

학위 논문

영상검색을 위한 다중 영상특징 추출과 결합 방법이 관한 연구

□ 송석진 / KBS 부산총국 기술국

요 약

현재 사회 전반에 걸쳐 급격히 증가하고 있는 멀티미디어 정보를 효율적으로 관리, 활용할 수 있는 방법이 다양하게 연구되고 있다. 본 연구에서는 내용기반 영상검색을 위한 다중 영상특징 추출 방법과 특징결합 방법을 제시한다. 우선 전처리 및 캐니 에지 검출법으로 질의영상내 물체영역의 에지를 검출한다. 그 다음에 제안한 블록 다각형 알고리즘을 통해 분할된 물체영상을 획득한다. 분할된 물체영상은 HSV 공간으로 변환되고 히스토그램 인터섹션 방법으로 유사도가 측정된다. 또한 분할된 물체영상은 웨블릿 변환 영상으로도 변환된다. 이러한 변환후 웨블릿 부밴드의 LL 영역에 제한하는 거리밴드 평균 오토코릴로그램 알고리즘을 적용하여 오토코릴로그램 유사도를 측정한다. 그리고 GLCM을 이용한 엔트로피와 콘트라스트 유사도는 LH, HL 영역에서 측정된다. 전 과정을 통해 얻은 4개의 다중 영상특징은 수정된 보다 카운트 방법으로 결합되고 최종 유사도가 결정된다. 실험결과 제안한 다중 영상특징을 사용한 검색 방법이 단일 영상특징을 사용하는 검색 방법보다 소환성과 정확성의 성능에 있어 우수함을 보였다. 그리고 NMRR 측정에서도 개선된 성능을 보였다.

I. 서 론

현재 사회 전반의 다양한 분야에서 디지털화된 정보들은 기하급수적으로 증가되고 있다. 디지털 영상, 동영상, 음성 데이터 등과 같은 비정형 멀티미디어 정보의 활용은 대용량 저장장치의 발전, 컴퓨터 처리속도 향상, 초고속 통신망 구축, 그리고 정보처리 기술등의 발전으로 더욱 활발하게 진행되고 있다. 이에 따라 이들 멀티미디어 정보를 체계적이고 효율적으로 데이터베이스에 저장하고 관리할 필요가 대두되었다. 그러나 디지털 멀티미디어 정보는 양적으로도 엄청나며 분산화 되어 관리가 용이하지 않고 기존의 정보관리 시스템보다는 데이터의 저장, 전송, 디스플레이 등에 있어 성능이 우수해야 한다[1].

가장 이상적인 정보시스템은 기존 정보시스템의 유용한 기능과 계층적 저장시스템 지원기능,

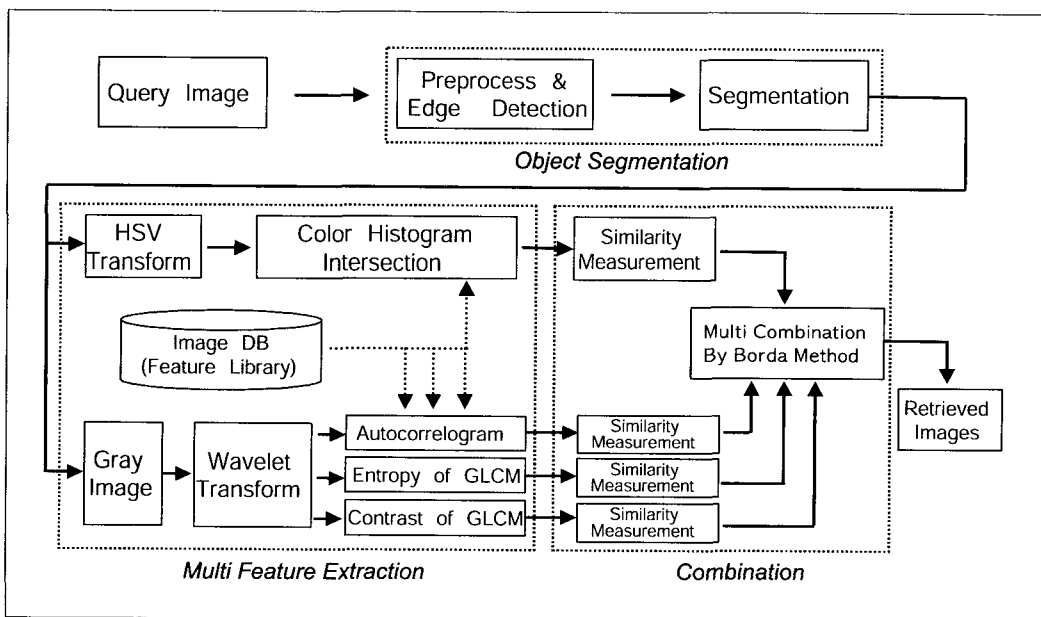
그리고 내용기반 정보검색 기능을 밀접하게 통합시킨 형태를 갖춰야 한다. 영상검색이란 영상을 분석하여 특징을 추출한 다음 이를 색인화하여 유사한 특징을 가지는 영상을 검색하는 기술로 정의할 수 있다. 대표적인 영상검색 시스템으로는 IBM에서 개발한 QBIC가 있는데 내용과 더불어 스케치, 사용자가 구성한 질의화상, 컬러, 레이아웃 기반으로 대규모 데이터베이스 질의가 가능하도록 개발하였다. 그리고 콜롬비아 대학의 Web Seek & Visual Seek은 사용자가 텍스트나 컬러 등을 질의요소로 하여 비주얼 쿼리를 작성하여 검색한다.

