

중앙 영역의 컬러 특징과 최적화된 빈 수를 이용한 내용기반 영상검색

류 은 주* · 송 영 준* · 박 원 배** · 안 재 형***

요 약

본 논문은 중앙 영역에서의 컬러 특징 추출 기법과 추출된 컬러 특징들의 비교 빈(bin) 수를 최적화한 새로운 내용기반 영상 검색 방법을 제안한다. 인간의 시각적 특징이 중심 객체의 유무에 영향을 받고, 대부분의 영상의 중심 객체는 중앙에 존재한다는 가정 하에 영상의 중앙 영역에서 컬러 특징을 추출한다. 따라서, 배경이 단순한 경우 영상의 전체영역을 특징으로 하여 검색했을 때 배경의 영향에 좌우되는 단점을 극복할 수 있다. 또한 영상의 컬러 특징값은 HSV 컬러 공간으로 변환한 후 16레벨로 양자화를 하여 추출한다. 실험값을 통해 기존의 16개 빈을 모두 비교하여 검색한 경우에 비해 상위 8개 빈만을 가지고 검색한 경우 주관적인 평가와 객관적인 평가 모두 다 좋은 결과를 보인다. 영상 전체를 특징으로 추출한 경우보다 중앙 영역만으로 특징을 추출한 경우 평균 precision이 약 5%정도 좋았다.

Contend Base Image Retrieval using Color Feature of Central Region and Optimized Comparing Bin

Eun-ju Ryu* · Young-jun Song* · Won-bae Park** · Jae-hyeong Ahn***

ABSTRACT

In this paper, we proposed a content-based image retrieval using a color feature for central region and its optimized comparing bin method. Human's visual characteristic is influenced by existent of central object. So we supposed that object is centrally located in image and then we extract color feature at central region. When the background of image is simple, the retrieval result can be bad affected by major color of background. Our method overcome this drawback as a result of the human visual characteristic. After we transform image into HSV color space, we extract color feature from the quantized image with 16 level. The experimental results showed that the method using the eight high rank bin is better than using the 16 bin. The case which extracts the feature with image's central region was superior compare with the case which extracts the feature with the whole image about 5%.

키워드 : 특징 추출(Feature Extraction), 비교 빈 수 최적화(Optimized Comparing Bin), 내용기반 영상검색(CBIR)

1. 서 론

현대 사회는 하루가 다르게 디지털 영상을 만들어내는 장비가 증가하고 또 이것을 이용하는 사람의 수가 늘어나고 있다. 디지털 카메라, 캠, 카메라가 달린 휴대폰에 이르기까지 일반 대중들이 디지털 영상을 만들어내는 장비를 일반적으로 사용하게 되면서 디지털 영상의 수는 헤아릴 수 없을 정도로 증가하였다. 또한 일반 포털 사이트에서 블로그를 제공함으로써 특정 인물들만 홈페이지를 소유하던 예전과는 달리 많은 사람들이 홈페이지를 소유하고 쉽게 디지털 영상들을 웹으로 올릴 수 있게 되어 현재 많은 디지털 영상이 웹으로 쏟아지고 있다. 따라서 개인이 하루에 접하는 영

상의 수는 많아졌다. 이렇게 많아진 디지털 영상들이 저장 도구가 발달함에 따라 개개인이 소유할 수 있는 영상의 수는 늘어나게 되었으나 영상의 관리는 어려워졌다. 이렇게 많아진 영상들을 쉽게 관리하기 위해서 내용기반 영상 검색 시스템이 필요하다.

