

스케치 질의를 통한 웹기반 영상 검색과 분류 시스템

(Web-based Image Retrieval and Classification System using Sketch Query)

이 상 봉[†] 고 병 철[†] 변 해 란^{**}

(Sang-Bong Lee) (ByoungChul Ko) (Hye-Ran Byun)

요 약 디지털 기술의 발달과 인터넷의 대중화에 더불어 영상데이터의 생산과 교환이 더 자유로워짐에 따라 디지털 도서관, 영상처리, 데이터베이스 시스템과 같은 연구분야에서 내용기반 영상검색에 대한 관심이 높아지고 있다. 일반적으로 '영상에 의한 질의'의 경우 사용자가 마음에 드는 영상이 없더라도 반드시 질의 영상을 데이터베이스로부터 선택해야 하지만, '스케치에 의한 질의'는 사용자의 생각에 따라 영상을 그림으로 표현할 수 있으므로 최근에 가장 많이 사용되는 질의 방법 중 하나이다. 본 논문에서는 스케치 질의와 영상 분류 방법을 이용하는 자바 기반의 영상검색 시스템을 제안한다. 본 시스템에서는 유사 영상을 검색하기 위해 영상으로부터 색상 히스토그램과 Haar-웨이플릿 계수를 사용하고, leave-one-out 방법을 이용하여 영상을 분류하도록 하였다. 본 논문에서는 사진·그림, 자연·도시 등의 영상 분류를 통해 영상의 의미정보를 추출할 수 있을 뿐 아니라, 사용자 질의 영상을 분류하여, 질의 영상이 갖고 있는 의미 공간으로 검색 공간을 축소하여 검색 시간을 단축시키는 효율성을 얻을 수 있었다.

키워드 : 내용기반 영상 검색, 스케치 질의, 영상 분류

Abstract With the explosive growth in the numbers and sizes of imaging technologies, Content Based Image Retrieval (CBIR) has been attracted the interests of researchers in the fields of digital libraries, image processing, and database systems.

In general, in the case of query by image, the user has to select an image from database to query, even though it is not his completely desired one. However, since query-by-sketch approach draws a query shape according to the user's desire, it can provide more high-level searching interface to the user compared to the query by image. As a result, query-by sketch has been widely used.

In this paper, we propose a Java-based image retrieval system that consists of sketch query and image classification. We use two features such as color histogram and Haar wavelets coefficients to search similar images. Then the Leave-One Out method is used to classify database images. The categories of classification are photo & painting, city & nature, and sub-classification of nature image. By using the sketch query and image classification, we can offer convenient image retrieval interface to user and we can also reduce the searching time.

Key words : Content-based Image Retrieval, Query-by-sketch, Image Classification

1. 서 론

멀티미디어 기술의 발달로 세계 여러 곳에서 디지털 형태의 시청각(audiovisual) 정보의 생산과 소비가 증가하고 있다. 디지털 기기의 발달로 영상의 생산 속도와 양이 급속히 증가하고 있어, 우리가 원하는 정보를 빠르고, 효율적이고, 편하게 검색하는 것이 중요한 문제로 대두되고 있다. 이러한 문제를 인식하여 이들을 효과적으로 검색할 수 있는 내용기반 영상 검색에 관련된 연구들이 진행되어 왔다[1]. 현재, 대부분의 영상검색 시스

· 본 논문은 학술진흥재단의 지원(2002-005-H20001)으로 연구되었음

† 비 회 원 : 연세대학교 컴퓨터과학과

soopi@nownuri.net

soccer1@cs.yonsei.ac.kr

** 종신회원 : 연세대학교 컴퓨터학과 교수

hrbyun@cs.yonsei.ac.kr

논문접수 : 2001년 8월 27일

심사완료 : 2003년 5월 2일

템이 데이터베이스에 포함되어 있는 예제 영상들로 검색을 시작하는데, 이러한 시스템의 경우 예제 영상에서 사용자가 원하는 영상이 나오지 않았을 때에는 원하는 예제가 나올 때까지 검색을 해야하는 문제점을 지니고 있다. 이러한 방식에 대한 대안으로 사용자 스케치에 기반을 둔 검색 방법과 영상 분류를 통해 얻어진 의미 클래스에서 각 클래스의 영상들을 예제로 제시함으로써, 검색의 용이성과 효율성을 획득할 수 있다. 그리고, 영상 분류를 함으로써 영상의 의미를 얻을 수 있으며, 질의 영상의 검색 공간을 해당 의미 클래스로 줄여 검색에 소요되는 시간 또한 줄일 수 있다[2].

본 논문에서는 영상의 특징추출을 통해 영상 분류를 시행하며, 이를 통해 검색 효율을 높이며, 사용자의 검색 용이성과 편이성을 위한 스케치 질의 검색 인터페이스와 그 결과로 선택된 영상들을 예제로 하여 영상을 검색하는 웹 기반 영상 검색시스템을 구축 하였다.

제안한 시스템은 사용자의 접근이 용이한 웹 브라우저를 통해 수행될 수 있도록 Java 로 설계되었고, 컬러 히스토그램, Haar 웨이블릿 변환을 통해 얻어진 계수들, 영상이 갖는 채도(saturation), 명도(intensity), 컬러영역이 갖는 상관관계를 색인함으로써, 다양한 특징을 이용한 영상 검색이 가능하도록 시스템을 설계하였다. 시스템을 설계함에 있어서, 스케치질의에 대해 Haar 웨이블릿 변환을 통해 얻어진 계수와 컬러 히스토그램을 함께 사용하여, 컬러히스토그램과 [4]에 제시된 방법을 HSV 컬러 공간에 적용했을 때 보다 좋은 성능을 보여 주었으며, [10]에서 제시한 영역 관계 추출 방법을 수정함으로써, 큰 컬러 영역이 적은 빈도수로 나타나는 영상도 상대적 영역질에서 검색될 확률을 높였다.