본 논문에서는 사용자가 시각적인 예를 제시하는 방식(query by example)으로 영상내의 질의를 원하는 물체영역을 선택하면 이를 배경으로부터 물체

를 분할한다. 그리고 분할된 물체영역에서 색상정보, 공간적 그레이분포 및 질감특성을 추출하여 영상 데이터베이스 내에서 유사 물체를 검색 가능하게 구현하였다.

II. 물체 분할

전체 블록도는 <그림 1>과 같다. 질의영상은 메디안 필터링과 캐니 에지 검출 과정을 거친 후 볼록 다각형 알고리즘을 통해 물체 분할된다. 그리고 분할된 영상을 HSV 변환시킨 후 히스토그램을 구하고 공간적 그레이분포 및 질감특성을 추출하기 위해 오토코릴로그램과 엔트로피, 콘트라스트 특징을 구한다. 이렇게 구한 다중 영상특징



<그림 1> 전체 블록도.

정보를 이용하여 데이터베이스내의 영상들과 4개의 유사도 거리를 구한후 수정한 보다 카운트 방법으로 유사도를 결합하여 최종 유사영상들을 출력한다.

1. 캐니 에지 검출법(Canny edge detection)

질의영상에는 잡음이 포함될 수 있으므로 전처리 과정이 필요하다. 일반적으로 잡음을 제거하기 위해서 저주파 통과 필터를 이용하는데 이것은 가우시안 잡음을 제거 하는데 적합하지만 솔트 및 페퍼 잡음, 임펄스 잡음을 제거하기 위해서는 적합하지 않다. 따라서 메디안 필터로 전처리하여 이런 단점을 보완한다(2). 그다음 캐니 에지 검출법으로 전처리된 영상에서 에지를 구한다.

2. 볼록 다각형 (convex hull) 알고리즘

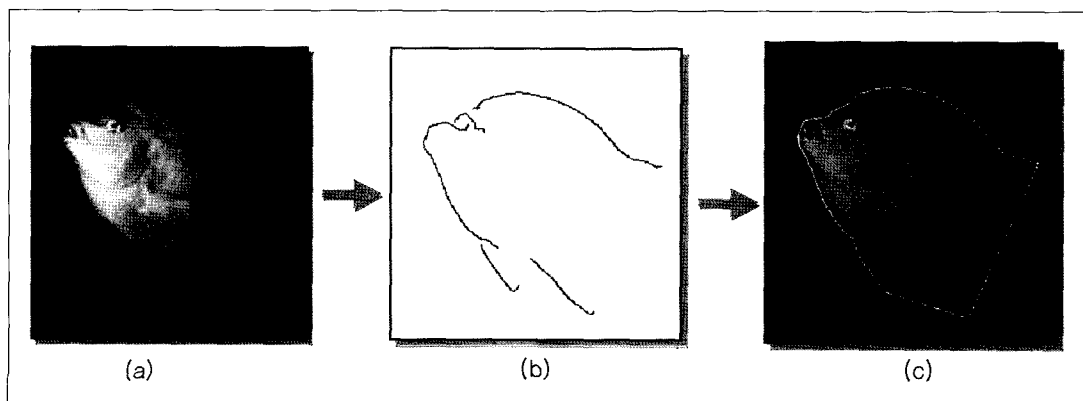
캐니 에지 검출법을 거친 에지는 물체 전체의 윤

곽을 나타내는 것이 아니고 <그림 2(b)>에서 보듯 물체의 강한 에지만이 나타나는 끊어진 에지이다. 따라서 다음과 같이 제안하는 볼록 다각형 알고리즘을 사용하여 <그림 2(c)>처럼 배경으로부터 물체를 분할한다.