일반적으로 내용기반 영상검색은 영상의 '컬러', '질감', '형태'의 세 가지 특징을 주로 사용한다[1]. 이 중에서도 특히 컬러는 영상을 구성하는 가장 지배적인 요소로 복잡한 배경에서 다른 특징들보다 상대적으로 강건하고 영상의 크기와 이동 및 회전에 민감하지 않다. 또한 특징의 추출과 활용이 쉽기 때문에 질감 특성이나 형태 특성보다 훨씬 많이 사용되고 있다. 컬러를 영상의 특징으로 사용한 방법 중 컬러 히스토그램이 가장 많이 사용되었는데 이 컬러 히스토그램 방법은 이동이나 회전에는 강하지만 공간 정보를 가지고 있지 않아서 전혀 다른 모양의 영상이지만 컬러 분포가 같

* 준 회원 : 충북대학교 대학원 정보통신공학과
** 준 회원 : LG전자 CDMA 단말 사업부 선임연구원
*** 정 회원 : 충북대학교 전기전자공학부 교수
논문접수 : 2004년 5월 15일, 심사완료 : 2004년 7월 29일

은 경우 같은 영상으로 볼 수 있는 단점을 가지고 있다.

Swain과 Ballard[2]가 제안한 히스토그램 인터섹션은 사용자가 제시한 질의 영상의 컬러 히스토그램과 영상 데이터베이스 내에 있는 모든 영상에 대한 각각의 컬러 히스토그램들을 비교해 유사한 영상을 검색하는 방법으로 계산이 간단하다는 장점을 지니고 있다.

이밖에 영상 정보를 표현하기 위해 Stricker와 Orengo[3]는 누적 히스토그램을 제안하였고 Stricker와 Dimai[4]는 히스토그램 인터섹션(histogram intersection) 기법에 공간 매칭을 결합한 방법을 제안하였다. 그러나 이 방법은 조명이 변화하는 경우 성능이 저하되는 단점을 가지고 있다. 이러한 문제를 해결하기 위해서 Funt[5] 등은 컬러 비율의 히스토그램을 비교함으로써 검색을 수행하는 불변 컬러 인덱싱(constant color indexing)이라는 알고리즘을 제시하였다. 그러나 이 방법은 일정한 조명에서는 Swain의 방법보다 조금 성능이 떨어지는 단점이 있다.

본 논문에서는 컬러 히스토그램 인터섹션 방법의 조명 변화와 영상 내의 물체의 크기에 민감하다는 단점과 또한 영상의 공간 정보를 이용하지 않고, 단순한 컬러의 빈도값만을 특징 값으로 사용하여 잘못된 영상 검색 결과를 초래할 수 있는 히스토그램의 최대 단점을 극복하고자 픽셀의 R, G, B 컬러 값뿐만 아니라 그 컬러의 사용 빈도수 또한 컬러 특징 요소로 사용하는 방법을 사용하였다.

본 논문의 구성은 관련 연구에서 이전에 컬러 특징으로 많이 사용되었던 컬러 히스토그램 방법에 대해 설명하고, 제안방법에서는 이전에 제안되었던 방법의 단점을 보완하여 컬러 특징을 추출하는 제안한 방법에 대한 기술과 비교 빈(bin) 수를 최적화 하는 방법에 대해 설명한다. 실험 결과를 통해 효율성을 보이고, 결론으로 글을 맺는다.

2. 기존의 컬러 히스토그램 방법

컬러 특징을 표현하는 가장 일반적인 기법이 컬러 히스토그램이다. 컬러 히스토그램은 영상의 전체적인 특징(global feature)을 표현한다. 이러한 컬러 히스토그램을 구축하기 위해서 먼저 적절한 컬러 공간을 선택하고, 해당 컬러 공간에서 양자화를 수행하여 정보의 크기를 줄인다. 컬러 히스토그램은 영상 내의 각 화소에 대해 동일하게 양자화 된 컬러를 계수함으로써 얻을 수 있다. 이렇게 얻어진 컬러 히스토그램은 해당 영상의 컬러 특징으로 사용된다.

Swain과 Ballard가 제안한 히스토그램 인터섹션은 사용자가 제시한 질의 영상의 컬러 히스토그램과 영상 데이터베이스 내에 있는 모든 영상에 대한 각각의 컬러 히스토그램들을 비교해 유사한 영상을 검색하는 방법으로 계산이 간단하다는 장점을 지니고 있다. 식 (1)은 이러한 히스토그램 인터섹션을 정의한 식이다.