본 논문은 다음과 같다. 2장에서는 관련 연구들을 알아보고, 3장에서는 시스템의 구조에 대해 살펴본다. 4장과 5장에서는 영상 분류와 검색을 위한 특징 추출에 대해 설명한다. 6장에서는 실험 결과를 보이고, 결과를 통해 영상 분류와 스케치 질의의 효율성에 대한 평가 및 분석을 한다. 마지막으로 7장에서는 결론 및 향후 연구 계획에 대하여 서술한다.

2. 관련 연구

내용기반 영상검색은 영상이 갖고 있는 물리적, 논리적인 의미를 추출, 색인을 하여 방대하게 존재하는 영상 데이터를 효과적이고 빠르게 검색하는 방법으로[1][3], 본 논문에서 수행할 연구의 초점인 스케치 질의, 영상이 갖는 상위 단계의 의미정보를 추출, 영상검색의 효율을 높이기 위한 영상분류에 관련된 연구들을 좀 더 자세히

살펴보겠다.

Jacobs가 제안한 시스템[4][5]의 경우 영상의 다해상도 분석이 가능한 웨이블릿 변환 중 쉽고, 계산량이 적은 Haar 웨이블릿을 사용하여 절대값이 큰 계수의 위치를 색인하여 검색한다. 컬러 컴포넌트들의 전체 평균값을 이용함으로써 컬러 정보를 충분히 반영하지 못한다. Wang이 제안한 시스템의 경우[6] Daubechies 웨이블릿 변환을 사용하여 2단계 검색을 한다. 2단계 검색을 통해 검색의 정확성을 높였지만, 이 방법에서 사용한 Daubechies 웨이블릿의 경우 Haar 웨이블릿보다 시간이 오래 걸리고, 영상 색인에 많은 정보를 사용하고 있다. Smith가 제안한 VisualSEEK [7]의 경우 영상에서 HSV 컬러 공간에서 두드러지게 나타나는 컬러 영역을 추출하여 각 영역의 중심점, 크기를 색인하여, 컬러영역의 위치 비교를 통해 영상을 검색한다. [표 1]은 스케치를 기반으로 하는 시스템과 본 논문에서 설계한 시스템의 특징을 요약하고 있다.

표 1 스케치 기반 영상 검색 시스템

시스템	특징 추출	특징 비교법	비교
Jacob s [4]	·컬러공간 : YIQ ·Haar 변환 이용, 계수 추출 ·전체 평균, 계수 위치	·평균과 계수의 이용	·빠른 검색 ·컬러 정보의 반영 미흡
Wang [6]	·컬러공간 : 보색공간 ·Daubechies 웨이블릿 변환 ·저주파 부분의 표준편차 ·고주파 부분의 계수	·2단계 비교 - 저주파 영역 - 고주파 영역	·2단계 검색으로 검색의 효율, 정확성 높임 ·웹과 연계 ·많은 수의 계수 색인
Smith [10]	·컬러공간 : HSV ·컬러 영역의 상대적인 위치	·컬러 쌍의 빈도수	·쉬운 인터페이스 ·검색의 용이성 ·웹과 연계
제안된 시스템	·컬러공간 : HSV ·컬러 히스토그램 ·Haar 변환 이용, 계수 추출 ·컬러 영역의 위치 ·직선 성분 히스토그램	-히스토그램 Haar 계수 매칭	·웹과 연계 ·분류를 통한 검색의 효율성 높임 ·다양한 검색 방법 제공 ·체계적인 인덱스 파일구성

영상분류에 관련된 연구는 비교적 최근에 많이 연구되고 있는 분야이다. Huang[2]은 컬러 correlogram을 구하여 SVD(Singular Value Decomposition)를 통해 계층적으로 분류를, Szummer[8]는 K-NN, 블록 단위의 비교를 통해 실내의 영상 분류에 관한 연구를, Vailaya

[9]는 Canny 윤곽선 추출방법으로 윤곽선을 추출하여 도시 자연 영상 분류에 관한 연구를 수행했다. 또 Smith[10]는 영상에서 나타나는 두드러진 영역을 추출하여 각 영역의 상하관계를 색인하여 분류하였다. 하지만 이 방법은 색상의 상하 관계만으로 분류할 수 있는 분류범위가 제한적이라는 단점을 가진다.

본 논문에서 제안된 시스템은 위에서 언급된 스케치 질의 검색 시스템 및 영상 분류방법을 바탕으로 효과적인 영상 검색을 수행하기 위한 다양한 검색 방법을 제공하고, 웹을 통한 검색의 용이성, 영상의 분류를 통한 검색의 효율성을 고려한 영상검색 시스템을 설계하였다. 그리고, 영상검색, 분류를 위해 여러 가지 특징들을 추출하기 때문에 이 특징들을 효율적으로 사용하기 위해 XML을 이용하여 특징 인덱스 파일을 구성하였다.

3. 시스템 구조

검색시스템은 [그림 1]과 같이 구성되어 있고, Java 기반의 시스템으로 영상 데이터 베이스는 서버에 구성돼 있으며, 클라이언트에서는 영상 검색 인터페이스를 제공한다.

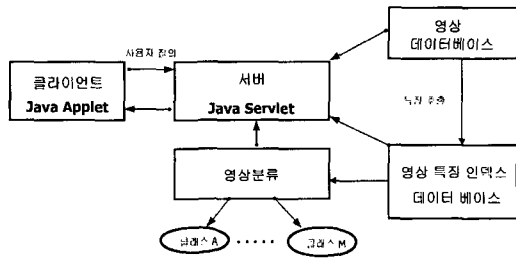


그림 1 전체 시스템 개요

그림 1에서 보는 것과 같이 영상 데이터베이스로부터 off-line으로 특징 벡터를 추출하고 이 특징 벡터를 인덱스로 데이터베이스에 다시 저장한다. 그런 뒤에 영상특징 인덱스를 이용하여 영상 데이터 베이스는 시스템에서 정의한 각 클래스로 분류된다. On-line상에서 사용자가 스케치 질의를 할 경우, 서버에서는 마찬가지로 특징 값을 데이터베이스로부터 추출하여 어떤 클래스에 속하는지 결정하고 전체 데이터베이스를 검색하는 것이 아니라 해당 클래스에서만 검색하도록 한다. 이렇게 함으로써 보다 정확한 검색이 이루어질 수 있고, 검색 시간도 단축할 수 있다. 그림 2는 사용자로부터 입력된 스케치 질의 영상을 데이터베이스와 비교하여 가장 유사한 영상을 우선 순위대로 보여주는 검색 과정을 보여준다.

