

■ 2000년 정보과학 논문경진대회 수상작

영역기반 영상 검색을 위한 FRIP 시스템 (FRIP System for Region-based Image Retrieval)

고 병 철[†] 이 해 성^{**} 변 혜 란^{***}

(Byoungchul Ko) (Hae-Sung Lee) (Hyeran Byun)

요약 본 논문에서는, 영역 기반 영상 검색 시스템인 FRIP(Finding Region In the Pictures)을 제안한다. 이 시스템은 크게 색상과 방향성 질감 성분을 결합하는 굳건한 영상 분할 알고리즘과, 분할된 각 영역으로부터 특징 정보들을 추출하고 검색하는 3개의 알고리즘을 포함하고 있다. 영역 분할을 위해서, 영상으로부터 확장 및 이동된 색상 좌표계와, 방향성 질감 성분을 추출하여, 본 시스템에서 제안하는 원형 필터에 적용시킨다. 원형 필터에 의해, 영역의 경계선이 자연스럽게 유지 될 수 있고, 또한 일반적인 영역 병합 알고리즘에 의해 병합되지 않던 의미 없는 줄무늬나 작은 점 영역들도 몸체 영역으로 병합 될 수 있다. 영상을 분할한 후에, 효율적인 저장 공간의 관리와 특징 정보 계산 시간을 줄이기 위하여 각 영역으로부터 최적의 특징 정보만을 추출하고 이것을 색인화 하여 데이터베이스에 저장하고 검색에 사용한다. 사용자 인터페이스를 위해서는, 영역의 '색상', '크기', '모양', '위치' 와 같은 4개의 질의 조건을 주고, 사용자의 요구에 따라 정합 점수를 계산한 뒤, 그 점수에 따라 상위 검색 결과를 보여 주도록 설계되었다.

Abstract In this paper, we have designed a region-based image retrieval system, FRIP(Finding Region In the Pictures). This system includes a robust image segmentation scheme using color and texture direction and retrieval scheme based on features of each region. For image segmentation, by using a circular filter, we can protect the boundary of round object and merge stripes or spots of objects into body region. It also combines scaled and shifted color coordinate and texture direction. After image segmentation, in order to improve the storage management effectively and reduce the computation time, we extract compact features from each region and store as index. For user interface, by the user specified constraints such as color-care / don't care, scale-care / dont care, shape-care / dont care and location-care / dont care, the overall matching score is estimated and the top k nearest images are reported in the ascending order of the final score.

1. 서론

영상 기술의 급속한 성장과 더불어, 내용기반 영상 검색은 전자 도서관, 영상처리, 멀티미디어 데이터베이스 시스템 등과 같은 응용 분야에서 연구자들의 큰 관심을 끌고 있다. 더 나아가, 최근 빠른 하드웨어 기술의 성장

으로 인해, 대용량의 멀티미디어 데이터를 빠른 시간 안에 저장하고 검색하는 기술이 가능해 졌다[1]. 기존의 전통적인 색인 기술은 관리자가 영상영상 기술의 급속한 성장과 더불어, 내용기반 영상 검색은 전자 도서관, 영상처리, 멀티미디어 데이터베이스 시스템 등과 같은 응용 분야에서 연구자들의 큰 관심을 끌고 있다. 더 나아가, 최근 빠른 하드웨어 기술의 성장으로 인해, 대용량의 멀티미디어 데이터를 빠른 시간 안에 저장하고 검색하는 기술이 가능해 졌다[1]. 기존의 전통적인 색인 기술은 관리자가 영상을 직접 보면서 적절한 텍스트 내용을 입력하는 형식이었다. 이 방법은 실제적으로 가장 많이 사용되는 방법이지만, 대용량의 영상 데이터를 처리하는 데는 시간상으로 매우 비효율

[†] 본 연구는 한국과학재단 목적기초연구(98-0701-0201-3)지원으로 수행 되었습니다.

[†] 학생회원 : 연세대학교 컴퓨터과학과
soccer1@aipiri.yonsei.ac.kr

^{**} 비 회원 : 연세대학교 컴퓨터과학과
geneel@saver.co.kr

^{***} 종신회원 : 연세대학교 컴퓨터과학과 교수
hrbyun@aipiri.yonsei.ac.kr

논문접수 : 2000년 2월 17일

실사완료 : 2001년 4월 2일

적이며, 관리자의 주관에 의해 다른 내용이 색인어로 입력 될 수 있는 오류를 범할 수 있다. 따라서 내용기반 영상검색의 초점은 영상의 시간적인 내용을 자동으로 추출하고 색인화 하여 검색에 사용할 수 있도록 하는데 모아지고 있다.

현재, '영상에 의한 질의(query-by-image)'는 가장 대표적인 내용기반 영상검색 기술의 하나이다. 하지만, 만약 영상 검색 시스템이 단지 영상의 전역적 특성만을 이용한다면, 많은 잘못된 결과를 나타낼 수 있다. 예를 들어, 전역적 특성만을 사용하였을 경우, 사용자가 파란 하늘을 배경으로 하는 작은 크기의 흰색 기구를 질의하였을 때, 배경이 다른 색으로 바뀐다면 이 기구 영상은 유사 영상으로 검색 될 수 없다. 따라서, 이러한 단점들을 해결하기 위해, 최근에 대부분의 영상 검색 시스템은 객체(object) 또는 영역(region)에 의한 영상 검색에 초점이 맞춰지고 있다.

1.1 관련 연구

현재 몇몇 시스템에서 영상 데이터베이스로부터 색상, 질감, 모양 성분들을 이용하여 객체 또는 영역 단위로 검색하는 연구가 진행되어 오고 있다(예, Blob-world[3], QBIC[6], VisualSEEk[12], Netra[17]). QBIC의 초기 버전에서는 전체 영상과 수동으로 분할된 영역에 의한 질의 환경을 제공하고 있다. 하지만, 최근의 QBIC에서는 검색 정확성을 향상시키기 위해서 전경/배경(foreground/background) 분할 알고리즘을 추가하고 있다. Netra는 대용량의 영상 데이터베이스를 검색할 수 있는 도구로 개발되었다. 이 시스템에서는 Java 언어와 색상, 질감, 모양 정보 등을 사용하여 웹을 통해 작동되도록 설계되었다. 이 시스템의 가장 두드러진 특징은 객체 또는 영역

기반 영상 검색이 가능하도록 자동 영상 분할 알고리즘을 개발하였다는 것이다[17]. VisualSEEk는 영역을 추출하고 색상 집합에 의해 영역의 특성을 기술하여 검색에 응용하는 시스템이다. 색상 집합 back-projection 과정을 통해서, 영상으로부터 두드러진 영역을 자동으로 추출 한 뒤, 이를 이용하여 일반적인 내용기반 영상 검색뿐만 아니라 영역의 공간적 배치에 의한 검색 기능도 제공하고 있다. Blobworld는 영상의 화소로부터 연관성 있는 색상과 질감을 가지는 영역으로 변환하는 과정을 통해, 영상을 분할하고, 이렇게 분할된 영역(blob)들로 부터 각각의 색상, 위치, scatter matrix를 이용한 모양 정보 등을 추출하여 검색에 사용한다.

위에서 설명한 각각의 시스템들은 나름대로의 장점이 있지만, 각 시스템에 따라서 다음과 같은 몇 가지의 단점을 가지고 있다. QBIC의 경우, 사용자가 다양한 형태의 특징 값들을 사용할 수 있지만, 영상 분할에 사용자의 수 작업이 필요하다는 단점을 가지고 있다. Netra는 분할된 영상을 이용하여 영상을 검색하지만, 실제적으로 분할된 영역 수를 사용자가 임의로 넣어 주어야 한다는 단점이 있다[3]. VisualSEEk는 영상내의 객체 또는 영역의 공간 정보를 이용한 검색에 적합하지만, 색상 이외에 중요한 질감이나 모양 정보들을 사용하지 않고 있다. 마지막으로 Blobworld는 색상과 질감, 위치 정보를 결합하여 좋은 영상 분할 결과를 보여주지만, 한 영상을 분할하는데 너무 많은 시간(약 5분)이 걸리는 단점이 있다.

대부분의 시스템은 이러한 장단점에도 불구하고 다음과 같은 몇 가지의 공통된 특징을 가지고 있다.

- (1) 영상은 몇 개의 두드러진 영역들로 분할된다.