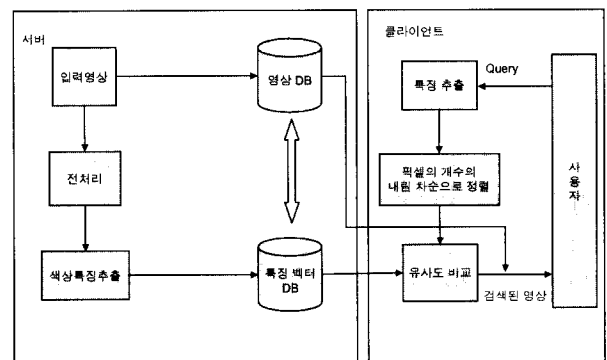
$$H(I, M) = \sum_{j=1}^n \min(I_j, M_j) \tag{1}$$

여기에서 I, M은 각각 질의 영상과 목표 영상을 나타내며, j는 히스토그램의 빈, n은 해당 영상의 총 빈의 수를 나타낸다. 위 식 (1)은 동일한 영상 크기일 경우이고, 질의 영상과 목표 영상의 크기가 다를 경우에는 다음 식 (2)와 같이 0과 1의 값을 갖도록 정규화 하여 나타낸다. 만일 질의 영상과 목표 영상이 동일하다면 식 (2)의 결과값은 1이 되고, 두 영상의 컬러가 전혀 다른 값들을 갖는다면 0이 된다. 그러나 이 방법은 조명의 변화와 영상 내의 물체의 크기에 민감하다는 단점이 있다. 또한 영상의 공간 정보를 이용하지 않고, 단순한 컬러의 빈도값만을 특징값으로 사용하는 히스토그램의 최대 단점을 극복하지 못함으로써 잘못된 영상 검색 결과를 초래할 수 있다.

$$H(I, M) = \sum_{j=1}^n \min(I_j, M_j) / \sum_{j=1}^n M_j \tag{2}$$

3. 제안 방법

본 논문에서 제안하는 방법은 기존의 r, g, b 값만 가지고 컬러 특징을 표현했던 방법[6]을 보완하는 방식으로 컬러 특징으로 r, g, b, p 값을 사용한다. 컬러 특징을 각 영역의 r, g, b 분포 특성을 고려하고, 배경의 의존적이지 않게 하기 위해 영상의 중앙 영역만을 특징 추출의 대상으로 삼았다. 또한 기존의 16개 빈으로 하여 특징 추출한 방식을 최적의 빈으로 비교하도록 하였다. 전체 시스템 구성도는 (그림 1)과 같다.

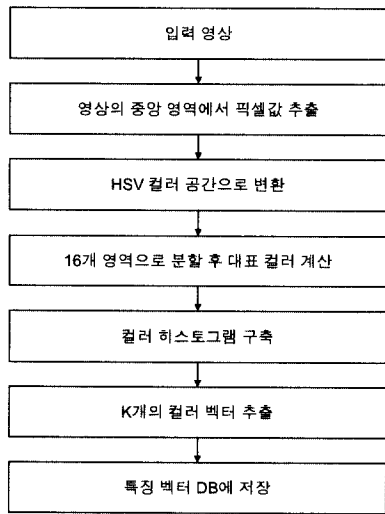


(그림 1) 전체 시스템 구성도

3.1 컬러 특징 추출 처리 기법

컬러는 인간이 물체를 인지하고 분별하는 가장 중요한 특징이다. 따라서 내용기반 영상 검색 시스템에서도 컬러 특징은 가장 중요한 수단이 된다. 또한 컬러 특징은 다른 저수준 특징들보다 특징 추출과 검색이 상대적으로 용이하다. 이러한 컬러 특징을 가장 잘 표현하는 방법이 컬러 히스토