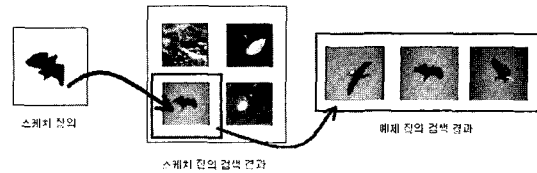


그림 2 영상 검색 과정

스케치 질의에 대한 유사도 측정은 Haar 웨이블릿 변환을 통해 추출된 계수와 컬러 히스토그램 값을 함께 이용하여 최종 유사 영상을 검색하며, 그 외 여러 특징을 이용해 검색을 할 수 있다.

4. 영상 분류

4.1 분류의 범위

시스템 구조에서 설명한 것과 같이 영상을 검색하기 위한 전 단계로 데이터베이스 영상 및 질의 스케치 영상을 클래스 별로 분류하게 되면 검색 성능뿐만 아니라 검색 시간을 줄일 수 있다. 하지만, 영상의 경우 같은 의미를 갖는다고 하여도 그 표현되는 형태가 매우 다양함으로 분류의 범위는 공통되는 의미의 중복을 최대한 줄일 수 있는 클래스로 분류로 하도록 하였다.

분류의 범위는 [그림 3]과 같이 사진 그림의 분류, 자연 도시 분류, 자연 영상의 세부 분류로 한정하였다.

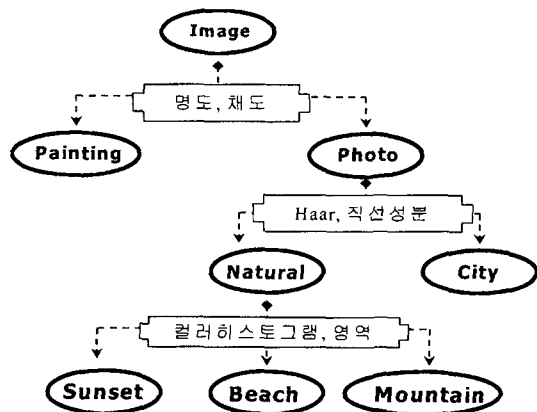


그림 3 분류 범위

4.1.1 사진과 그림 분류

사진과 그림의 분류에는 의미적인 중복보다는 그림과 사진 영상의 표현형태가 얼마나 다른가 하는 문제가 대두된다. 사진과 그림 분류의 경우는 사진과 그림의 차이

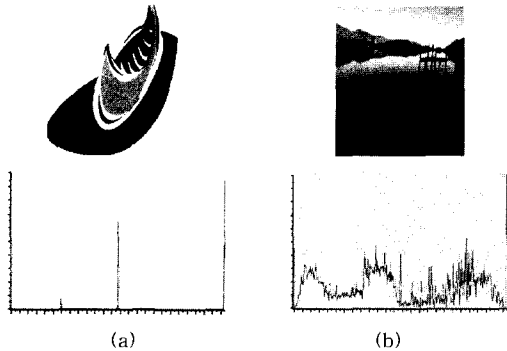
를 명확하게 알 수 있는 데이터, 즉, 사진과 클립아트 영상의 분류를 시행하였다.

그림의 경우 사진에 비해 색이 비교적 적게 사용되기는 하지만, 수채화, 판화 같은 경우는 [그림 4]의 (b)와 같이 색의 변화도 다양하며, 사진과 매우 흡사하기 때문에 두 카테고리를 명확히 할 수 있는 특징을 추출해 내기가 어렵다. 그래서, 사진-그림의 분류대상은 [그림 4]의 (a)와 [그림 4]의 (c)와 같은 영상을 대상으로 하였다.



(a) 사진 (b) 사진과 유사한 그림 (c) 그림
그림 4 사진-그림 분류 대상

[그림 5]는 분류할 대상인 사진과 그림이 갖는 컬러 특징을 보여준다. [그림 5]의 (a)와 같이 사진인 경우 컬러의 분포가 한 곳에 집중되어 나타나는 반면, (b)와 같이 사진의 경우 전체에 걸쳐 넓게 분포하는 모습을 볼 수 있다. 따라서, 영상 안에 포함된 컬러 분포가 다양할 경우 일반적인 자연 영상으로 분류하고 컬러 분포가 몇 개의 빈에서 높은 값을 갖는다면 그림 영상으로 판단하여 분류하도록 하였다.



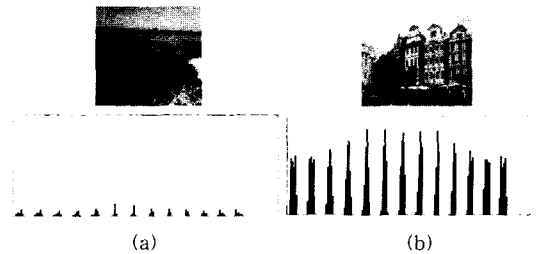
(a) (b)
그림 5 사진과 그림의 특성

4.1.2 도시영상과 자연영상 분류

도시와 자연영상의 분류는 직선성분을 중심 특징으로 이용한다. 본 논문에서는 Vailaya[9]의 방법을 응용하여 도시와 자연 영상을 분류하도록 하였다. 우선 각 데이터 베이스로부터 Haar웨이블릿 변환을 통해 에지 성분을

구한다. 직선 성분의 히스토그램을 구하는 방법은 5.2절에 자세히 설명되어 있다. 그런 뒤에 각 에지 성분의 방향성 정보를 추정하여 총 73개 빈(bin)을 갖는 히스토그램을 생성한다. 이것은 각 5개의 빈을 하나의 빈으로 양자화 한 것이다.

