정). 1992년 ~ 1997년 영동전문대학 전자계산과 조교수. 1998년 ~ 현재 컴퓨터응용과 조교수. 관심분야는 지능형 정보검색 시스템, 웹 멀티미디어 컨텐츠 저작도구, 이미지 및 비디오 정보검색, 인터넷 방송, 웹 에이전트, 멀티미디어 및 웹 DB

김 민 구
1977년 서울대학교 계산통계학과(이학사). 1979년 한국과학기술원 전산학과(공학석사). 1989년 Pennsylvania 주립대(박사). 1999년 ~ 2000년 Louisiana 대학 연구과학자. 1981년 ~ 현재 아주대학교 컴퓨터공학과(교수). 관심분야는 지능형 정보검색 시스템, 지능형 교수 시스템, 지능형 캐릭터 에이전트