알고리즘 절차

1. 가로방향으로 스캔을 하여 좌우의 볼록셋 (convex set)을 검출한다.
2. 세로방향으로 스캔하여 상하의 볼록셋을 검출한다.
3. 위에서 구한 상하좌우의 다각형 셋을 AND처리하여 공통되는 볼록셋만을 찾는다.
4. 공통되는 볼록셋에서 외곽 부분만을 차례로 연결하여 볼록 다각형을 구한다.

질의물체 내에 여러 색상과 에지가 있고 배경이 포함된 영상이라도 캐니 에지 검출법과 볼록 다각형 알고리즘을 사용하여 배경과 물체내부의 에지를 제외하고 원하는 물체의 윤곽을 연결하여 물체를 분할한다.



<그림 2> 물체분할과정 (a) 전처리된 영상 (b) 캐니 에지 검출 (c) 분할된 그림

Ⅲ. 컬러 영상에서의 특징 추출

1. HSV변환

선택된 질의영상은 HSV 좌표계로 변환되는데 이는 인간의 색인지를 가장 잘 표현해 준다. Travis 방식으로 HSV 좌표계로의 변환은 식(1)과 같다(3).

$$\begin{aligned}
 V &= \max(r, g, b), \quad S = \frac{V - \min(r, g, b)}{V} \\
 r' &= \frac{V - r}{V - \min(r, g, b)}, \quad g' = \frac{V - g}{V - \min(r, g, b)}, \\
 b' &= \frac{V - b}{V - \min(r, g, b)} \quad (\text{with } r', g', b') \quad (1) \\
 h &= 5 + b', \text{ if } r = \max(r, g, b) \text{ and } g = \min(r, g, b) \\
 1 - g', \text{ if } r = \max(r, g, b) \text{ and } g = \min(r, g, b) \\
 1 + r', \text{ if } g = \max(r, g, b) \text{ and } b = \min(r, g, b) \\
 3 - b', \text{ if } g = \max(r, g, b) \text{ and } b = \min(r, g, b) \\
 3 + g', \text{ if } b = \max(r, g, b) \text{ and } r = \min(r, g, b) \\
 5 - r', \text{ otherwise} \\
 H &= h \quad 60
 \end{aligned}$$

본 연구에서는 HSV변환을 위해 채도 S는 $0.2 \leq S \leq 0.6$, $0.6 < S \leq 1.0$ 로, 명도 V는 $0.2 \leq V \leq 0.6$, 0.6

〈표 1〉 컬러 매핑 테이블

색상 (Color)	색도 (Color Angle)	채도 (Saturation)	명도 (Value)
빨강	$0 \leq H < 18$	$0.2 \leq S \leq 0.6$ $0.6 < S \leq 1.0$	$0.2 \leq V \leq 0.6$ $0.6 < V \leq 1.0$
주황	$36 \leq H < 54$	"	"
노랑	$72 \leq H < 90$	"	"
연두	$108 \leq H < 126$	"	"
녹색	$144 \leq H < 162$	"	"
청록	$180 \leq H < 198$	"	"
파랑	$216 \leq H < 234$	"	"
남색	$252 \leq H < 270$	"	"
보라	$288 \leq H < 306$	"	"
자주	$324 \leq H < 342$	"	"
흰색	Don't Care	$0 \leq S < 0.2$	$0.8 < V \leq 1.0$
회색	Don't Care	$0 \leq S < 0.2$	$0.2 \leq V \leq 0.8$
검정색	Don't Care	Don't Care	$0 \leq V < 0.2$

$\leq V \leq 1.0$ 로 각각 2레벨 분할하여 정확하고 세밀한 색상 분석이 되도록 양자화 하였다. 그리고 색상 H는 먼셀(Munsell) 색상환에 따라 20색상 범위로 나누었다. 그리고 〈표 1〉과 같이 무채색은 따로 3레벨로 할당하여 모두 83레벨로 양자화하였는데 그중 몇 가지 색은 〈표 1〉에 나타내었다.