[그림 6]에서 보는 것과 같이 자연 영상의 경우는 직선 히스토그램 값이 작게 나타나는 반면, 도시 영상, 즉 건물이 많이 나타나는 영상의 경우는 히스토그램의 값이 매우 크게 나타나는 것을 볼 수 있다. 이러한 직선 히스토그램의 성질을 바탕으로 직선성분이 뚜렷하게 나타나는 도시 영상과 자연영상의 분류를 시행한다.



(a) (b)
그림 6 자연 영상과 도시 영상의 특성

4.1.3 자연 영상의 세부 분류

자연 영상의 경우 두드러진 컬러 영역들이 뚜렷하게 드러나는 경우가 많으며, 컬러 영역의 위치 또한 비슷한 경우가 많다. 해변 영상의 경우 파란색이 많이 나타나고, 하늘, 땅, 바다의 컬러 배열을 볼 수 있으며, 일출몰 영상의 경우에는 붉고, 검은 색이 많이 사용된다. 눈 덮인 산 같은 경우에는 흰색 계열이 많이 사용되며, 하늘, 산 등 이러한 컬러 영역의 배열을 찾아 볼 수 있다. 그래서, 자연 영상의 세부분류에는 컬러히스토그램과 컬러 영역의 상대적인 관계를 주 특징으로 하여 분류를 시행한다.

5. 검색을 위한 특징 추출

5.1 컬러분포 추출

Smith[11]는 컬러공간이 갖는 균일성(Uniformity), 완전성(Completeness), 간결성(Compactness), 자연성(Naturalness) 등의 모든 성질을 만족하는 컬러공간이 HSV 컬러 공간을 사용하여, Androutsos [12][13]의 실험결과를 바탕으로 흰색, 검은색 부분을 추출해 냈다. 나머지 컬러 부분에 대해 똑같은 범위를 갖도록 색상(hue)에서 18 구간, 채도(saturation)에서 3구간, 명도(brightness)에서 3구간으로 164 색으로 줄여 전체 컬러의 분포를 나타내는 컬러 히스토그램을 구하여, 히스

토그래프 교차법을 이용하여 유사도를 측정하였다.

5.2 방향정보 추출 및 직선 성분 추출

영상이 포함하고 있는 질감 정보(texture)정보는 Haar 웨이블릿 변환(wavelets transform)을 통한 계수를 사용하며, 직선 성분은 Haar 웨이블릿을 통해 얻어진 고주파 영역의 성분을 Hough 변환을 사용하여 추출한다.

Haar 웨이블릿 변환의 경우 저주파 성분과 고주파 성분의 분리 능력이 부족하다는 단점을 갖고 있다[10]. 하지만, 사용자 스케치 그림의 경우 비교적 같은 컬러 영역이 넓게 나타나며, 윤곽선이 뚜렷하여 Haar 웨이블릿 변환으로도 고주파 성분을 잘 분리할 수 있다[5]. Haar 웨이블릿 변환을 통해 얻어진 계수들 중에서 절대값이 큰 계수들을 200개 추출한다. 이렇게 추출된 계수의 비교는 식 (1)과 같다.

$$D_{haar} = \frac{1}{Total\ number} \sum_{i=0}^M \sum_{j=0}^N 1 \quad (1)$$

if $T_{component}[i] \equiv Q_{component}[j]$

T : 데이터베이스 영상의 Haar 색인 배열

Q : 질의 영상의 Haar 색인 배열

수식 (1)은 영상검색 시 컬러정보외에 추가적인 질감정보로 사용된다.

다음으로 영상분류를 위해 웨이블릿 변환된 고주파 성분만을 이용하여 직선의 방향성을 추정하도록 하였다.

Haar 웨이블릿 1단계 변환 후 얻어진 가로, 세로, 대각 방향의 성분을 추출한 후, 고주파 성분만을 가진 영상을 만들고, 이진 영상으로 변환한 후, Hough 변환을 이용하여 직선 성분의 히스토그램을 구성한다. [그림 7]은 2차원 직선을 극좌표계의 θ, r 로 표현한 Hough 변환과정을 보여준다. 본 시스템에서는 직선 성분을 추출하기 위해 원점과 직선상의 한점과의 길이를 나타내는 r 을 15단계, 직선의 기울기를 나타내는 θ 를 13단계로 나눠 r, θ 공간상의 빈도수를 구한다. 추출과정을 보면 [그림 8]과 같다.

이렇게 추출된 직선 성분 히스토그램은 히스토그램 비교법을 이용하여 비교를 수행한다.

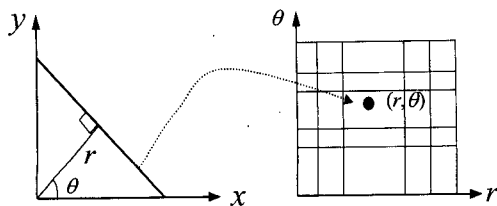


그림 7 Hough Transform

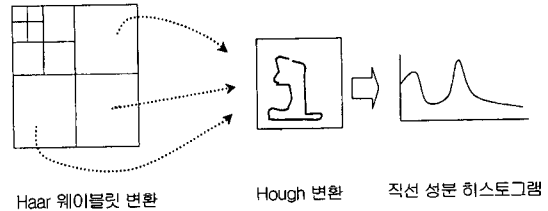


그림 8 직선성분 추출

5.3 컬러영역추출

컬러 영역을 추출하여, 각 컬러 영역의 상대적인 위치를 색인 함으로써, 유사한 컬러 배치를 갖는 영상을 검색할 수 있다. 영역추출의 전체적인 과정은 [그림 9]와 같다.

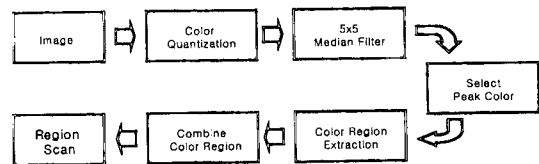


그림 9 컬러 영역 추출 과정

영역 추출 후, 영역의 색인은 Chang[14]이 제안한 2차원 문자열 비교법을 더욱 간략화 시킨 Smith[10]의 방법을 변형하여 사용한다. Smith[10]가 제시한 방법은 수직으로 영역을 스캔하면서 스캔 테이블을 구성했다. 이 과정을 간단히 설명하면 다음과 같다.