표 1 대표적인 영역 기반 영상 검색 기법에 대한 요약

시스템	영상 분할	특징 값	모양 특징 값
QBIC	<ul style="list-style-type: none"> · foreground/background algorithm · 제한 범위의 영상 사용 · 영상 분할과정 : 부분적으로 사용자의 수 작업 필요 	<ul style="list-style-type: none"> · 색상 · 위치 · 질감 · 모양 	<ul style="list-style-type: none"> · 영역 넓이(area) · 원형성(circularity) · 이심률(eccentricity) · 주축 방향성 · 모멘트
VisualSEEk	<ul style="list-style-type: none"> · color set back-projection · 영상분할: 색상 정보만 사용 	<ul style="list-style-type: none"> · 색상 · 배치 정보 · 영역 	<ul style="list-style-type: none"> · 사용 안함
Netra	<ul style="list-style-type: none"> · edge flow · 영상분할: 사용자에 의해 조절 	<ul style="list-style-type: none"> · 색상 · 위치 · 질감 · 모양 	<ul style="list-style-type: none"> · contour representation · 푸리에 변환
Blobworld	<ul style="list-style-type: none"> · EM algorithm (색상 및 질감 결합) · 영상 분할: 한 영상에 대해 약 5분 소요 	<ul style="list-style-type: none"> · 색상 · 위치 · 질감 · 모양 	<ul style="list-style-type: none"> · 무게중심(centroid) · scatter matrix (영역 넓이, 이심률, 기울기)

- (2) 분할된 각 영역들로부터 최적의 특징 벡터들이 추출된다.
- (3) 특징 벡터들의 집합은 데이터베이스에 색인으로 저장된다.

1.2 FRIP 시스템의 개요

FRIP(Finding Region In the Pictures) 시스템은 영역기반 영상 검색 시스템이다. 자동적인 영상 분할을 위해서 본 시스템에서는 원형 필터(circular filter)를 설계하였다. 원형 필터는 확장(scaled) 및 이동(shifted)된 색상 좌표계와 방향성 질감 성분을 결합하여 적용된다. 확장 및 이동된 색상 좌표계는 채도와 명도의 변화에 민감하지 않고, 색상 변환 시간을 줄 수 있는 장점이 있다. 방향성 질감 성분을 위해서는 쌍직교 웨이블릿 프레임(biorthogonal wavelet frame)을 사용하였다. 이때, 쌍직교 웨이블릿 프레임에 의한 고주파와 저주파 성분 중에서 방향성 성분을 나타내는 고주파 성분의 계수들만을 이용하여 방향성 질감 성분으로 사용한다. 본 시스템에서는, 이렇게 추출된 두 개의 특징 값을 이용하는 원형 필터와 보조적으로 영역 병합(region merge) 알고리즘을 적용하여 우수한 영역 분할 결과를 얻을 수 있었다.

영역이 분할된 후에, 각 영역으로부터 색상, 질감, 크기, 모양, 위치 정보 등을 색인화 하여 검색하는 작업이 필요하다. 본 시스템에서는 저장공간을 효율적으로 관리하고 계산 시간을 줄이기 위해, 각각의 영역으로부터 최적의 특징 값을 추출하도록 설계하였다. 이 중에서 영역 기반 검색을 위해 가장 중요한 모양 정보 기술(description)을 위해서 본 시스템에서는 이심률(eccentricity)과 변형된 형태의 반경 기반 시그니처(radius-based signature)를 사용하였다. 이를 이용하여, 영역의 전역적 특징뿐만 아니라 지역적인 특징도 함께 고려할 수 있다. 검색에서, 이심률은 대략적으로 모양이 일치하지 않는 영역들을 제거하는데 사용되고, 유사한 영역들로 선언된 영역들에 대해서만 시그니처를 이용하여 본 시스템에서 제안하는 유사성 측정과정을 통해 우선 순위대로 결과영상들을 사용자에게 보여 주게 된다.

본 논문은 다음과 같이 구성된다. 2장에서는 본 시스템에서 제안하는 영상 분할과정을 설명하고, 3장에서는 특징 값 추출 기술을 소개한다. 4장에서는 유사성 측정 알고리즘을 제안하고, 5장에서 실험 및 평가 결과를 설명한다. 마지막으로 6장에서 결론과 함께 향후 개선 방향을 소개한다.

2. 영상 분할 기법

영상 분할은 영상 분석을 위한 첫 단계로 영상 처리

분야에서 가장 오래된 연구분야 중 하나다[1]. 또한 이것은 영상 이해(image understanding) 또는 영역/객체 단위 영상 검색을 위한 필수적인 단계다. 특히, 영역기반 영상 검색의 경우, 검색의 성능은 영상검색의 정확성에 크게 좌우된다. 따라서, 대부분의 '영역에 의한 질의' 시스템[3][6][12][17]들은 정확한 색인과 검색이 가능하도록 영역의 자동 추출에 초점을 맞추고 있다.

본 장에서는 FRIP 시스템에서 제안하는 원형 필터를 이용한 영상 분할 알고리즘을 소개한다. 원형 필터를 위한 두 개의 특징 값 중 색상을 위해서는 확장 및 이동된 색상 좌표계[4]를 사용하고(2.1절), 방향성 질감을 위해서는 쌍직교 웨이블릿 프레임을 이용하였다(2.2절).

2.1 확장 및 이동된 색상 좌표계(scaled and shifted color coordinate)

2.1.1 일반적인 색상 좌표계를 이용한 영상 분할의 문제점

컬러 영상 분할은 많은 응용 분야에서 유용하게 사용되고 있다. 하지만, 영상으로부터 영역을 정확하고 자동적으로 추출하는 일은 현재의 컴퓨터 비전 기술로는 한계가 있다[16]. 이것은 일반적인 자연 영상들이 매우 다양한 형태로 외부로부터 간섭을 받고 있기 때문이다. 실제로 대부분의 자연 영상들은 짐승, 그림자, 표면 곡률(surface curvature), 폐색(occlusion)등으로 인해 간섭을 받고 있다.

컬러 또는 gray-level을 이용하는 몇 개의 영상 분할 알고리즘[14][16]등은 일반적으로 영상 내에 포함된 색상의 채도와 명도의 불안정성으로 인해 의미 있는 영역을 분할해 내는데 실패하고 있다. 이러한 문제들을 극복하기 위하여, 본 논문에서는 컬러에 대해서 확장 및 이동(scaling and shifting) 기술[4]을 적용하였다.

2.1.2 색상 좌표계 변경

인간의 시각 시스템(HVS)은 3가지의 지각 성분(색조, 채도, 명도)을 이용하여 모델링 된다. 여기서 색조(Hue) 성분은 일반적으로 영상에 대한 빛과 같은 자연적인 간섭현상에 대해 영향을 받지 않는다[4]. 따라서, 균정한 영상 분할을 위해서는 색상 좌표계 자체를 다른 색상 좌표계(예, HSV, munsell color, CIE-Lab)등으로 변경시키는 것보다는 영상에서 채도와 명도 성분만을 변형시켜 주는 작업이 더 효율적이다.

일반적으로 컬러 영상은 채도와 명도 성분을 변경시키기 위해서 RGB 색상으로부터 YIQ 색상으로 변경된다. 변경 후에 명도와 채도 성분을 변형시키고 다시 RGB 색상으로 변경시킨다. 하지만, 이러한 작업은 계산상으로 매우 비효율적이다. 따라서, 본 시스템에

서는 영상 분할을 위한 전처리 단계로 영상의 질을 향상시키기 위해서 확장과 이동 색상 좌표를 위한 하이브리드 기술[4]을 적용하였다.

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = k \frac{L'}{L} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + (1-k) \left(\begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} L'-L \\ L'-L \\ L'-L \end{bmatrix} \right) \quad (1)$$

$R'G'B'$: RGB 색상으로 부터 변형된 색상 좌표계

L' : 원 명도(L)값 으로부터 *contrast-stretch*된 명도 값

k : 확장(*scaling*) 및 이동(*shifting*)을 위한 *factor*

식(1)에서 명도 성분(L)은 LHS 색상 좌표계로 부터 다음과 같은 수식에 의해 얻어진다. $L=0.299R+0.587G+0.114B$. 본 시스템에서는 L' 을 *contrast-stretch*된 명도 값으로 정의한다. 명도 성분의 대비는 영상내의 밝고 어두운 화소 값의 분포를 의미한다. 따라서, 좋은 대비를 보여주는 영상은 화소 값이 한쪽에 치우치지 않고 고르게 분포하게 된다. 이 성질을 이용하여, 본 시스템에서는 전체 화소 값에 대한 영상의 히스토그램이 고르게 분포하도록 명도(L)성분에 대하여 *contrast-stretch*를 적용 시켰다. 식(1)에서, 만약 $k=0.5$ 일 경우, $R'G'B'$ 색상 성분은 원래의 RGB 색상으로부터 중간 정도의 확장과 이동이 발생하는 색상으로 변환되어, 채도 값과 명도 값이 모두 향상된 영상을 얻을 수 있다. 하지만, 본 시스템에서는 실험에 의해 채도 값을 확장시키는 것이 명도 값을 이동시키는 것보다 영상 분할을 위해서 중요하다는 것을 발견하였다. 따라서, 본 시스템에서는 k 값을 0.7로 정하여 사용하였다. 식(1)의 하이브리드 기술은 색상의 확장과 이동을 통해서 채도 성분이 과도하게 증가하거나 감소하는 문제를 막아준다[4]. 또한, 영상 분할을 위해 RGB 색상을 다른 색상계로 변환하지 않으므로써, 본 시스템에서는 향상된 영상을 얻을 수 있었을 뿐만 아니라 계산 시간도 단축시킬 수 있었다. 그림1은 수식(1)을 이용하여 원 영상을 확장 및 이동시킨 결과를 보여준다.