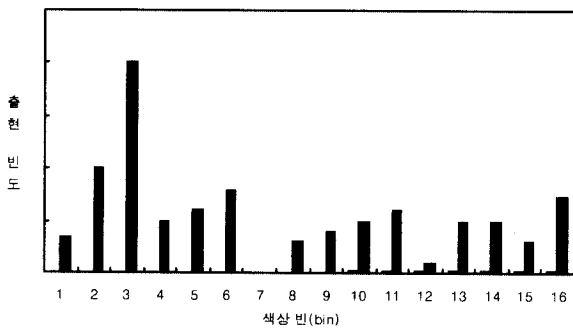
그램이다. 기존의 컬러 히스토그램 방식은 해당 컬러의 r, g, b 값이나 hue 값과 같은 컬러 성분을 이용하지 않고, 단순히 같은 컬러의 빈도값만을 이용한다. 이로 인해 컬러 히스토그램을 구성하기 위해 컬러 양자화를 수행함으로써 발생하는 양자화 에러를 고스란히 흡수하게 된다. 따라서 본 논문에서는 이러한 단점을 극복하기 위해 컬러값과 빈도수를 모두 이용하는 방법을 제안한다. 이를 위해 본 절에서는 컬러값과 빈도수 모두를 컬러 특징으로 추출하는 새로운 방법에 대해 기술한다. (그림 2)는 데이터베이스에 들어가는 특징 벡터의 처리 과정을 보여주고 있다.



(그림 2) 특징 벡터 DB 구성 흐름도

영역	R	G	B	빈도
1	255	0	0	340
2	30	80	180	1000
3	55	0	90	2000
..
..
..
16	0	255	33	730

(a) 인덱스 테이블



(b) 빈에 따른 색상 히스토그램

(그림 3) 컬러 히스토그램

본 논문에서는 HSV 컬러 공간[7]을 사용하였고, 유사한 컬러를 포함하는 영상 내의 영역을 분리하기 위해 균일 양자화를 적용하였다. 변환된 HSV 컬러 공간에서 H, S, V 값 각각을 4, 2, 2개의 값으로 양자화하여 16개의 컬러 영역으로 분리하였다.

영상의 컬러 특징을 추출하기 위해 HSV 컬러 공간 영역을 이용하여 16개의 영역들로 분리한 후, 각 영역의 대표 컬러와 빈도수를 특징으로 추출하였다. 각 영역의 대표 컬러는 각 영역의 r, g, b 평균값을 취하였다. 이를 이용해 컬러 히스토그램을 구한 것이 (그림 3)이다.

영상에서 추출한 컬러 특징은 다음 식 (3)과 같은 벡터 형식으로 표현된다.

$$\text{color feature} = [r_i, g_i, b_i, p_i], (i = 1, 2, \dots, M) \quad (3)$$

여기서 M은 컬러의 총 수를 나타내고 p_i는 해당 r, g, b 값을 갖는 컬러의 비율값으로 $\sum_{i=1}^M p_i = 1$ 을 만족한다. 한 영상에서 추출한 컬러 특징 벡터의 크기는 M × 4byte로 작은 공간을 차지한다. 컬러는 r, g, b, p로 하였고 K는 최적 빈으로써 실험에 의해 8로 정하였다.

3.2 중앙 영역의 특징 추출

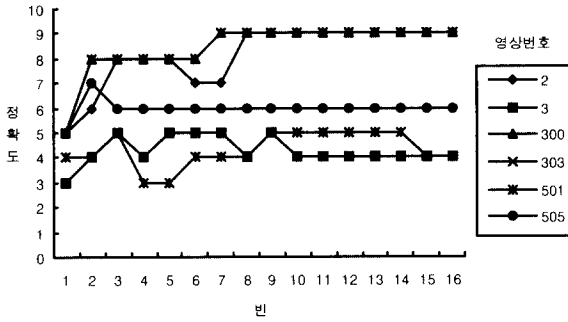
데이터베이스에 영상 정보가 들어갈 때 각기 다른 크기를 가진 영상을 일률적으로 처리하는 것이 쉽지 않기 때문에 영상의 제한된 일정 부분만을 가지고 특징을 추출하게 된다. 사람의 주관적인 유사성 판단 기준이 영상 내에 존재하는 특정한 객체의 존재 유무와 그 객체의 유사성에 크게 의존한다[8]. 이에 대부분의 영상들은 중심 객체가 중앙에 있다는 가정을 하고 중앙 영역의 픽셀 값만을 특징으로 추출하였다.