2. 대표색상을 이용한 컬러 히스토그램 인터섹션

히스토그램은 적은 계산량으로 색상분포에 대한 유사성을 효율적으로 비교할 수 있다. 따라서 히스토그램 인터섹션 기법으로 유사도 S를 측정한다(4). 두 영상의 유사도는 식(2)와 같이 표현된다

$$S_h = \sum_{j=1}^n (H_q(j) - H_t(j)) \quad (2)$$

여기서 $H_q(j)$ 와 $H_t(j)$ 는 각각 질의영상과 타겟영상의 히스토그램이고 j는 빈을 나타낸다. 하지만 분할할 때 조금 포함된 배경색을 제거하기 위해 식(3)처럼 유사도 S_h 는 질의영상의 히스토그램 $H_q(j)$ 에서 가장 큰 빈 값을 가지는 10가지 색상을 대표색상으로 선택하여 타겟영상의 히스토그램 H_t 와 컬러 히스토그램 인터섹션을 취한다.

$$S_h = \sum_{j=s_1}^{s_{10}} \min(H_q(j), H_t(j)) \quad (3)$$

Ⅳ. 그레이 영상에서의 특징 추출

1. 웨블릿 변환

웨블릿 이론의 주된 개념은 스케일링 함수

(scaling function)와 웨블릿 함수의 확장(scaling)과 천이(translating)에 있으며 웨블릿 변환의 기저 함수(basis function)는 스케일링 함수 $\phi(x)$ 의 스케일링 매개변수를 확장하고 천이함으로써 생성된다. 마더 웨블릿(MOTHER wavelet)이라 불리는 $\Psi(x)$ 함수의 확장과 천이를 통해 얻어지는 $\Psi_{m,n}(x)$ 는 다음과 같다[5].

$$\Psi_{m,n}(x) = 2^{-\frac{m}{2}} \Psi(2^{-m}x - n) \quad (4)$$

여기서 m, n 은 정수이다. 정규직교(orthonormal)성 질에 의해 신호 $f(x)$ 의 웨블릿 계수(wavelet coefficient)는 다음 식에 의해 계산된다.

$$c_{m,n} = \int_{-\infty}^{\infty} f(x)\Psi_{m,n}(x) dx \quad (5)$$

그리고 합성식은 아래 식과 같다.

$$f(x) = \sum_{m,n} c_{m,n}\Psi_{m,n}(x) \quad (6)$$

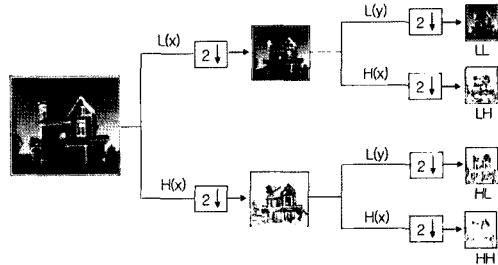
식(6)의 이들 웨블릿 계수로부터 $f(x)$ 를 복구하는데 이용한다. 마더 웨블릿 $\Psi(x)$ 를 만들기 위해 우선 스케일링 함수 $\phi(x)$ 를 결정하는데 식(7)과 같다.

$$\phi(x) = \sqrt{2} \sum_k h(k)\phi(2x - k) \quad (7)$$

웨블릿 함수 $\Psi(x)$ 는 스케일링 함수에 관계되고 식(8)과 같다.

$$\Psi(x) = \sqrt{2} \sum_k g(k)\phi(2x - k) \quad (8)$$

여기서 $g(k) = (-1)^k h(1-k)$ 이다. 웨블릿 변환을 수행하는 것은 $\phi(x)$ 와 $\Psi(x)$ 의 확실한 형태를 요구



〈 그림 3 〉 웨블릿 부대역 분해도.

하는 것이 아니고 단지 $h(k)$ 와 $g(k)$ 에 의존한다. 2차원 영상의 웨블릿 변환은 〈그림 3〉의 대역분해 부분에서 나타난 것과 같이 필터뱅크 구조를 구성하여 수평방향과 수직방향으로 1차원 이산 웨블릿 변환을 반복적으로 수행해서 얻어진다.

2. 오토코릴로그램 특징추출

1) 오토코릴로그램

오토코릴로그램은 영상의 각 화소가 갖는 색상의 공간적 상호관계가 거리에 따라 어떻게 변하는지를 나타내고 위치, 카메라줌 변화에 의한 모양 변화에도 강인한 특징을 보인다.

우선 $N \times N$ 의 크기를 갖는 영상을 I 로 두고 I 의 그레이 레벨을 m 개의 레벨로 양자화 시킨다. 영상 I 내의 위치 (x, y) 에서의 화소를 p 라고 두고 $I(p)$ 로 나타낸다. 따라서 $p \in I_c$ 는 $p \in I, I(p)=c$ 와 같은 의미를 갖는다. 여기서 c 는 화소의 색상을 나타낸다. 내용 전개에 편리를 위하여 두 화소 $p_1 = (x_1, y_1), p_2 = (x_2, y_2)$ 사이의 거리를 다음과 같이 정의한다[6].