- ㉠ 질의 레이아웃의 숫자만큼 가로 세로의 행과 열을 만든다.
- ㉡ 데이터베이스 영상을 세로축으로 임의의 N개 부영역으로 분리한다.
- ㉢ 영상에 포함된 각 영역을 레이블(예, [그림 10]에서 1 과 2)하고, 레이블된 영역을 이용하여 행렬 1, 2와 같은 스캔 테이블을 만든다.
- ㉣ 행렬의 가로, 세로축은 영역의 번호를 나타내고, 행렬 안의 숫자는 영역들의 상대적 위치에 대한 빈도수를 나타낸다. 예를 들어 행렬 1에서 행1번과 열1번은 데이터베이스 영상 1에서 '1번 영역이 1번 영역 보다 위에 있는 경우의 빈도'를 나타내는 것으로 자기 자신의 위치를 나타낼 경우 0의 값을 갖는다. 역시 마찬가지로 1번 행의 2번 열의 경우 '1번 영역이 2번 영역보다 위에 있는 경우의 빈도'를 나타내는 것으로 그림 6의 (b) 행렬 1에서 보는 것과 같이 세로로 분리된 4개의 1번 영역이 2번 영역의 위에 위치함으로 4의 값을 갖는다.

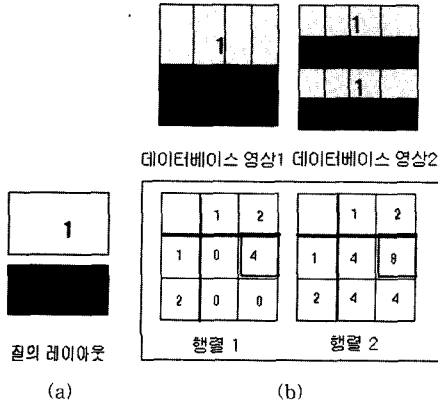


그림 10 영역 질의 문제점

이 방법의 경우 영역들간의 상대적 위치를 효과적으로 레이블할 수 있지만, 영역의 개수가 많아질 경우 비슷한 영상이라 하더라도 다른 영역으로 판단해 버리는 경우가 많고 [그림 10]-(b)의 우측 예에서 보는 것과 같이 실제로는 비슷한 영역이라 하더라도 약간의 영역간에 위치변동으로 인해 전혀 다른 값을 갖게 되는 문제점이 있다.

따라서, 영역 스캔을 하면서 같은 색을 가진 영역이 일정 크기 이상 나타나게 되면, 이를 서로 다른 영역으로 간주하여 영역간의 관계를 구하여, [그림 11]의 (b)에서 보는 바와 같이 영상1과 영상 2의 컬러1, 컬러2가 나타나는 빈도수를 유사하게 할 수 있어, 유사 영상으로 선택될 확률을 높일 수 있다.

[그림 12]는 15개의 샘플 데이터에 대해 영역질의 한 결과를 보여주고 있다. 앞에서 언급한 바와 같이, 질의 영상에 대해 영역이 크게 나타나는 영상을 수정된

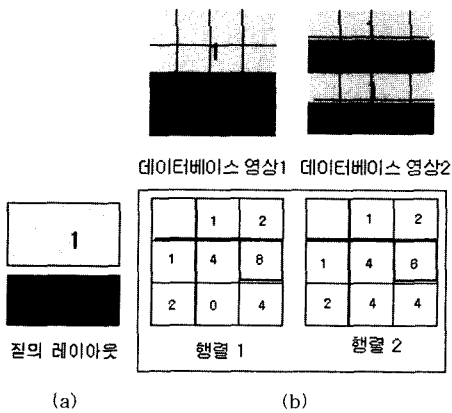


그림 11 수정된 영역질의

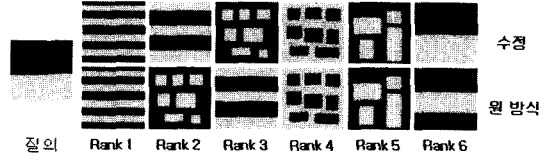


그림 12 영역질의 결과 비교

방법에 의해 6번째 유사 영상으로 찾았지만, 원 방식을 사용했을 때는 13번째 유사 영상으로 검색되었다.

스캔 테이블 구성은 [그림 13]이 자세하게 보여준다. [그림 13]의 (a)와 같이 최종적으로 영역이 추출되었다면, 이 영상을 스캔하며, 컬러 쌍의 관계를 색인 한다. [그림 13]의 (a)에서 보면, 영역을 스캔하면서 10이라는 컬러가 12 라는 컬러 보다 위에 있다는 것을 알 수 있으므로, 스캔테이블의 가로 10, 세로 12 되는 부분에 1을 증가시키면서, 컬러쌍의 빈도수를 색인한다.

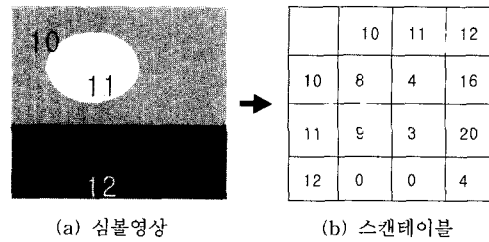


그림 13 영역관계 색인

유사도 측정은 질의 영상, 또는 사용자 스케치 영상에서 나타난 컬러 쌍이 데이터베이스 영상의 각각 스캔 테이블에 나타난 빈도수로 유사도의 기준으로 삼는다.

$$D_{region} = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} T[a_m][a_n]$$

if $a_m, a_n \in t_n (n=1,2,...)$ and $Q[a_m][a_n] \neq 0$ (2)

- T : 데이터베이스 영상의 스캔 테이블
- Q : 질의 영상의 스캔 테이블
- $a_m (m=1,2,...)$: 질의 영상의 선택된 컬러들
- $t_n (n=1,2,...)$: 데이터베이스 영상의 선택된 컬러들

5.4 기타 특징 추출

영상의 채도, 명암을 색인하기 위해 HSV 컬러 공간을 이용하여 쉽게 분리가 가능하다. 이 HSV 공간의 S, V 컴포넌트에 대해 100 개의 빈을 갖는 히스토그램을 구하여, 히스토그램 교차법을 이용하여 유사도 측정을 하도록 하였다.