(그림 4)에서 보는 것과 같이 영상에서 중앙 픽셀을 찾는다. 그 중앙 픽셀의 좌표값을 (x, y)라고 놓았을 때 (x - N, y - N) 위치의 픽셀을 시작점으로 해서 N × N 크기의 영상부분만을 가지고 특징을 추출한다.

(그림 4) 중앙 영역 특징 추출

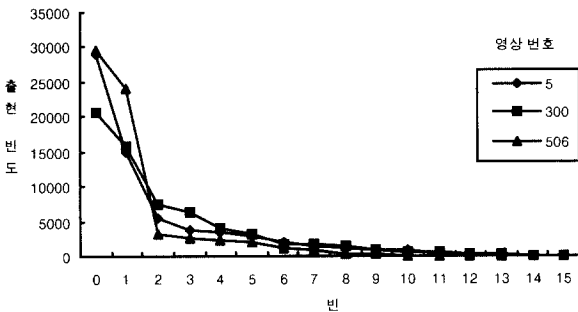
3.3 비교 빈(bin) 수 최적화

영상을 16개 레벨로 양자화를 하여 히스토그램을 작성하였는데 이 16개 빈을 모두 비교하여 유사성을 계산하는 것은 시간도 오래 걸리고 효율적인 면에서도 좋지 않다 그래서 본 논문에서는 실험 결과를 바탕으로 일정한 임계값으로 비교 빈 수를 줄이려 한다.



(그림 5) 비교 빈(bin) 수에 따른 검색 결과 꺾은 선 그래프

(그림 5)는 세 개의 그룹(아프리카, 버스, 코끼리)에서 각각 두 개의 영상을 뽑아 비교 빈 수를 1에서 16까지 바꾸며 검색 결과를 나타낸 꺾은선 그래프이다. (그림 5)에서 보듯이 16개를 모두 비교한 결과와 더 적은 수의 빈으로 비교한 결과의 차이가 많이나지 않는 것을 알 수 있다. 어떤 경우에는 적은 수의 빈으로 비교 했을 때 더 좋은 결과를 보이기도 한다.



(그림 6) 빈에 따른 픽셀의 분포

(그림 6)은 세 개의 그룹에서 뽑은 임의의 영상들의 빈에 따른 분포를 보여주는 꺾은선 그래프이다. 그래프에서 보는 것과 같이 상위 8개의 빈에 속해있는 픽셀의 수가 전체 픽셀의 95%를 차지한다. 그렇기 때문에 상위 8개 빈만 가지고 비교를 수행해도 영상의 지배적인 컬러 특징을 모두 고려해서 비교했다고 볼 수 있다.

4. 실험 결과

본 논문은 각각 100장씩 10개의 다른 주제를 갖는 총 1,000장의 자연 영상[9]으로 제안된 방법의 성능을 평가하였다. 본

논문에서는 기존의 컬러 히스토그램의 유사도 측정 방식의 단점인 해당 컬러의 값을 이용하지 않음으로서 생기는 인식을 저하의 단점을 해결하고자 퍼지 컬러 거리 함수를 사용해서 유사도를 계산하였다. 아래 식 (4)은 이 퍼지 컬러 거리 함수를 나타내는 식이다.

$$D^1(q, t) = \sum_{m=0}^{M-1} |h_q(m) - h_t(m)| \tag{4}$$

먼저 제안한 컬러 특징 추출 방법을 이용해 컬러 히스토그램을 구한다. 컬러 히스토그램은 각 컬러의 RGB 값과 빈도수를 모두 가진다. 본 논문에서는 이러한 컬러의 RGB 값과 빈도수를 모두 이용하는 퍼지 컬러 함수를 이용해 질의 영상과 대상 영상의 유사도를 측정한다. 질의 영상과 대상 영상을 각각 I_1, I_2 라고 하면, 각 영상의 컬러 특징은 $[r, g, b, p]$ 와 같이 구성된다. 이때 두 영상의 컬러간의 유사도는 식 (5)와 같은 멤버쉽 함수를 사용하여 구해진다. 즉 두 컬러가 유사하다면 식 (5)는 1에 가까운 값을 갖고, 유사하지 않다면 0의 값을 갖는다.