$$|p_1 - p_2| \triangleq \max\{|x_1 - x_2|, |y_1 - y_2|\} \quad (9)$$

그리고 집합 $\{1, 2, \dots, n\}$ 은 $[n]$ 으로 정의한다. 거리 $k \in [d]$ 를 $d \in [n]$ 로 먼저 고정을 시킨 후, $i \in [m]$ 인 조건을 만족하는 영상 I의 오토코릴로그램은 다음과 같이 나타낸다.

$$\gamma_{c_i, c_i}^k(I) \triangleq P_r \left[p_2 \in I_{c_i} \mid |p_1 - p_2| = k \right] \quad (10)$$

$p_1 \in I_{c_i}, p_2 \in I_{c_i}$

위식은 그레이 레벨이 c_i 인 임의의 화소가 주어질 때, γ_{c_i, c_i}^k 은 주어진 화소로부터 거리 k 만큼 떨어진 화소의 레벨이 c_i 일 확률을 나타낸다. 오토코릴로그램 \mathbf{A} 은 다음과 같다.

$$\mathbf{A}(k, c_i) = \frac{\Gamma_{c_i, c_i}^k}{8k \times h_{c_i}} \quad (11)$$

여기서 분모의 $8k$ 는 주변화소 갯수를, h_{c_i} 는 색상 c_i 인 히스토그램이므로 분모는 임의의 화소로부터 거리 k 인 화소의 총합을 나타낸다. $k \in [d]$ 을 만족하는 각 k 에 대해서 색상 c_i 인 $p_1 \in I$ 와 색상 c_i 인 $p_2 \in I$ 가 $|p_1 - p_2| = k$ 을 만족하는 경우의 개수를 구한 후 룩업 테이블을 만든다.

2) 거리밴드 평균에 의한 오토코릴로그램

일반영상은 텍스처 영상과 달리 화소간 변화가 심하지 않고 타겟영상과 유사한 영상을 검색해야 하므로 영상간의 차이에 민감한 오토코릴로그램은 오히려 에러를 유발할 수 있다.

따라서 일반 오토코릴로그램을 수정한 거리밴드 평균에 의한 오토코릴로그램을 제안한다. 룩업 테이블에서 일정거리를 밴드로 묶어 각 그레이 레벨 별 평균값으로 룩업 테이블을 형성하는데 이때 밴드 거리 k 의 거리 범위는 다음과 같이 정의한다.

$$k'_j = (1 + b'_{j-1}) \sim (N_{db} + b'_{j-1}) \quad (12)$$

여기서 b_{j-1} 는 이전 거리밴드의 최대 거리값이고 N_{db} 는 거리밴드로서 N_{db} 는 3으로, k_j 의 j 값은 1~4의 범위로 정하였다. 그리고 그레이 레벨은 128로 하였다. 따라서 다음과 같이 오토코릴로그램을 나타낸다.

$$\mathbf{A}' = \sum_j \frac{\mathbf{A}^b(k, c_i)}{N_{db}} \quad (13)$$

질의 영상에서 위와 같이 구한 오토코릴로그램 $A_q(k, c_i)$ 은 타겟영상의 오토코릴로그램 $A_t(k, c_i)$ 과 비교하여 오토코릴로그램 유사도 S_a 를 구한다.

$$S_a = \arg \max_{1 \leq n \leq N} \left\{ \sum_{j=1}^k \sum_{i=1}^m \left[1 - \frac{|A_q(i, j) - A_t(i, j)|}{A_q(i, j) + A_t(i, j)} \right] \right\} \quad (14)$$

3. 엔트로피와 콘트라스트 특징 추출

질감은 2차원 공간적인 특성도 가지므로 단순한 1차원의 히스토그램 방법으로 영상의 질감 특징을 분석하는 것보다 효율적이다. 본 연구에서는 GLCM으로 알려진 p 매트릭스를 웨블릿 변환된 대역인 LH, HL에 적용하여 각 영역에서의 엔트로피와 콘트라스트 특징을 아래 식과 같이 구한다(7).

$$\mathbf{E} = - \sum_i \sum_j P(i, j) \log P(i, j) \quad (15)$$

$$\mathbf{C} = \sum_i \sum_j (i - j)^2 P(i, j) \quad (16)$$

웨블릿 변환된 영상의 LH, HL 두 영역에서 구한 엔트로피, 콘트라스트 값은 각각 더하여 평균하여 질의영상과 타 영상간의 회전으로 인한 유사도 오차를 줄인다. 이에 따른 유사도 거리는 다음식과 같

이 구한다.