5.5 색인 파일의 구조

다양한 검색과 분류를 수행하기 위해서 한 영상에서

많은 특징들을 추출해 낸다. 검색이나 분류 시에는 이렇게 추출된 모든 특징들을 다 사용하는 것이 아니라, 각각의 검색 방법에 맞는 특징들을 이용하게 된다. 그래서, 영상에서 추출한 이러한 많은 특징들을 XML을 이용하여 체계적으로 구성함으로써, 각 검색, 분류 시 필요한 특징들만을 손쉽게 이용할 수 있다.

표 2 색인 파일 구조

```

<color type = "COLOR_TYPE">
<histogram size = "BIN_NUMBER"> 빈 값들 </histogram>
<color_select size = "PEAK_COLOR_NUMBER"> 값 </color_select>
<scantable type = "SCAN_TYPE">
  <row number = "ROW_NUMBER"> 값들 </row>
</scantable>
</color>
<saturation>
<histogram size = "BIN_NUMBER"> 빈 값들 </histogram>
</saturation>
<intensity>
<histogram size = "BIN_NUMBER"> 빈 값들 </histogram>
</intensity>
<haar>
<colorcomponent type = "TYPE"> 빈 값들 </colorcomponent>
<positive size = "SIZE"> 값들 </positive>
<negative size = "SIZE"> 값들 </negative>
</haar>
<line>
<hueline size = "BIN_NUMBER"> 빈 값들 </hueline>
<satline size = "BIN_NUMBER"> 빈 값들 </satline>
<intline size = "BIN_NUMBER"> 빈 값들 </intline>
</line>
    
```

6. 실험 결과

6.1 실험 환경 및 데이터

실험 환경은 윈도우즈2000에서 JDK1.2, 아파치 웹서버, tomcat 서버릿 엔진을 이용하여 실험을 수행했다.

일반적으로 영상검색에서 성능평가를 위한 테스트 데이터가 없고 평가에 대한 기준들도 각 시스템마다 다르므로 본 논문에서는 코렐 데이터와 웹을 통해 사진, 그림을 포함한 1500여 개의 데이터를 수집하여 사용했다. 특징추출의 편의를 위해 256 × 256의 동일한 크기의 JPEG 영상으로 수정하여 특징추출을 통한 검색과 분류를 시행하였다. 일반적으로 성능평가를 위해서 precision과 recall을 많이 사용하는데 recall의 경우는 모든 데이터베이스안에 존재하는 영상들이 미리 분류되어 있어야 하지만 실제로 복잡한 영상을 접치지 않게 하나의 영상 카테고리만으로 분류하는 것이 쉬운 작업이 아니므로 본 논문에서는 precision만을 이용하여 성능을 평가하였다. 평균 검색 시간은 사용자의 스케치 질의 시간을 제외하고, 1500개의 영상에 대해 약 1.25초가 걸렸다.

6.2 영상 검색 결과

영상 검색은 스케치 질의와 예제 질의에 대해 수행하였다. 사용자 스케치 영상의 경우 정확한 그림을 그릴 수 없기 때문에, 컬러 영역이 크게 나타나며, 사용된 컬러 수도 제한적으로 나타나기 때문에, 사용자 스케치 질의의 경우 영상전체의 컬러 분포를 나타내는 컬러 히스토그램과 Haar 웨이블릿 변환을 통해 얻어진 계수를 이용하여 유사도 측정을 한다. 이렇게 함으로써, HSV 컬러 공간만을 사용한 Smith [10]의 방법보다 좋은 결과를 얻을 수 있었다([그림 16]의 상대적 영역질의 참조).

[그림 14]는 스케치 질의를 통한 결과를, [그림 15]는 상대적인 컬러 영역관계를 통한 검색 결과를 보여준다.

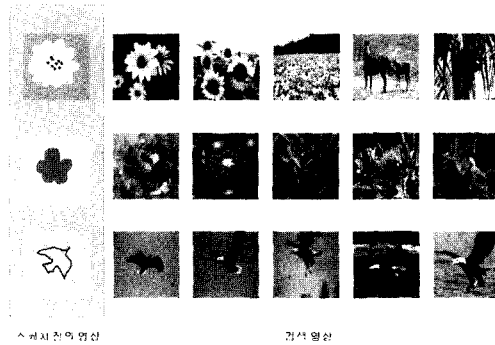


그림 14 스케치 질의 검색 결과

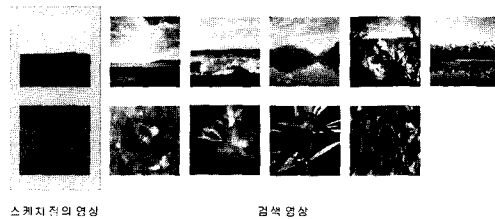


그림 15 영역 관계를 통한 검색 결과

[그림 16]은 새, 호랑이, 꽃, 일출물 영상에 대해 검색율을 측정하는 것이다. HSV 컬러 히스토그램과, Smith [10]가 제안한 상대적 영역간의 관계만을 이용한 방법, Haar 웨이블릿 변환 계수만을 사용한 방법, 그리고, 본 논문에서 구성한 시스템의 검색 결과를 보여주고 있다. 그림에서 보는 것과 같이 검색 성공률은 컬러 히스토그램만을 사용한 경우 평균 55%, 상대적 영역 질의는 58%, Haar 계수만을 사용한 방법의 경우 74%, 본 논문에서 제안하는 방법의 경우 87%의 성능을 보여주었다.