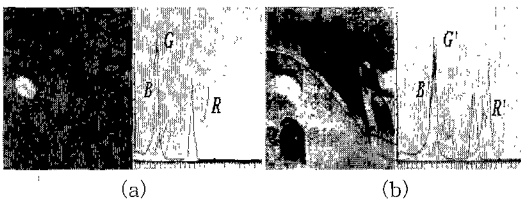


그림 1 (a) 원영상 (b)확장 및 이동 된 영상

2.2 쌍직교 웨이블릿 프레임을 이용한 방향성 질감 성분 추출

질감 성분은 기본적인 패턴들이 반복되어 나타나는 표현의 시간적인 특성으로 매우 중요한 특징 중의 하나다. 대비성(contrast), 균일성(uniformity), 거칠기(coarseness), 규칙성(regularity), 주기성(frequency), 조밀성(density), 방향성(directionality)과 같은 질감 성분은 영상을 해석하거나 분류하는데 있어 중요한 정보를 제공한다[8]. 대부분의 질감 성분은, Harr 웨이블릿 변환[5], Gabor decomposition[18], MSAR(Multiresolution Simultaneous Autoregressive Model)[10]이나 웨이블릿 변환 등에 의해 추출되고 있다.

본 시스템에서는 질감 성분을 추출하기 위해 고전적인 접근 방법을 사용하지 않고, 방향성 질감 성분을 추출하기 위해 쌍직교 웨이블릿 프레임(biorthogonal wavelet frame)을 사용한다. 프레임은 basis 집합의 over-complete 버전으로, 만약 함수나 벡터의 집합이 독립적이고 식(2)로의 확장을 허용한다면 그 집합은 프레임으로 불린다[7].

$$g(t) = \sum_k \langle g(t), \tilde{f}_k(t) \rangle f_k(t)$$

$g(t)$: 원 신호, $f_k(t)$: Basis 집합 (2)

$\tilde{f}_k(t)$: 이중(dual) basis 집합

쌍직교 웨이블릿 프레임을 이용하면, 원 영상으로부터 같은 크기의 저주파와 고주파 영상을 추출 할 수 있다. 다시 각각의 고주파 영상은 X-Y 방향을 가지는 2개의 부 영상으로 분할된다. 본 시스템에서는 첫 번째 단계의 고주파 부 영역으로부터 각 화소의 X-Y 성분을 추출해 낸다. 각 화소의 X-Y 방향성분은 2.3.2절에서 원형 필터를 위한 임계값으로 사용된다. 그림 2는 쌍직교 웨이블릿 프레임을 이용하여 원영상으로부터 1단계 X-Y축 고주파 성분을 추출한 결과를 보여주고 있다.

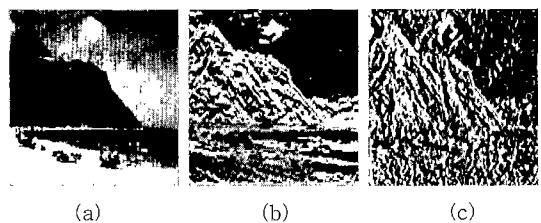


그림 2 1단계 고주파 영상 (a) 원 영상 (b) X-축 방향 고주파 영상 (c) Y-축 방향 고주파 영상

2.3 원형 필터를 이용한 영상 분할

본 시스템에서 영상은 변형된 색상 값과 쌍직교 웨이

블릿 변환으로부터 추출한 방향성 질감 성분을 이용하여 영역으로 분할된다. 이 과정은 다음과 같은 두 단계로 나뉘어 진다.

- (1) 원형 필터와 영역 병합(region merge) 방법을 이용한 1-단계 분할.
- (2) 영역 표시(region labeling) 방법과 영역 병합을 이용한 반복-단계 분할.

2.3.1 1-단계 영상 분할

우선, 영상의 변형된 색상 값들은 두드러진 색상 값으로 대략적인 양자화 과정을 거친다. 예를 들어, 만약 양자화 간격(interval)이 10이고 RGB의 색상 값이 각각 18, 32, 251이라면, 이 값들은 간격의 중간 값인 15, 35, 255로 양자화 된다. 다음으로, 영상에 포함된 잡음을 제거하기 위하여 메디안 필터가 적용된다. 세 번째로, 영상의 평균화(image averaging) 작업을 실시한다. 평균화 작업에 의해 영상에 포함된 두드러진 점들이나, 큰 영역 속에 포함된 작은 영역들이 흐려진(blurred) 영상으로 변화되므로, 원형 필터 작업이 쉽게 이루어 질 수 있도록 돕는 역할을 한다.

$$f(x, y) = \frac{1}{N^2} \sum_{x=1}^N \sum_{y=1}^N f(x, y) \quad (3)$$

$f(x, y)$: 평균화된 화소

N : 윈도우 크기

영상의 평균화 작업 후에, 개개의 화소들을 영역에 병합시키거나 영역 내에 포함된 잘못된 작은 영역들을 제거하기 위해, 본 시스템에서 제안하는 원형 필터를 적용한다. 원형 필터는 일반적인 영역뿐만 아니라 영역 병합 방법에 의해 병합되지 않는 줄무늬나 얼룩 등을 포함하는 동물 영역을 병합하는데 필수적인 과정이다. 적용과정은 우선, 필터내에 존재하는 화소에서 중심점을 제외한 모든 점들의 히스토그램을 구하고, 여기서 가장 많은 색을 가지는 화소들의 색상이 중심점 색상과 임계값 이하(수식(4) - T_1)에 존재하고, 마찬가지로 해당 색상 영역의 방향성 성분의 계수가 중심화소의 계수와 임계값(수식(4) - T_2) 이하에 존재한다면, 중심점을 가장 많은 색을 가지는 인접 화소의 색으로 변화 시켜준다. 이렇게 함으로써, 큰 영역내에 포함된 두드러진 작은 영역들은 큰 영역으로 병합 될 수 있게 된다. 여기서, 원형 모양의 필터를 이용한 것은 다른 모양의 필터에 비해 대부분의 자연 영상 또는 인공 영상들에 포함된 영역의 모양을 병합이 끝난 후에도 대체적으로 자연스럽게 유지시켜 줄 수 있기 때문이다. 본 시스템에서는 2가지 종류의 다른 크기를 갖는 필터를 사용하였다. 첫 번째 필터는 그림3-(b)에 나와 있는 것과 같이 11x11 크기의 윈

도우를 사용하고, 두 번째 필터는 7x7 크기를 갖는 윈도우를 사용한다. 이때, 첫 번째 필터는 1-단계 분할에 사용되고, 두 번째 필터는 반복-단계에서 사용된다. 본 논문에서 제안하는 원형 필터의 크기는 영상의 크기에 따라 변화 될 수 있다. 즉, 필터의 크기는 영상의 크기가 논문에서 사용하고 있는 데이터베이스 영상의 크기(320x210) 보다 2배 이상 커질 경우 마찬가지로 2배로 확장되고, 2배 이하로 작아질 경우 2배 작게 적용된다.