$$S_e = 1.0 - \frac{\left| \frac{\left(\sum_{i=0}^1 E_q \right)}{2} - \frac{\left(\sum_{i=0}^1 E_t \right)}{2} \right|}{\frac{\left(\sum_{i=0}^1 E_q \right)}{2} + \frac{\left(\sum_{i=0}^1 E_t \right)}{2}} \quad (17)$$

$$S_c = 1.0 - \frac{\left| \frac{\left(\sum_{i=0}^1 C_q \right)}{2} - \frac{\left(\sum_{i=0}^1 C_t \right)}{2} \right|}{\frac{\left(\sum_{i=0}^1 C_q \right)}{2} + \frac{\left(\sum_{i=0}^1 C_t \right)}{2}} \quad (18)$$

V. 다중 영상특징 결합

1. 보다 카운트

다음은 일반적인 보다 카운트 계산식이다[8].

$$B_j = \sum_{k=1}^K (w^k \cdot G^k(i)) \quad (19)$$

각 유사도 측정에서 j 영상의 G 값을 합해 보다 카운트 값을 구한다.

$$G^k(i) = M - r + 1, \text{ if } C_r^k = i \quad (20)$$

식(20)에서 M 값은 전체영상의 개수이고 C_r^k 의 의미는 j 영상이 k 번째 유사도측정의 출력에서 순위가 r 이라는 것을 의미한다. 그리고 1를 더한 이유는 어떤 분류기의 맨 마지막 순위의 클래스는 다른 분류기에서의 순위에 상관없이 항상 마지막 순위가

되지 않게 하기 위해서이다.

2. 수정한 보다 카운트

결과 출력부에서 데이터베이스내의 입력 질의 영상과 유사한 영상 모두를 결과로 출력하기 위해 다음 식과 같이 변경한다.

$$CME(x) = C_0^{cme}, C_1^{cme}, \dots, C_{M-1}^{cme} \quad (21)$$

그리고 데이터베이스의 증가에 따른 전체 검색기의 분별력 저하를 방지하기 위해 아래와 같이 수정한다.

$$M = \max [N(e_1(x)), N(e_2(x)), \dots, N(e_k(x))] \quad (22)$$

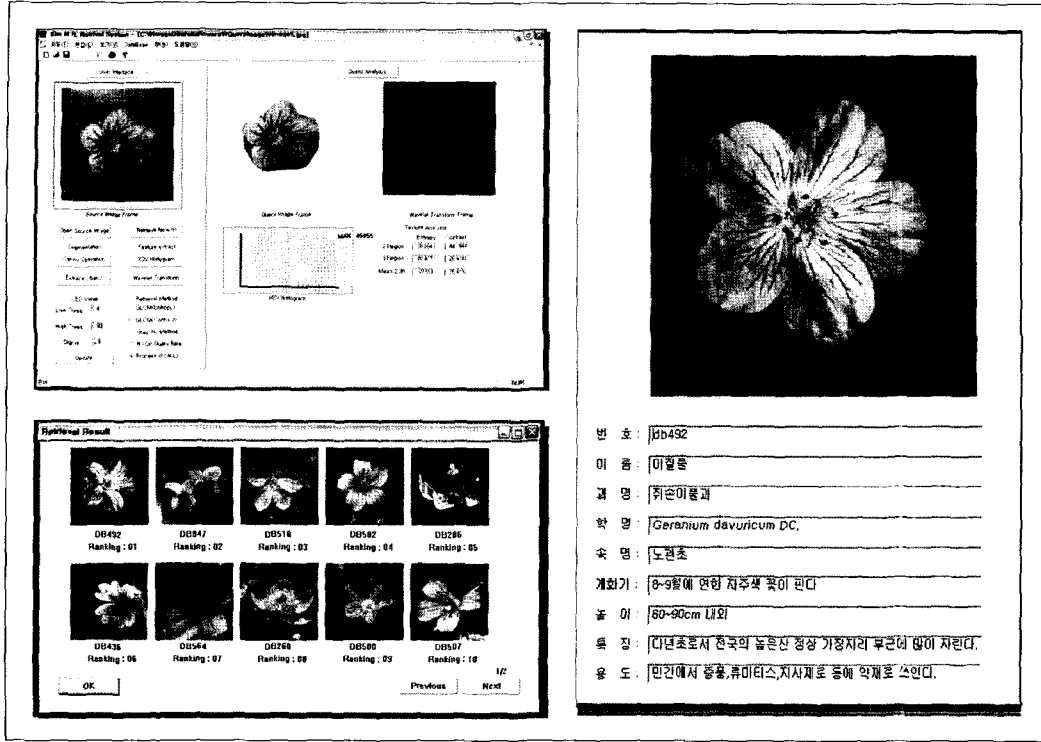
여기서 $N(e_k(x))$ 는 k 번째 검색기의 검색 결과 개수이다. 검색기가 저성능일 때 그 결과가 전체 결과에 큰 영향을 끼치므로 이에 대한 보완으로 식(19)를 다음식과 같이 수정한다.

$$B_i' = \prod_{k=1}^K (w^k \cdot G^k(i)) \quad (23)$$

그리고 여러 검색기의 결과에 포함된 영상일수록 질의 영상과 유사할 확률이 높으므로 아래식과 같이 가중치를 더한다.

$$B_i' = \prod_{k=1}^K (w^k \cdot G^k(i)) + M^{f-1} \quad (24)$$

여기서 f 는 빈도수이다.