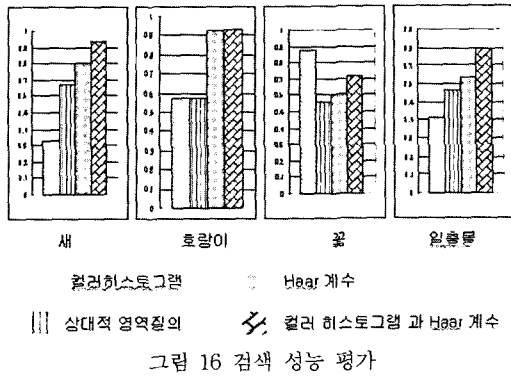


그림 16 검색 성능 평가

6.3 영상 분류 결과

영상의 분류는 4장에서 추출한 영상의 특징들과 그 특징들의 유사도 측정을 이용한 K-NN(K-Nearest Neighbor) 방법을 사용 하였다. 실험 데이터가 많지 않을 때 유용한 Leave-One-Out 방식을 K-NN과 함께 사용하여 영상의 분류를 하였고, 영상 데이터는 반드시 한 클래스에 속한다는 가정을 갖고 분류하였다.

6.3.1 사진 vs 그림

실험 데이터로는 사진 1000개, 그림 550 개의 데이터를 사용하여 K-NN과 Leave-One-Out 방법을 사용하여 분류 결과를 측정하였다. [그림 17]은 주 특징으로 사용한 명도와 채도를 이용한 방법과 영상 검색에서 유사도 측정으로 사용한 컬러 히스토그램과 Haar 계수를 이용한 방법을 K 값을 변화시키면서 측정한 결과를 보여주고 있다. 결과에서 볼 수 있듯이 직선 성분을 사용하지 않고 컬러 히스토그램과 Haar 계수를 이용한 분류

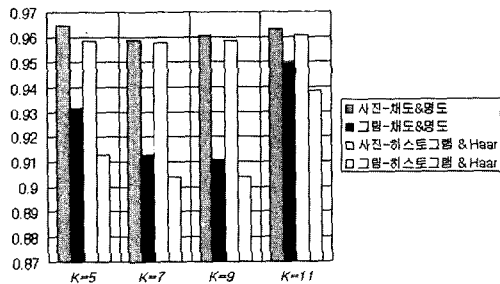


그림 17 K 값에 따른 사진-그림 분류 결과

표 3 사진-그림 분류결과

	사진	그림	총계	결과
사 진	960	40	1000	96 %
그 립	517	33	550	94 %
총 계	1477	73	1550	95.3 %

에서도 좋은 성능을 보여주고 있다.

6.3.2 자연 vs 도시

자연과 도시 영상의 분류는 직선 건물이 많은 특징을 이용하여 분류를 시행하였다. 실험 데이터로는 자연 영상 534개, 도시 영상 250개를 사용하였다. [그림 18]은 주 특징으로 사용한 직선 성분 히스토그램과 Haar 계수를 이용한 방법과, 영상 검색에서 유사도 측정으로 사용한 컬러 히스토그램과 Haar 계수를 이용한 방법을 K-NN의 k 값을 변화시키면서 측정한 결과를 보여주고 있다. 분류의 결과는 [표 4]와 같다.

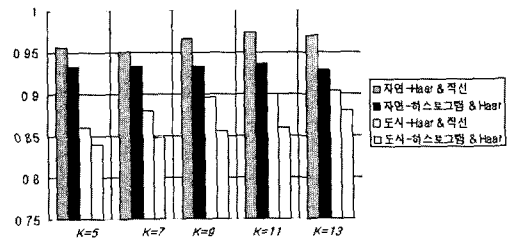


그림 18 K 값에 따른 도시 - 자연 영상 분류

표 4 도시-자연 영상의 분류결과

	자연	도시	총계	결과
자 연	517	17	534	97 %
도 시	25	225	250	90 %
총 계	542	242	784	94.6 %

6.3.3 자연 영상의 세부 분류

자연 영상의 세부분류로는 일 출물, 눈덮힌 산, 해변으로 분류한다. 자연 영상의 경우 의미상 중복되는 경우가 많아 분류에 어려움이 있었다.

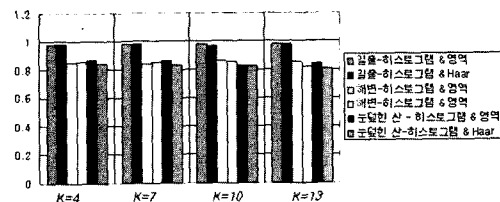


그림 19 K 값에 따른 자연 영상 세부분류 결과

표 5 자연 영상의 세부분류 결과

	일출물	눈덮힌 산	해변	총계	결과
일출물	223	3	4	230	97 %
눈덮힌 산	8	104	15	127	82 %
해변	6	12	107	125	86 %
총계	237	117	126	482	90 %

일출몰의 경우 분류의 결과가 상당히 좋은데 그 이유는 영상 대부분이 검은 배경을 가지고 있어서, 그 특징이 매우 뚜렷하게 나타나기 때문이다.

7. 결론 및 향후과제

본 논문에서는 영상 분류와 스케치 질의를 통한 웹 환경에서의 영상검색 시스템을 제안하였다. 영상을 분류하여 영상의 높은 단계의 의미 정보를 얻을 수 있고, 검색 시 검색 공간을 줄여 빠른 검색이 가능하며, 스케치 질의를 통해 사용자가 원하는 영상을 적극적으로 나타냄으로써, 검색에 있어 보다 사용자가 원하는 것을 잘 반영할 수 있었다.

본 시스템의 경우 스케치 질의 검색의 경우 컬러 히스토그램과 Haar 웨이블릿 변환을 통해 얻어진 계수를 이용하여 검색을 수행하고, 이 특징 외에 컬러 영역의 상대적인 위치를 통해 스케치 질의에 대한 검색을 수행할 수 있으며, 히스토그램과 웨이블릿 변환 계수를 함께 이용하여 검색 성능의 향상을 얻을 수 있었다. 또한, 영상 분류를 통해 영상 데이터베이스를 유사 영상끼리 분류하여, 검색 공간을 특정 클래스로 한정함으로써, 검색에 소요되는 시간을 줄일 수 있었다. 분류의 문제는 이전부터 연구가 되어 왔지만, 모든 데이터에 대해서 완벽한 결과를 보여준다는 것은 어렵고, 더욱이 영상 분류에 있어 분명 같은 의미의 영상임에도 불구하고 눈에 나타나는 성질이 크게 다를 수 있다는 점에서 분류의 어려움은 커진다. 그리고, 영상 분류를 함으로써 검색의 공간을 줄일 수는 있지만, 분류 오차로 인해 발생할 수 있는 검색 오차를 간과할 수 없기 때문에, 영상 분류를 최대한 정확하게 해야 한다. 본 시스템의 분류의 결과는 90~96%의 성능을 보여주었지만, 실험의 결과를 객관적으로 증명해 줄 수 있는 표준화된 데이터가 없기 때문에, 분류나 검색의 결과를 단정지어 판단하기가 어렵다.