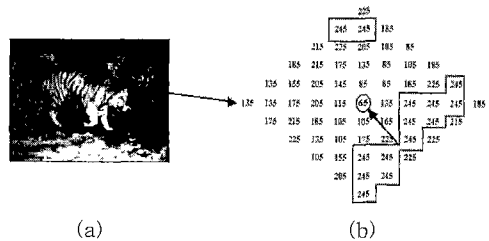


그림 3 본 시스템에서 제안된 원형 필터

그림 3-(b)에서 보는 것과 같이, 만약 ‘245’ 값을 갖는 화소의 수가 다른 값을 갖는 화소의 수보다 많고 식(4)를 만족한다면 원형 필터의 중심 화소인 ‘65’는 ‘245’로 변하게 된다. 수식(4)는 원형 필터내의 중심점이 가장 많은 화소를 갖는 색상의 값으로 변화되기 위한 조건 식이다. 여기서 첫 번째 임계값인 T_1 은 색상차를 위한 임계값이고, T_2 는 방향성 질감 성분을 위한 임계값으로 적용된다. 수식(4)에 의해서, 만약 중심화소가 2개의 임계값을 동시에 만족할 경우, 중심 화소는 가장 많은 색상을 가지는 이웃 영역으로 병합되게 되고, 그렇지 못할 경우 원래의 색을 유지하게 된다. 여기서, 원형 필터는 R'G'B' 3개의 색상 좌표계에 대해 각각 적용된다.

$$|C(a, b) - M_c| \leq T_1 \cap \left| \frac{C_y}{C_x} - \frac{M_y}{M_x} \right| \leq T_2 \quad (4)$$

$C(a, b)$: 중심 화소의 색상 값

M_c : 최대 화소수를 갖는 영역의 색상 값

C_y, C_x : 중심 화소의 X-Y 방향성 질감 성분

M_y, M_x : 최대 화소수를 갖는 영역의 X-Y 방향성 질감 성분 평균

본 시스템에서는 원형 필터를 이용하여 영역의 꼭뎀을 보호하고 영역 안에 포함된 크기가 작은, 의미 없는 영역들을 제거함으로써, 보다 효율적인 영상 분할을 할 수 있었다. 특히, 동물 영역의 경우, 몸 영역 안에 포함된 줄무늬나 점들은 대부분 일반적인 영역병합 알고리즘으로 제거 할 수 없었지만, 원형 필터를 적용시킨 후에 모두 하나의 영역으로 분리 될 수 있었다(그림 4).

지금까지 5가지 과정동안 접이나 줄무늬 혹은 영역 안에 포함된 의미 없는 작은 영역들을 제거 할 수 있었다. 하지만, 영상 안에는 아직도 큰 영역으로 병합되지 못한 의미 없는 작은 영역들이 많이 존재한다. 따라서, 마지막으로 이러한 작은 영역들은 영역 병합 방법에 의해 색상 값이 근사한 큰 영역으로 병합되게 된다. 영역 병합(region merging)방법은 동질성 측정 함수 식(5)를 만족 할 경우 이웃하는 영역으로 병합시킨다. 여기서 임계값 T 는 1-단계 과정에서는 고정된 값을 사용하지만, 반복-단계에서는 영역의 수에 따라 증가된다. 영역 병합 과정에서 영역 R 은 이웃하는 영역과 색상 값을 이용하여 병합을 시도한다. 만약, 병합이 실패하게 되면, 해당 영역은 새로운 영역으로 선언되게 된다. 이때 병합 규칙은 영역들의 평균 색상 값인 Rmi 에 기초한다. 즉, 한 영역이 다른 이웃 영역에 병합이 허용된다면, 두 영역간의 평균 색상 값은 갱신되게 되고, 이 값을 이용하다 다시 이웃 영역들과 식(5)를 이용하여 병합을 시도한다. 1-단계 과정은 그림 4에 나타나 있다.

$$\sum_i^N |R_{RGB} - Rm_{RGB}^i| \leq T \quad (5)$$

N : 이웃하는 영역의 수

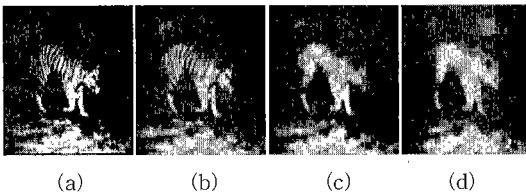


그림 4 1-단계 분할 결과 : (a) 원영상 (b) 영상 평균화 영상 (c) 원형 필터 적용 영상 (d) 1-단계 최종 결과 영상

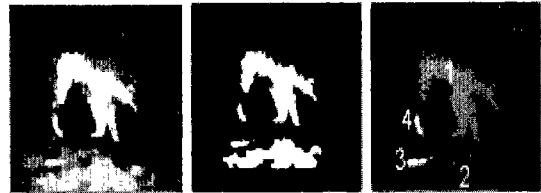
2.3.2 영역 표시(region labeling)와 반복-단계 영상 분할 1-단계 과정을 통해 분할된 영역들은 변화된 색상과 방향성 질감 성분을 사용하였다 하더라도 아직 과도하게 영역이 분할된 상태다. 또한, 경우에 따라서, 어떤 영역들은 서로 의미적으로 다른 영역임에도 불구하고, 유사한 속성을 갖기 때문에 한 개의 영역으로 선언될 수 있다. 이러한 문제점을 해결하기 위해 본 시스템에서는 영역을 구분하여 표시(region-labeling)하는 connected-component 알고리즘을 사용하였다.

이 과정은 3개의 색상 좌표계를 그레이 영상(I)으로 변환시킨 뒤에 적용된다($I=(3*R'+4*G'+2*B')/9$). 그레이 영상으로부터 같은 색상을 갖는 영역들이 비퍼에 '255'의 값으로 투영(project)되고, 이 이진 영상은 좌우,

상하 순서로 검사된다. 이 단계에서 같은 색을 가지는 영역들은 각기 다른 영역 번호로 표시 되게 되고(그림 5), 추가적으로 영역의 화소수가 30이하일 경우, 잡음 영역으로 판단하여, 영역으로 표시하지 않는다.

영역이 표시 작업이 끝난 뒤에, 영역의 총 수가 30 이상일 경우 7x7 크기의 원형 필터와 영역 병합 과정을 반복하게 된다. 동시에 영역 병합 임계값인 T 는 영역의 수가 30 이하가 될 때까지 계속적으로 증가한다.

본 시스템에서는 모양 기술(description)의 저장공간을 효과적으로 관리하기 위해 최대 분할 될 수 있는 영역의 수를 30으로 한정하였다.



(a) 1-단계 영상 분할에 의해 분할된 영역 (b) 비퍼에 투영된 같은 색을 가지는 영역 (c) connected-component 방법에 의해 다른 번호로 표시된 영역

그림 5 영역 표시(Region labeling) 과정

2.4 영상 분할 실험 결과

영상 분할 실험은 4개의 색상 좌표계(RGB, HSI, CIE-XYZ, R'G'B')에 대해서 각각 본 시스템의 알고리즘을 적용하여 실시되었다. 실험결과는 그림 6과 같다.

본 시스템을 이용한 영상 분할 실험 결과, 다음과 같은 결론을 얻을 수 있었다.

- (1) 색상을 이용한 영상 분할은 명도의 변화에 큰 영향을 받는다. HSI 색상 좌표계가 인간의 시각과 유사하게 설계되었다[14] 하더라도, 영상 안에서 명도 값이 상대적으로 작을 경우(어두울 경우) 오히려 잘못된 분할 결과를 보여 주었다(그림 7).
- (2) 영상 안에 채도와 명도가 균일하게 분포할 경우, 대부분의 색상 좌표계는 모두 좋은 결과를 보여 준다. 특히, RGB와 CIE-XYZ의 경우 HSI 색상 좌표계와 유사한 좋은 결과를 보여주었다.
- (3) 확장 및 이동된 색상 좌표계인 R'G'B'는 명도가 작거나 영상의 일부가 그림자 또는 표면 곡률에 의해 변경되었더라도 다른 색상 좌표계 보다 월등히 좋은 분할 결과를 보여 주었다.

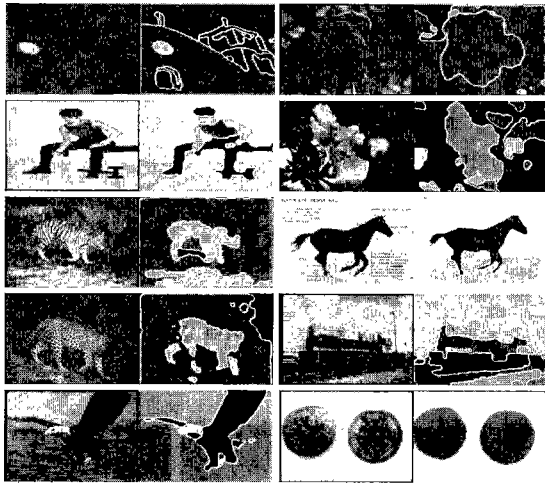
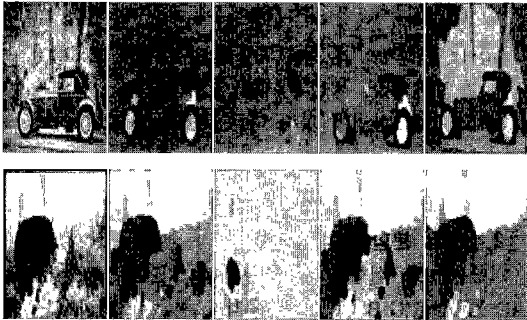


그림 6 영상 분할 결과

그림 7은 잘못된 영상 분할 결과를 보여준다.