$$H(I, H) = \sum_{j=1}^n \min(I_j, M_j) \tag{5}$$

이러한 두 컬러값 간의 유사도를 사용하여 변형된 D^2 거리 함수인 식 (6)을 이용해 두 영상간의 유사도를 구한다.

$$D^2(I_1, I_2) = \sum_{i=1}^B p_{1i}^2 + \sum_{j=1}^B p_{2j}^2 - \sum_{i=1}^B \sum_{j=1}^B 2\mu(I_1, I_2) p_{1i} p_{2j} \tag{6}$$

여기에서 I_1 은 질의 영상의 컬러 특징이고, I_2 는 대상영상의 컬러 특징, B는 컬러 수를 나타낸다.

(그림 7)은 검색 결과를 보여준다. (그림 7)(a)를 질의 영상으로 주었고 질의 영상과 유사한 순서로 10개의 영상을 보여준다. 질의로 준 영상은 버스 영상으로 이와 같은 주제로 묶여있는 영상들은 보통 도시를 배경으로 버스가 전면에서 사선 방향으로 나와 있는 영상이다. 버스의 색은 보통 파란색, 녹색, 노란색, 빨간색등으로 이루어져 있다. 질의로 주어진 영상은 하늘과 도로가 보이고, 뒤에 나무를 배경으로 흰색과 빨간색으로 구성된 버스가 중앙에 놓여있는 그림이다. (그림 7)에서 보이는 10개의 검색 결과 영상 역시 질의 영상과 같이 빨간 버스가 중앙에 놓여있는 영상이 검색되었음을 볼 수 있다.

<표 1>은 기존에 제안되었던 히스토그램 방법과 영상의 크기를 256×256으로 정규화 한 후 영상 전체를 특징으로 추출한 방법, 제안한 방법의 정확도를 비교한 것이다. 결과에서 볼 수 있듯이 기존에 제안된 히스토그램 방법보다 r, g, b 값을 특징값으로 하고 영상 전체의 픽셀값을 특징으로 사용한 방법[10]이 더 좋은 정확도를 보이고 있고, 이 방법보다는 제안한 방법이 더 좋은 정확도를 보이고 있음을 알 수 있다.

(그림 7) (a)영상을 질의하였을 때 검색 영상

<표 1> 제안 방법의 정확도

영상 \ 방법	기 존	영상 전체 특징	제안방법
버 스	74.51	80.32	85.48
코 끼 리	50.96	53.54	58.06
아프리카	62.90	68.38	73.33

5. 결 론

본 논문에서는 영상의 중앙 영역의 픽셀만으로 컬러 특징으로 추출하여 데이터베이스를 구성하고, 영상검색을 할 때 양자화된 모든 빈의 값을 비교하여 유사도를 계산하지 않고, 일정한 개수의 빈만 가지고 계산하여 유사 영상을 검색하는 내용기반 영상검색 방법을 제안하였다.

단순화된 배경이 중심 객체보다 더 많이 차지할 경우 배경색이 영상 특징에 많은 비중을 차지하게 된다. 유사도를 계산할 때에도 이렇게 배경이 단순한 그림의 전체를 특징값으로 추출했을 경우 검색 우선순위에 배경색이 큰 영향을 미친다. 그렇기 때문에 대부분의 영상의 중심 객체는 중앙에 놓여있다고 가정하고 영상의 중앙 영역의 컬러 특징만을 영상의 특징으로 추출하였다. 실험 결과에서도 영상을 일정한 크기로 정규화하는 전처리를 걸친 후 영상 전체의 픽셀값으로 특징을 추출한 경우보다 중앙 영역의 픽셀 값만으로 특징을 추출한 방법이 더 좋은 결과를 나타냄을 알 수 있었다.