차후 연구 과제로는 다음과 같이 정리 할 수 있다.

첫째, 내용기반 영상 검색, 분류의 성능 평가를 위한 표준화된 데이터가 필요하다.

둘째, 영상을 분류한 결과로 얻어진 의미 클래스들의 데이터를 체계적으로 정렬, 또는 연결을 함으로써 검색의 효율성을 얻고, 영상 데이터베이스를 브라우징 하는 데에도 적절하게 사용될 수 있을 것이다

셋째, 본 시스템 자체가 웹을 통해 연결이 되어 있으며, 많은 영상 데이터들이 웹을 통해 교환되는 점을 감안해 본다면, 에이전트를 통한 영상 자동 수집, 특성 추출, 분류를 시행함으로써 데이터베이스 구축시 드는 비용을 절감할 수 있을 것이다.

넷째, 분류의 다양화를 통해 향상된 검색 환경을 제공할 수 있을 것이다. 상표영상, 우표 영상 등을 구분하여 해당 영상에 대해서는 그 영상 분야에 적합한 알고리즘을 사용하여 영상을 검색할 수 있다면 시스템의 편의성을 증대시킬 수 있을 것이다.

참고 문헌

- [1] Venkat N. Gudivada, Vijay V. Raghavan, "Content Based Image Retrieval Systems," IEEE Computer, 1995.
- [2] Jing Huang, S Ravi Kumar, Ramin Zabih, "An Automatic Hierarchical Image Classification scheme," ACM Multimedia 1998.
- [3] <http://www.cselit.it/mpeg/>
- [4] Charles E. Jacobs, Adam Finkelstein, David Salesin, "Fast Multiresolution Image Querying," SIGGRAPH 95, NewYork, 1995.
- [5] Eric J. Stollnitz, Tony D, DeRose, David H. Salesin, Wavelets for computer Graphics: Theory and Applications, p.245, Morgan Kaufmann Publishers, Inc. 1996.
- [6] James Ze Wang, Gio Wiederhold, Oscar Firschein, Sha Xin Wei, "Wavelet-based Image indexing techniques with partial sketch retrieval capability," Proceedings of the Fourth forum on Research and Technology Advances in Digital Libraries, 1997.
- [7] J. R. Smith and S. F. Chang, "VisualSEEK : A Fully Automated Content-Based Image Query System," ACM Multimedia 1996, Boston MA, Nov, 1996.
- [8] Martin Szummer and Rosalind W. Picard, "Indoor-Outdoor Image Classification," IEEE International Workshop on Content-based Access of Image and Video Databases, Jan 1998.
- [9] A. Vailaya, A. K. Jain and H. J. Zhang, "On Image Classification : City Images vs. Landscapes," Pattern Recognition, vol.31, no. 12, 1998.
- [10] J. R. Smith, Chung-Sheng Li, "Decoding image semantics using composite region templates," IEEE Workshop on Content-based Access of Image and Video Libraries, June, 1998.
- [11] 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..
- [12] D. Androutsos, K.N. Plataniotis and A.N. Venet-sanopoulos, "Image region extraction for content-based image retrieval," EUSIPCO 1998, Rhodes, Greece, September 7-10, 1998.
- [13] D. Androutsos, K. N. Plataniotis, A. N. Venet-

- sanopoulos, "Vector Angular Distance Measure for Indexing and Retrieval of Color," SPIE Electrical Imaging '99, vol. 3656, p.604-613, January 1999.
- [14] Shi-Kuo Chang, Qing-Yun Shi, Cheng-Wen Yan, "Iconic Indexing by 2-D Strings," IEEE Transaction on Patter Analysis and Machine Intelligence, Vol. 9, No. 3, May, 1987.
- [15] Rafael C. Gonzales and Richard E. Woods. Digital Image Processing, p.716, Addison Wesley Publishing Company, Reading, Massachussetts, 1993.
- [16] Earl Gose, Richard Johnsonbaugh, Steve Jost, Pattern Recognition and Image analysis, p.484, Prentice Hall, 1996.
- [17] Ioannis Pitas, Digital image processing algorithms, p.362, Prentice Hall international series in acoustics, speech, and signal processing, 1993.



이 상 봉

1994~1999년 연세대학교 컴퓨터과학과 졸업(공학사). 1999년~2001년 2월 연세대학교 대학원 컴퓨터과학과 졸업(공학석사). 2001년 2월~현재 대우정보통신 연구원. 관심분야는 영상검색, 영상처리



고 병 철

1998년 경기대학교 전자계산학과 졸업(이학사). 1998년~2000년 연세대학교 대학원 컴퓨터과학과 졸업(공학석사) 2000~현재는 연세대학교 대학원 컴퓨터과학과 박사과정 재학중. 관심분야는 영상검색, 비디오 인덱싱, 패턴 인식, 인

공지능



변 혜 란

1980년 연세대학교 수학과 졸업(이학사). 1983년 연세대학교 대학원 수학과 졸업(이학석사). 1987년 Univ. of Illinois, Computer Science(M.S.). 1993년 Purdue Univ., Computer Science (Ph.D.). 1994년~1995년 한림대학교 정보공학과 조교수. 1995년~1998년 연세대학교 컴퓨터과학과 조교수. 1998년~현재는 연세대학교 컴퓨터과학과 부교수. 관심분야는 인공지능, 영상인식, 영상처리