(a) 원 영상 (b) RGB (c) HSI (d) CIE-XYZ (e) R'G'B'

그림 7 잘못된 영상 분할 결과

3. 특징 값 추출

영상을 영역 단위로 분할 한 뒤에, 각 영역들로부터 영역을 표현 할 수 있는 특징 값들을 추출하는 과정이 필요하다.

각 영역의 특징 값(색상, 질감, 위치, 모양)들은 영상의 내용을 표현하는데 사용된다. 하지만, 저장공간의 효율적인 관리와 계산 시간 단축을 위해서는 다음과 같은 3가지 조건을 만족 시켜야 한다.

- (1) 특징 값들은 반드시 간략하고 효율적으로 기술되어야 한다. 자세한 특징 정보는 대략적인 특징 정보 보다 정확한 영상 검색을 할 수 있지만, 과다

한 저장 공간의 관리와 계산 시간은 검색 정확률에 비해 오히려 비효율적이다.

- (2) 특징 값들은 반드시 특정 분야에 독립적이어야 한다. 즉, 검색 환경이 바뀌어도 한 개 혹은 그 이상의 특징 값들은 바뀐 분야에 적용될 수 있어야 한다.

- (3) 특징 값들은 각 영역의 일반적인 내용(색상, 질감, 모양 등) 뿐만 아니라 영역간의 구조적 관계(위치, 공간적 배치)도 포함해야 한다.

본 시스템에서는, 위의 3가지 조건을 고려하여 각 영역으로부터 색상, 질감, 영역의 크기, 모양, 위치 정보 등을 추출하였다.

3.1 색상 정보 추출

색상은 영상을 표현하는 가장 중요한 특징이다. 영상의 색상 분포는 흔히 색상 히스토그램을 사용하여 표현된다. 하지만 이 경우 각 영역을 모두 히스토그램으로 기술하기에는 bin의수가 너무 많으므로 본 시스템에서는 각 영역의 RGB 평균값을 색상 정보로 추출하였다.

3.2 질감 정보 추출

질감 정보를 추출하기 위해 본 시스템에서는 2.3.1 장에서 구해진 각 영역의 방향성 질감 성분을 질감 정보로 추출하였다. 역시, 첫 번째 단계의 고주파 영상으로부터 X-Y 방향의 크기(amplitude)를 추출하여 저장한다.

3.3 영역 크기 정보 추출

영역의 크기(scale or area)는 영역의 정보를 표현하는데 유용하게 사용될 수 있는 저 수준의 또 다른 특징 값이다. 본 시스템에서는 영역의 크기를 영역 내에 포함된 화소의 수로 정의하고 이 값을 크기 정보로 추출하였다.

3.4 영역의 모양 및 위치 정보 추출

모양 정보는 영역의 색상과 질감 성분이 유사할 경우, 영역을 구분해 낼 수 있는 가장 중요한 특징 값이다. 하지만 모양의 유사성에 의한 영상 검색은 상당히 어려운 문제다. 왜냐하면, 인간이 모양을 판단하는 것과 같이 의미적 유사한 모양을 판단해 낼 수 정확한 수학적 정의를 내리는 것이 힘들기 때문이다[8]. 더 나아가, 모양 기술(description)에 대한 연구들의 일반적인 문제점은 다양한 모양 기술 방법들의 정확성을 어떻게 판단할 것인가에 대한 기준이 없다는 것이다.

현재까지, 체인 코드(chain code), 다각형 근사(polygon approximation), 곡률, 푸리에 기술자(descriptor), 모멘트와 같은 여러 가지 방법들이 제안되어 왔고, 여러 분야에 적용되어 왔다[9].

표 2 일반적으로 사용되는 모양 기술 방법의 장단점

	체인코드	푸리에 기술자	모멘트
장점	- 구현이 용이 - 이동에 불변	- 구현이 용이 - 이동, 회전, 크기 변화에 불변	- 수학적으로 간결 - 이동, 회전, 크기 변화에 불변
단점	- 회전, 크기 변화에 민감	- 영역의 지역적인 정보를 제공 못함	- 항상 high order 모멘트가 모양 특징과 상관되어 있는 것은 아님 - 영역의 지역적인 정보를 제공 못함

하지만, 표 2와 같은 단점을 해결하고, 적절한 저장 공간의 관리와 효율적인 유사성 측정을 위해서, 모양 기술은 반드시 다음과 같은 기준들을 만족해야 한다.

- (1) 양 기술(description)은 반드시 이동(translation), 회전(rotation), 크기 변화(scaling)와 같은 3가지 조건에 불변(invariant)해야 한다.
- (2) 모양 기술은 반드시 최적의 특징들만을 저장하고 사용하기 쉬워야 한다.
- (3) 모양 기술은 반드시 전역적인 모양 정보뿐만 아니라 지역적인 모양 정보도 제공해야 한다.

본 시스템에서는 크기 변화, 회전, 이동 등에 불변하는 모양 기술을 위해 변형된 형태의 반경 기반 시그네처(radius-based signature)와, 이를 위한 빠른 모양 유사성 측정을 위한 식을 제안하였다.

3.4.1 FRIP의 모양 및 위치 정보 추출 알고리즘

영상 분할 작업이 끝난 뒤에 각 영역은 베퍼에 '255'의 값으로 투영된다. 그런 뒤에, 식(6)에 의해 영역의 무게중심(centroid)이 계산된다. 영역의 위치 정보는 이 무게중심으로 정의된다.

$$C_x = \frac{1}{N} \sum_{x,y} f(x,y)x \quad (6)$$

$$C_y = \frac{1}{N} \sum_{x,y} f(x,y)y$$

$f(x,y)$: 영역안의 한 화소
 x, y : 영역 R 안에 포함된 화소 좌표

투영된 영역으로부터, sobel 경계선 추출 알고리즘과 edge thinning 알고리즘에 의해 영역의 경계선을 추출한다. 경계선이 추출되고 나서, 모양은 시그네처에 의해 표현 되게 된다. 시그네처는 경계선을 1차 함수로 표현한 것으로, 본 시스템에서는 몇 가지 형태의 시그네처 방법(곡률 기반, 반경 기반, 경계선 좌표 기반)중 반경 기반 시그네처(radius-based signature) 방법을 사용하였다. 반경기반 시그네처는 영역의 무게중심으로부터 경

계선까지의 거리로 구성된다[8].

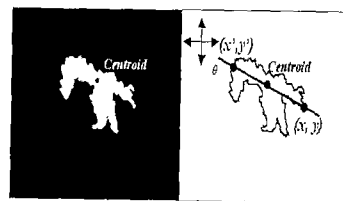
일반적인 방법에 의한 시그네처는 영역의 이동에는 불변하지만, 회전이나 크기 변화에는 민감하다. 따라서, 우선 회전에 불변하게 하기 위해, 본 시스템에서는 영역의 기울어짐 각도(θ)를 측정한다. 기울어짐 각도는 영역의 최소 관성 모멘트를 가지는 축의 각도로 정의된다 [2]. 이것은 식(7)에 의해 얻을 수 있다.

$$\theta = \frac{1}{2} \tan^{-1} \left[\frac{2C_{m2,1}}{C_{m2,0} - C_{m0,2}} \right] \quad (7)$$

$$C_{m \ n, \ q} = \sum_x \sum_y (x - C_x)^n (y - C_y)^q \quad (8)$$

여기서, 식(8)은 중심 모멘트(central moment)의 (p, q)번째 순서(order)가 된다. 영역의 기울어짐 각도로부터, 영역에 대한 중심 축을 만들고 무게 중심으로부터 중심 축과 맞닿는 2개의 경계선 좌표(그림 10-(b), (x' , y'), (x , y))까지의 거리를 구하여 그 중 가장 먼 거리를 갖는 경계선의 좌표를 시그네처의 시작점으로 간주한다. 그런 뒤에 반경 기반 시그네처는 시작점으로부터 시계방향으로 구해진다. 이때, 시작점에 대해서 2개의 시그네처가 존재할 수 있다. 다시 말해서, 만약 영역이 원 영역에 대해 x축 또는 y축으로 180° 회전되어 있다면, 시작점은 같더라도 시그네처의 내용이 달라질 수 있으므로 2개의 시그네처를 고려해 주어야 한다. 본 시스템에서는 저장 공간을 효율적으로 관리하기 위해, 데이터베이스 생성시에 시계방향으로 한 개의 시그네처를 저장해 두고, 검색시에는 질의 영역으로부터 양방향의 시그네처를 추출한 뒤 식(9)를 이용하여 분산을 구하고, 분산 값이 작은 쪽을 실제 시그네처로 정의하였다. 이런 방법으로 영역이 회전되어 있더라도, 두 영역이 동일하다면 유사한 시그네처를 구할 수 있다.