〈그림 4〉 구현한 검색시스템 및 검색결과.

VI. 실험 및 결과

1. 실험환경과 구현시스템

본 연구에서 제안한 내용기반 영상검색 시스템은 Visual C++으로 구현하였다. 그리고 데이터베이스 구축은 40종류, 660여개의 야생화 영상에 대해 256×256크기로 정규화하고 특징들을 자동으로 미리 추출하여 특성 라이브러리에 색인화하여 저장하였다.

2. 소환성과 정확성 측정

일반적으로 유사 매칭을 수행하는 시스템의 평

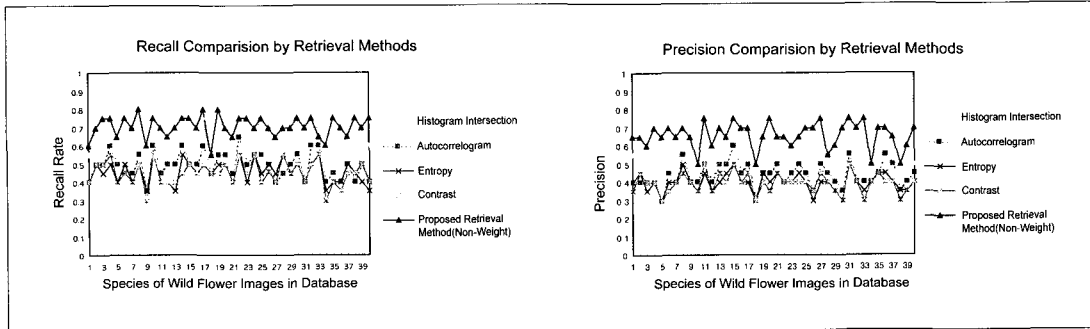
가로 소환성 및 정확성 측정 방법을 많이 사용한다. 소환성과 정확성은 다음식과 같이 정의될 수 있다.

$$Recall = \frac{a}{a + c} \tag{25}$$

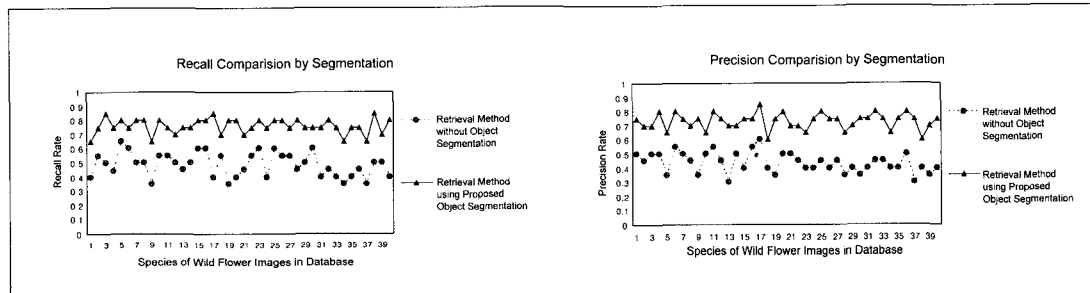
$$Precision = \frac{a}{a + b}$$

- a = Retrieved and Relevant (Detection)
- b = Retrieved and Not Relevant (False Alarm)
- c = Not Retrieved and Relevant (Miss)
- d = Not Retrieved and Not Relevant

각 검색방법에 따른 소환성 및 정확성의 실험결과는 그림 5와 같다. 제안한 다중 영상특징을 이용한 검색방법이 우수함을 알수 있다.



〈그림 5〉 제안한 검색방법과의 소환성, 정확성 비교



〈그림 6〉 물체 분할에 따른 소환성, 정확성 비교

그리고 물체분할에 따른 성능평가는 〈그림 6〉과 같다. 물체분할을 사용한 방법이 훨씬 나은 결과를 보였다.

3. 순위 측정

본 실험에서는 순위 측정의 일종인 NMRR(normalized modified retrieval rank)를 사용하는데 MPEG-7 그룹에서 정의된 것이다. NMRR은 다음의 과정을 통해 정의되며 $NG(q)$, $K(q)$ 을 표기한다.