영상의 컬러 특징을 추출할 때 영상을 16개 빈으로 나누어 양자화를 하였다. 그런데 유사도를 계산할 때 16개 빈 모두를 비교하며 유사도를 검사하면 효율이 떨어진다. 그래서 일정한 상위 빈만 가지고 유사도를 계산해 검색을 수행하였다. 상위 8개 빈의 누적합계를 계산했을 때 상위 8개 빈에 속한 픽셀이 전체 영상의 픽셀에서 차지하는 비중이 95%에 가까웠다. 상위 8개 빈만 비교하여 유사도 계산을 통해 검색한 결과가 모든 빈을 비교하여 검색했을 때보다 더 좋은 결과를 보인다. 추후 과제는 칼라 특징뿐만 아니라 모양

정보를 추가하여 웹 쇼핑몰에 적용해 볼 수 있을 것이다.

참 고 문 헌

[1] Y. Rui, T. S. Hang and S. Fu Chang, "Image retrieval: Current technique, promising directions, and open issues," *Journal of Visual Communication and Image Representation*, Vol.10, No.4, pp.39-62, Apr., 1999.

[2] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, Vol.7, No.1, pp.11-32, Nov., 1991.

[3] M. Stricker and M. Orengo, "Similarity of color images," in *Proceedings SPIE Storage and Retrieval for Image and Video Databases*, Vol.2420, pp.381-392, Feb., 1995.

[4] M. Stricker and A. Dimai, "Color indexing with weak spatial constraints," in *Proceedings SPIE Storage and Retrieval for Image and Video Databases*, Vol.2670, pp.29-40, Feb., 1996.

[5] B. V. Funt and G. D. Finlayson, "Color constant color indexing," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.17, No.5, pp.522-529, May, 1995.

[6] Yanmei Liang, Hongchen Zhai, Pierre Chavel "Fuzzy color-image retrieval," *Optics Communications*, pp.247-250, Nov., 2002.

[7] 고성제, 김재원, *디지털 영상처리*, 대영사, 2002.

[8] Hidenori Yamamoto, Hedeiko Iwasa, Naokazu Yokoya, Haruo Takemura, "Content-Based Similarity Retrieval of Images Based on Spatial Color Distribution," *ICLAP99*, pp. 951-956, 1999.

[9] James Z. Wang, Jia Li and Gio Wiederhold, "SIMPLiCity : Semantics-Sensitive Integrated Matching for Picture Libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.23, No.9, pp.947-963, 2001.

[10] 박원배, "내용기반 영상 검색을 위한 시각 특징 추출 기법", 충북대학교 박사학위논문, 2004.

류 은 주

e-mail : phoenixej@feechal.com
2003년 충북대학교 정보통신공학과
(공학석사)
2003년~현재 충북대학교 정보통신공학과
(석사과정)
관심분야 : CBIR

송 영 준

e-mail : songyjorg@dreamwiz.com
1996년 충북대학교 정보통신공학과
(공학석사)
1996년~1998년 LG전자 멀티미디어 사업
본부 주임연구원
1998년~2000년 LG반도체 메모리사업
본부 주임연구원
2000년~2003년 한국전자통신연구원 네트워크연구소 홈네트워
크팀 선임연구원
2000년~현재 충북대학교 정보통신공학과(박사과정)
관심분야 : 영상인식, 영상처리, 얼굴 인식

박 원 배

e-mail : nte92@dreamwiz.com
1998년 충북대학교 정보통신공학과
(공학석사)
2004년 충북대학교 정보통신공학과
(공학박사)
2004년~현재 LG전자 CDMA 단말 사업부
선임연구원
관심분야 : CBIR, Wavelet

안 재 형

e-mail : jhahn@cbucc.chungbuk.ac.kr
1981년 충북대학교 전기공학과(학사)
1983년 한국과학기술원 전기및전자공학과
(석사)
1992년 한국과학기술원 전기및전자공학과
(박사)
1987년~현재 충북대학교 전기전자공학부 교수
관심분야 : 영상 통신 및 영상 정보 처리, 멀티미디어 제작 및
정보제공, 인터넷 통신 및 프로그래밍