본 시스템에서는 모양 기술을 위한 색인의 크기를 줄이기 위해, 30도 마다 반경을 구하고 이 거리 값들을 영역의 색인으로 저장한다. 이렇게 저장된 시그네처는 본 시스템에서 제안하는 식(9)에 의해 크기 변화에 불변



(a) 영역의 무게 중심 (b) 영역의 중심축
 그림 8 시그네처를 위한 시작점 예측

하게 된다. 예를 들어, 만약 한 개의 영역을 A라고 가정하고, 크기가 변화된 영역을 B라고 할 경우, 영역의 유사성은 다음 식의 분산에 의해 결정 될 수 있다.

$$E_s = \frac{1}{12} \sum_{i=1}^{12} \left| \frac{B_i}{A_i} \right|$$

$$V_s = \frac{1}{12} \sum_{i=1}^{12} \left(E_s - \left| \frac{B_i}{A_i} \right| \right)^2 \quad (9)$$

E_s : 두 영역의 시그네처 사이의 평균값

V_s : 두 영역의 시그네처 사이의 분산값

식(9)에 의해 만약 크기가 다른 두 영역이 유사한 모양을 가진다면, 분산 값은 0이 된다. 그림 9는 변형된 반경 기반 시그네처를 이용하여 시작점을 예측하고 중심점으로부터 반지름의 거리를 구한 뒤 반지름 값이 정확함을 비교해 보기 위해 시그네처를 폴리곤 형태로 표현해 본 그림이다. 본 시스템에서 제안하는 모양 기술 방법은 영역의 회전 및 크기변화에 불변 할 뿐만 아니라 영역의 지역적인 정보를 함께 제공함으로써 영역의 작은 변형(distortion)에도 불변한 특징을 가진다. 본 시스템에서 제안하는 모양 기술 방법의 성능 평가는 4.1절에서 확인 할 수 있다.

시그네처와 더불어, 본 시스템에서는 각 영역에 대하여 bounding rectangle을 만든다. 검색시에 bounding rectangle을 통해 이심율을 구하고, 이를 이용하여 비유사 모양 영역들을 미리 제거해 줄으로써 모양 검색을 시간을 단축시킬 수 있다.

Bounding rectangle은 영역 포함을 포함하면서 기울어짐 각도가 일치하는 최소 사각형을 말한다. 여기서 bounding rectangle은 식(10)을 이용하여 영역의 경계 점들과 기울어짐 각도로부터 구할 수 있다.

$$a = x \cos \theta + y \sin \theta \quad (10)$$

$$b = -x \sin \theta + y \cos \theta$$

x, y : 영역의 경계 점

θ : 기울어짐 각도

각 경계 점들로부터 a_{max} , a_{min} , b_{max} , b_{min} 을 구해 내면, 바로 이 값들로부터 $l_1 = a_{max} - a_{min}$, $l_2 = b_{max} - b_{min}$ 를 구할 수 있고 이 두개의 길이 값으로 영역에 대한 사각형을 만들 수 있다. 여기서, 두 개의 길이 중 긴 쪽이 R_{max} 가 되고, 짧은 쪽이 R_{min} 이 된다. 이렇게 구해진 장축과 단축은 유사성 측정 단계에서 장축 대 단축의 비인 이심률을 구하는데 적용된다.

마지막으로, 이렇게 각 영역들로부터 구해진 12개의 특징 값들은 색인화 되어 데이터베이스에 저장된다. 최종적으로 저장되는 특징 값들은 표 3과 같다.

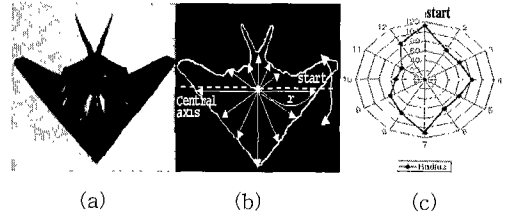


그림 9 변형된 반경기반 시그네처 (a) 원 영상 (b) 경계선 영역 (c) 12개의 시그네처를 이용하여 생성된 폴리곤

표 3 영역에 대한 기본적인 12개의 특징 값

특징 번호	기술 내용(description)	특징 번호	기술 내용(description)
1	영역 번호	7	무계중심의 y-좌표(Cy)
2	평균 Red 값(Ar)	8	장축(R_{max})
3	평균 Green 값(Ag)	9	단축(R_{min})
4	평균 Blue 값(Ab)	10	시그네처(Si, $i=1, \dots, 12$)
5	영역의 총 화소 수(NP)	11	영역의 y-방향성분 크기(Yd)
6	무계중심의 x-좌표(Cx)	12	영역의 x-방향성분 크기(Xd)

4. 유사성 측정

실제적인 정합(matching) 작업은, 데이터베이스에 저장된 영역들 중에서 질의 영상과 가장 유사한 특징을 가지는 상위 k개의 영상을 검색하는 것이다.

본 시스템에서, 사용자는 질의 하고자 하는 영상을 분할한다. 이때 특징 값들은 분할과 동시에 자동으로 추출된다. 그런 뒤에 질의 조건을 선택한다. 최종적으로 사용자가 입력한 질의 조건 즉, (1) 색상을 고려할 것인가?(color-care/don't care) (2) 크기를 고려할 것인가?(scale-care/don't care) (3) 모양을 고려 할 것인가?(shape care/don't care) (4) 위치를 고려할 것인가?(location care/don't care)에 따라서 식(17)에 의해 전체적인 정합 점수(matching score)가 계산된다. 여기서 (1)(2)(3)(4)의 각 조건들은 선형 결합(linear combination)에 의해 한 개의 통합된 값, 즉 최종 점수(score)로 나타내진다. 식(17)에서 사용된 각각의 특징 값을 위한 5개의 가중치 $a_1 \sim a_5$ 는 유사한 영상 집합의 600개 실험 영상을 사용하여, 휴리스틱(heuristic)한 방법으로 조절되었다. 검색 과정동안, 다음과 같은 6가지의 단계를 거쳐 유사한 영상들을 검색하고 최종적으로 통합된 점

수(score)에 의해 오름차순 형태로 결과를 사용자에게 보여 준다.

1. 초기 검색으로 검색 조건에 color-care가 선택되어 있을 경우, 질의 영상의 평균 색상과 이심률을 이용하고, color-don't care일 경우는 이심률만을 이용하여 데이터베이스에 저장된 영역과의 유사성을 비교한다(수식-(11), (12)).
2. 만약 두 영역간의 차이 값이 일정 임계값 이하라면, 후보 유사 영역으로 간주하고 다음 거리 측정을 허용한다. 이와 같은 초기 검색을 통해, 불필요한 검색 시간을 크게 단축시킬 수 있다.
3. 방향성 질감 성분의 차이 값을 계산한다(수식-(13)). 만약 이 값이 두 번째 임계값 이하라면, 사용자 질의 조건에 따라 다음 단계를 실행하고, 이상이라면, 벌칙 가중치(penalty weight)를 주고 다음 단계를 실행한다. 이것은 특히 자연 영상일 경우, 동일한 영역 혹은 물체라도 다른 질감 성분을 가질 수 있기 때문이다.
4. 사용자의 질의 조건에 따라, 모두 혹은 일부의 유사성 거리를 측정한다(수식-(14), (15), (16)).
5. 측정된 거리 값들로부터 최종 점수(final score)가 계산된다(수식-(17)).
6. 최종 검색결과가 점수의 오름차순에 따라 사용자에게 보여진다.