$NG(q)$: 질의 영상 q 에 대한 관련 영상 개수

$K(q) = \min(4 \times NG(q), 2 \times GTM)$

, where $GTM = \max NG(q)$ for all q 's

$Rank(I)$ 는 검색결과에서 영상 I 의 순위라고할 때 다음과 같이 표현한다.

$$Rank(I) = \begin{cases} R(I) & \text{if } R(I) \leq F(q) \\ (F+1) & \text{if } R(I) > F(q) \end{cases} \quad (26)$$

식(26)을 이용하여 AVR(average rank)는 다음과 같이 정의된다.

$$AVR(q) = \sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)} \quad (27)$$

MRR은 다음과 같이 정의된다.

$$MRR(q) = AVR(q) - 0.5 - \frac{NG(q)}{2} \quad (28)$$

식(28)을 정규화한 NMRR은 다음과 같이 나타낸다.

$$NMRR(q) = \frac{MRR(q)}{K + 0.5 - 0.5 * NG(q)} \quad (29)$$

실험에서 검색 방법에 따른 NMRR을 측정하였는데 <표 2>는 그 결과를 나타내며 제안한 검색 방법이 검색된 영상의 순위가 0.126으로 보다 정확함을 볼 수 있다.

< 표 2 > 검색 방법에 의한 NMRR 측정값

	Histogram Intersection	Autocorre-logram	Entropy	Contrast	Proposed Retrieval Method
NMRR	0.214	0.286	0.371	0.393	0.126

Ⅶ. 결론

본 논문에서는 배경이 포함된 일반영상을 대상으로 색상 및 질감 등의 영상 특징을 함께 추출한 뒤 다중 영상 특징을 결합하는 방법을 소개하였다. 질의영상을 제안하는 블록 다각형 알고리즘으로 물체 영역만을 분할한다. 그리고 분할된 영상에서 컬러 특징을 추출하기 위해 HSV 컬러 좌표계로 변환시킨 후 대표 색상을 이용한 컬러 히스토그램 인터섹션을 취해 유사도를 구하였다. 또한 분할된 영상에서 질감 등의 특징을 추출하기 위해 웨블릿 변환을 시도하여 LL-부밴드영역에서 제안하는 방법으로 오토코릴로그램 유사도를 구하였다. LH, HL-부밴드에서는 엔트로피와 콘트라스트 유사도를 구하였다. 이와같이 구한 4개 유사도를 수정된 보더카운트 방법을 사용하여 결합하여 검색영상을 출력하였다.

실험결과 제안한 다중 영상 특징 결합에 의한 검색법이 다른 검색법보다 우수한 성능을 보였다. 그리고 순위 결정에 있어서도 개선된 결과를 보였다.

참고 문헌

- (1) A. W. M. Smeulders, "Content-based image retrieval at the end of the early years", IEEE Trans. on PAMI, vol.22,no.12,pp.1349-1351, Dec. 2000.
- (2) Jain, Kasturi, Schunck, "Machine Vision", McGRAW-HILL, USA, pp.169~173, 2000.
- (3) J. R. Smith and S. F. Chang, "Tools and Techniques for Color Image Retrieval", IS&T/SPIEProceedings,vol.2670,pp.4-6,Feb,1996.
- (4) M. J. Swain and D. H. Ballard, "Color Indexing", International Journal of Computer Vision, vol. 7, no.1, pp. 11-32, Nov, 1991.
- (5) T. Chang and C.-C.Jay Kuo, "Texture Analysis and Classification with Tree-Structured Wavelet Transform", IEEE Transaction on Image Processing, Vol. 2, no. 4, pp. 430-434, Oct, 1993.
- (6) J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlograms", CVPR, 1997.
- (7) M. Nadler and E. P. Smith, "Pattern recognition engineering", USA: John Wiley & Sons INC., 1993.
- (8) Y. R. Bae, J. Y. Lee, B. T. Jeon, K. H. Kim, and S. Y. Jung, "A content based automatic video parsing s/w", ETRI research report, 1999
- (9) G. H. Park, Y. J. Baek, and H. K. Lee, " A ranking algorithm using dynamic clustering for content-based image retrieval ", Int. conf. on Image and Retrieval, pp. 316-324, Jul 1-19, 2002.

필자소개



송석진

- 1993년 : 동의대학교 전자공학과 학사
- 1998년 : 부산대학교 전자공학과 석사
- 1998년 ~ 현재 : 부산대학교 전자공학과 수료후 박사과정
- 1995년 ~ 현재 : KBS부산총국 기술국 근무
- 주관심분야 : 멀티미디어응용, 뉴미디어, 영상검색 및 처리