$$d_{Q,T}^C = |Ar_Q - Ar_T| + |Ag_Q - Ag_T| + |Ab_Q - Ab_T| \quad (11)$$

$d_{Q,T}^C$: 영역의 평균 색상차

$$E = \left| \frac{R_{Qmin}}{R_{Qmax}} - \frac{R_{Tmin}}{R_{Tmax}} \right| \quad (12)$$

E : 이심률, R_Q : 질의 영역, R_T : 데이터베이스 영역

$$d_{Q,T}^T = \left| \frac{Yd_Q}{Xd_Q} - \frac{Yd_T}{Xd_T} \right| \quad (13)$$

$d_{Q,T}^T$: 영역의 방향성 질감 성분차

$$d_{Q,T}^{Sc} = |NP_Q - NP_T| \quad (14)$$

$d_{Q,T}^{Sc}$: 영역의 크기차

$$d_{Q,T}^S = V_s \quad (15)$$

$d_{Q,T}^S$: 영역의 시그네처차, V_s : 식(9)에 의한 분산

$$d_{Q,T}^P = \sqrt{(Cx_Q - Cx_T)^2 + (Cy_Q - Cy_T)^2} \quad (16)$$

$d_{Q,T}^P$: 영역의 위치 거리

$$Score = a_1 d_{Q,T}^C + a_2 d_{Q,T}^T + a_3 d_{Q,T}^{Sc} + a_4 d_{Q,T}^S + a_5 d_{Q,T}^P \quad (17)$$

a_i : 특징들을 위한 가중치

식(17)에서 보는 것과 같이, 전체 영역의 거리는 색상, 질감, 크기, 모양, 위치 거리 값들의 합으로 나타내어진다. 이때, 전체영역간의 거리는 사용자의 선택에 의해 변경될 수 있다. 예를 들어, 만약 사용자가 크기에 대한 가중치를 다른 특징 값들에 비해 높여 줄 경우 사용자는 영역의 크기가 유사한 영상들을 결과로 검색할 수 있다.

4.1 모양 유사성 측정 테스트

시스템에서 제안하는 모양 유사성 측정 방식을 시험하기 위해서, 본 논문에서는 물고기 영상데이터베이스인 SQUID(<http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html>)로부터 100개의 무작위 영상을 다운 받고, 질의 영상으로 사용하기 위해 5 종류의 물고기를 선택하였다(그림 10). 본 실험에서는 FRIP에서 제안하는 모양 검색 방법의 성능을 평가하기 위해, shape-care 조건만을 선택하여 본 시스템에서 제안하는 방법이 실제로 영역의 이동, 회전, 크기변화에 불변한지를 검사해 보았다.

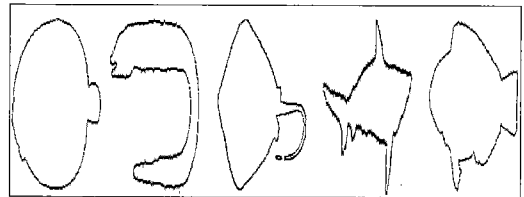


그림 10 모양 성능 평가를 위한 5종류의 질의 물고기 영상

모양 유사성 검색의 성능 평가를 위해, 본 논문에서 제안하는 12개의 시그네처를 사용하는 방법(MRS-12)과 24개의 시그네처를 사용하는 방법(MRS-24), 그리고 마지막으로 가장 많이 사용되는 푸리에 기술자(Fourier Descriptor)를 이용하여 성능을 비교하였다. 푸리에 기술자는 모양의 경계 점들을 64(FD-64)와 128(FD-128)개로 일정하게 샘플링하고 FFT를 이용하여 구현하였다. 이때, 회전에 불변하게 하기 위해, 위상(phase) 성분은 제외하고 계수(coefficient)의 크기(amplitude)성분만을 이용하였다. 또한, 크기 변화에 불변하게 하기 위해서 모양의 반경을 구하기 전에 최대 반경으로 반경들을 나누어 정규화 시키는 과정을 추가하였다. 최종적으로 푸리에 기술자의 유사성 측정은 유클리디언 거리(Euclidean distance)를 이용하여 계산하였다. 검색 성능 평가는 리콜(recall)과 프리시전(precision)을 통해 이루어졌다.

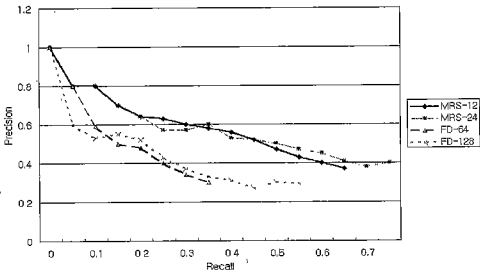


그림 11 푸리에 기술자와 본 논문에서 제안하는 방법의 검색 성능 비교

그림 11은 위에서 언급한 4가지 방법의 프리시전과 리콜의 성능 평가 결과를 보여준다. 그림 11에서 볼 수 있듯이, 본 논문에서 제안하는 방법(MRS-12)이 푸리에 기술자 보다 더 높은 프리시전과 리콜을 나타내는 것을 알 수 있다. 또한, 12개의 시그네처만을 사용하더라도 24개의 시그네처를 사용한 방법과 유사한 결과를 나타냄을 알 수 있다.

5. 실험 결과 및 평가

본 시스템에서는 웹과 Corel사의 Photo-CD로부터 얻은 자연 영상, 그래픽 영상, 그림 영상 등을 포함하는 2,700개의 영상 데이터베이스를 사용하였다. 본 시스템은 Visual C++ 6.0 언어를 사용하여 구현되었고, 오프라인(off-line) 시스템이다. 본 시스템에 대한 보다 자세한 내용 및 검색 결과는 <http://vip.yonsei.ac.kr/Frip>에서 확인 할 수 있다.

그림12는 FRIP 시스템의 사용자 인터페이스를 보여주고 있다. 현재까지 영역기반 영상검색에 대한 정확한 성능 평가 기준이 없으므로, 본 논문에서는 '해', '독수리', '호랑이', '꽃', '자동차'와 같은 5개의 객체에 대해서 질의한 결과를 프리시전으로 평가하였다. 이때, 검색 환경은 색상, 모양, 크기, 위치에 대한 조건을 모두 선택해 준 뒤에, 질의 영역을 분할시킨 뒤에 검색 버튼을 누르고 상위 18개의 영상에 대해 검색 결과를 판단하였다. 그림 13은 5개 객체에 대한 검색 성능을 보여준다. 그림에서 보는 것과 같이, 모양이 일정한 '해' 영역의 경우 높은 검색 결과를 보여주지만, '자동차'나 '꽃'의 경우 모양의 일그러짐, 또는 위치 변동에 따른 모양 변화 등으로 인해 낮은 검색 성능을 보여주고 있다. 마찬가지로, '호랑이'의 경우도 유사한 특성을 가지는 '표범'과 같은 동물들이 유사 영상으로 검색되는 오류가 나타난다. 그림 14, 15, 16은 5개 객체 중 '해', '독수리', '호랑이'에 대한 상위 9개의 검색 결과를 보여준다.

5.1 스케치에 의한 질의 환경

본 시스템에서는 또 다른 질의 환경으로써 QBIC과 같은 '스케치에 의한 질의' 환경을 추가적으로 제공한다. 이것을 이용하여, 사용자는 찾고자 하는 영역을 팔레트를 이용하여 그린 뒤에 똑같은 방법으로 질의에 사용한다.

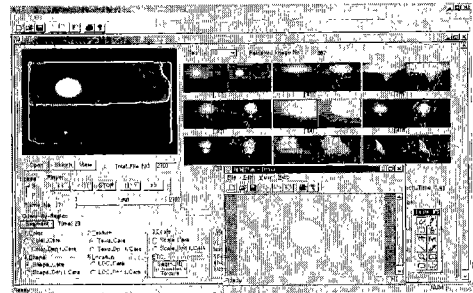


그림 12 FRIP 시스템 인터페이스

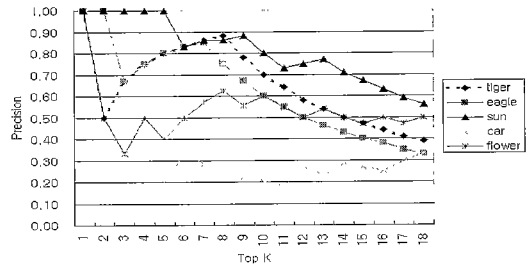


그림 13 5개 객체에 대한 검색 성능(Precision)

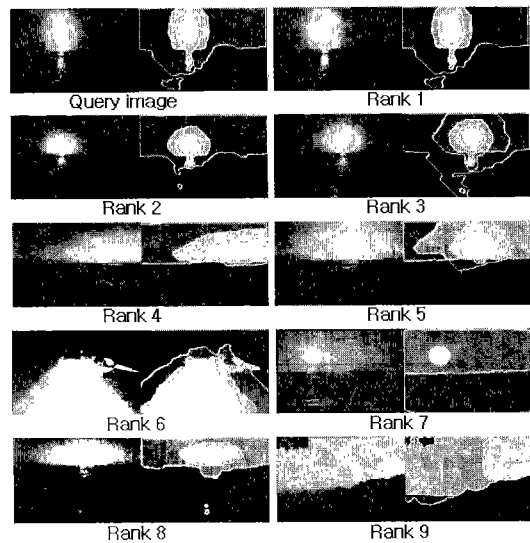


그림 14 '해' 영역 검색 결과(상위 9)

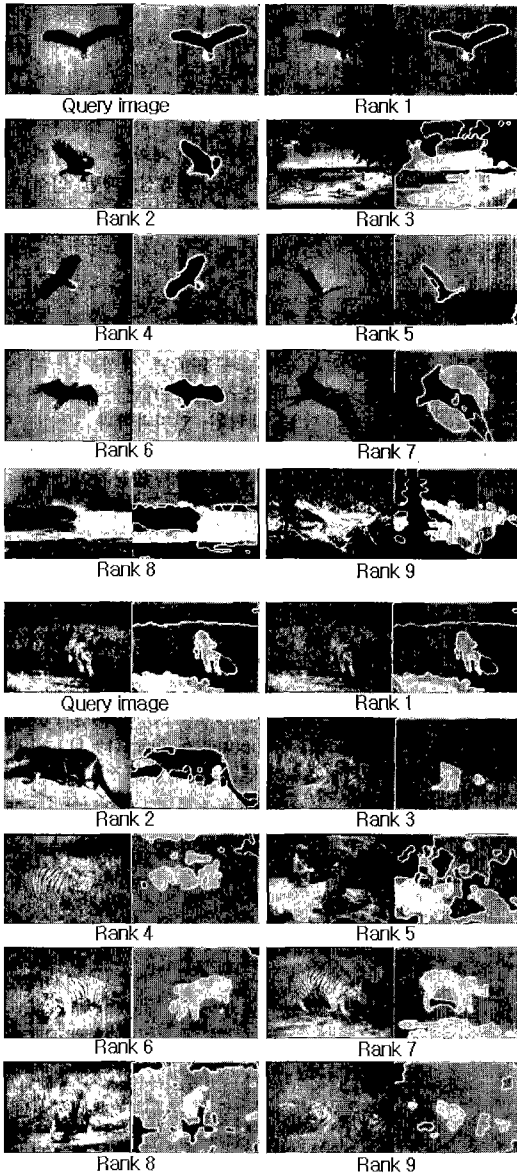


그림 16 '호랑이' 영역 검색 결과(상위 9)

6. 결론

본 논문에서는, 영역기반 영상 검색 시스템인 FRIP을 제안하였다. 이 시스템은 굳건한(robust) 영상 분할 알고리즘과 색상, 질감, 크기, 모양, 위치 정보 등을 이용한 검색 알고리즘을 포함한다. 더 나아가, 사용자에게 QBIC과 같은 '스케치에 의한 질의' 환경을 제공함으로써,

사용자가 보다 효율적으로 검색을 할 수 있도록 설계되었다. 결론적으로, 본 시스템의 목적은 사용하기 편하고, 검색 효율이 높은 영역 기반 영상 검색 시스템을 설계하는 것이다.

본 시스템의 주된 연구 결과를 요약하면 아래와 같다.

- (1) RGB 색상 좌표계를 다른 색상 좌표계로 변환하지 않음으로써, 계산량을 단축시킬 수 있었고, 하이브리드 기술에 의해 향상된 컬러 영상을 얻을 수 있었다.
- (2) 원형 필터를 사용함으로써, 영역의 경계선을 자연스럽게 보호 할 수 있었고, 영역 안에 포함된 의미 없는 줄무늬나 작은 점 영역들을 보다 크고 의미 있는 영역으로 병합 할 수 있었다.
- (3) 최적의 특징 정보만을 추출함으로써, 저장 공간을 효율적으로 관리할 수 있었고, 특징 값 계산 시간을 단축시킬 수 있었다.
- (4) 사용자에게 사용하기 쉽고, 검색하기 쉬운 사용자 인터페이스를 제공하고, 사용자의 목적에 따라 다른 영상을 검색 할 수 있는 환경을 제공하고 있었다.

참 고 문 헌

- [1] Moghaddamzadeh, and N. Bourbakis, "A fuzzy region growing approach for segmentation of color images," Pattern Recognition. Vol. 30, No. 6, pp. 867-881, 1997.
- [2] Anil K. Jain, Fundamentals of digital image processing, Prentice hall international editions, 1989.
- [3] Carson, M.Thomas, S. Belongie, J.M. Hellerstein, and J. Malik, "Blobworld: A system for region-based image indexing and retrieval," In Proceeding of International Conference Visual Information System, 1999.
- [4] Christopher C. Yang, Jeffrey J. and Rodriguez, "Efficient Luminance and Saturation Processing Techniques for Color Images," Journal of Visual Communication and Image Representation, Vol. 8, pp. 263-277, 1997.
- [5] Charles E. Jacobs, Adam Finkelstein and David Salesin, "Fast Multiresolution Image Querying," Proceeding of SIGGRAPH 95, NewYork, 1995.
- [6] Faloutsos, M. Flickner, W. Niiblack, D. Petkovic, W. Equitz and R. Barber "Efficient and Effective Querying by Image Content," Research Report #RJ 9203(81511), IBM Almaden Research Center, San Jose, Aug. 1993

- [7] Sideney Burrus, Ramesh A. Gopinath and Haitao Guo, Introduction Wavelets and Wavelet Transforms, A primer, Prentice-Hall, 1998.
- [8] Idris and S. Panchanathan, "Review of Image and Video Indexing Techniques," Journal of Visual Communication and Image Representation, Vol. 8, No.2, June, pp.146-166, 1997.
- [9] Guojun Lu, and Atul Sajjanhar, "Region-based shape representation and similarity measure suitable for content-based image retrieval," Multimedia Systems, pp. 165-174, 1999.
- [10] H. Zhang and D. Zhong, "A scheme for visual feature based on image indexing," Storage Retrieval Image Video Database III 2420, pp. 36-46, 1995.
- [11] J. L. Chen and A. Kundu. "Rotation and gray scale invariant texture identification using wavelet decomposition and hidden Markov model," IEEE Transactions on Pattern Analysis, March. pp. 208-214, 1994.
- [12] J.R. Smith and S.F. Chang, "VisualSEEK: A Fully Automated Content-Based Image Query System," ACM Multimedia, Boston MA, 1996.
- [13] L. Cinque, S. Levial, K.A. Olsen and A. Pellicano, "Color-Based Image Retrieval Using Spatial-Chromatic Histograms," International Conference on Multimedia Computing and Systems, Florence, Italy. June 7-11, 1999.
- [14] N. Ikonomakis, K.N. Plataniotis, M Zervakis and A.N. Venetsanopoulos, "Region growing and region merging segmentation," DSP-97, 1997.
- [15] Yining Deng, B.S. Manjunath and Hyundoo Shin, "Color Image Segmentation," Proceeding of IEEE Conference On Computer Vision and Pattern Recognition, 1999.
- [16] Yi Tao, William I. and Grosky, "Spatial color Indexing: A Novel Approach for Content-Based Image Retrieval," International Conference on Multimedia Computing and Systems, Florence, Italy. June 7-11, 1999.
- [17] W.Y.Ma and B.S. Manjunath, "Netra: A toolbox for navigating large image database," IEEE International Conference on Image Processing, 1997.
- [18] W. Y. Ma and B. S. Manjunath, "Image Indexing using a texture dictionary," Digital Image Storage Archiving Systems 2606, pp. 288-298, 1995.



고 병철

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

공지능



이해성

1995년 연세대학교 물리학과 졸업(이학사). 1995년 ~ 1997년 삼성코닝(주) 연구소 근무(연구원). 1997년 ~ 1999년 연세대학교 대학원 인지과학과 졸업(공학석사). 관심분야는 웨이블릿 이론 및 응용, 컴퓨터 그래픽스, 영상처리, 컴퓨터

비전, 신호처리, 인공지능



변혜란

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 ~ 현재 연세대학교 컴퓨터과학과 부교수. 관심분야는 인공지능, 영상인식, 영상처리