

# 웨이브릿 변환 영역에서 특징추출을 이용한 내용기반 영상 검색

최인호<sup>†</sup> · 이상훈<sup>\*\*</sup>

## 요 약

본 논문에서는 웨이브릿 변환 영역에서 추출된 특징을 기반으로 한 내용기반 영상검색 방법에 관해 연구하였다. 기존의 웨이브릿 기반의 방법에서의 문제점인 특징벡터의 크기를 줄이기 위해 웨이브릿 계수의 영역별 에너지 값을 이용하였으며, 대상물의 이동, 회전, 크기 변화에 영향을 받지 않는 모멘트 특성을 이용한 검색방법을 제안하였다. 본 방법은 특징벡터의 크기를 줄이고, 기존의 특징벡터와 비교해서 검색시간을 단축하면서 분류검색의 효율성을 향상시켰다. 영역기반 영상검색 기능을 제공하기 위해 영상분할 방법에 대해 연구하였으며, 불규칙한 광원에 의한 영향을 최소화할 수 있는 영상분할 방법을 제안하였다. 영상분할은 영역병합을 이용하였고, 병합 후보영역은 웨이브릿 변환의 고주파 대역 에너지 값을 이용하여 선정하였다. 분할된 영역정보를 이용하여 칼라와 질감, 모양 특징벡터를 구성하여 영역기반 영상검색을 수행하였다.

## Content-based Image Retrieval using Feature Extraction in Wavelet Transform Domain

In Ho Choi<sup>†</sup> and Sang Hun Lee<sup>\*\*</sup>

## ABSTRACT

In this paper, we present a content-based image retrieval method which is based on the feature extraction in the wavelet transform domain. In order to overcome the drawbacks of the feature vector making up methods which use the global wavelet coefficients in subbands, we utilize the energy value of wavelet coefficients, and the shape-based retrieval of objects is processed by moment which is invariant in translation, scaling, rotation of the objects.

The proposed methods reduce feature vector size, and make progress performance of classification retrieval which provides fast retrievals times. To offer the abilities of region-based image retrieval, we discussed the image segmentation method which can reduce the effect of an irregular light sources. The image segmentation method uses a region-merging, and candidate regions which are merged were selected by the energy values of high frequency bands in discrete wavelet transform. The region-based image retrieval is executed by using the segmented region information, and the images are retrieved by a color, texture, shape feature vector.

**Key words:** 내용기반검색, 웨이브릿 변환, 영상분할, 영역기반 영상검색

## 1. 서 론

최근 정보통신망 및 멀티미디어 기술의 발전으로 음성이나 정지영상, 동영상 등과 같은 대용량의 멀티미디어 데이터를 효율적으로 저장, 관리, 검색할 수

있는 방법들에 대한 요구가 증가하고 있다. 현재 사용되고 있는 멀티미디어 데이터 분류 및 검색 방법은 주석 기반의 검색방법(annotation-based retrieval technique)으로 사람이 직접 멀티미디어 데이터에 문자 키워드(keyword)를 부여하여 분류와 검색을 수행한다. 이러한 방법은 사용자가 직접 멀티미디어 데이터에 대하여 의미 정보(semantic information)

<sup>†</sup> 정회원, 김포대학 전자정보계열 조교수

<sup>\*\*</sup> 부산정보대학 정보통신계열 조교수

를 기술하는 것으로써 의미 정보에 의한 효율적인 검색이 가능하다는 장점을 가지고 있으나, 대용량의 데이터에 대하여 사람이 일일이 키워드를 부여해야 한다는 점과 동일한 영상인 경우에도 사람의 주관적인 판단에 의해 키워드가 다르게 부여될 수 있는 문제점이 있게 된다.

이러한 문제점을 해결하기 위해 멀티미디어 데이터에 대해 내용(content)으로 정의되는 특징을 자동으로 추출하여, 이를 기반으로 멀티미디어 정보를 검색하는 내용기반 검색 방법(content-based image retrieval)에 대한 연구가 활발히 진행 중에 있다 [1,2,10,17,19,20].

내용기반 영상 검색은 질의영상(query image)과 완전히 일치하는 영상을 데이터베이스로부터 검색하는 것이 아니라, 어느 정도의 오차를 허용하는 유사 일치(approximate matching) 검색을 사용한다.

질의에 사용되는 내용은 특징 추출(feature extraction) 방법에 의해서 얻어지며, 이러한 특징은 주로 칼라(color)나 질감(texture), 형태(shape) 등을 이용하고 있다. 특징 추출은 크게 공간영역(spacial domain)에서의 화소값(pixel value)이나 히스토그램(histogram) 등을 이용하는 방법[8,12,13,16,17]과 변환영역(transform domain)에서 주파수 계수(frequency coefficient)를 이용하는 방법[1,2,4,6,7,14,15]으로 분류되며, 현재 개발되어 있는 내용기반 영상 검색 시스템은 주로 공간영역에서의 특징추출 방법을 사용하고 있다.

공간영역에서 특징추출 방법은 알고리즘의 복잡도가 높고 특징벡터의 차원이 커서 특징추출 및 검색에 많은 시간이 소요되는 문제점이 있다. 이러한 문제점으로 인해 최근에 와서는 변환영역에서의 특징추출 방법이 연구되고 있다[1,2,4,6,7,14,15]. 대용량의 영상 데이터베이스 내 저장되는 영상들은 압축된 형태로 저장될 것이라는 점을 고려해볼 때, 별도의 복호화(decoding) 과정이 필요 없는 변환영역에서의 특징추출방법이 필요해진다. 변환영역에서의 특징추출방법은 초기에는 DCT(Discrete Cosine Transform)를 이용하였으나, 최근에는 공간-주파수 특성과 다중 해상도 특성을 갖는 웨이브릿(wavelet)을 이용한 방법[1,2,4,6,7,14,15]이 관심에 대상이 되고 있다.

본 논문에서는 웨이브릿 변환 영역에서 추출된 특징을 기반으로 영상의 전체적인 특징을 이용한 검색

과 부분적인 특징을 이용한 영역기반 영상검색(region-based image retrieval)에 관해 연구하였다. 기존의 웨이브릿 기반의 방법에서의 문제점인 특징벡터의 크기를 줄이기 위해 웨이브릿 계수의 영역별 에너지(energy) 값을 이용하였으며, 대상물(object)의 이동, 회전, 크기 변화에 영향을 받지 않는 모멘트(moment) 특성을 이용한 검색방법을 제안하였다.

영역기반 영상검색 기능을 제공하기 위해 칼라 영상을 대상으로 한 영상 분할 방법(image segmentation method)과 분할된 영역 정보를 이용한 영역기반 영상검색에 관해 연구하였다. 영역기반 영상 검색은 분할된 영역정보 이용하여 수행하였으며, 질의에 사용되는 특징은 칼라와 질감, 형태를 이용하여 분할된 영역마다 각각 특징벡터를 구성하였다.

동일한 영상 데이터베이스에서 기존의 방법과 비교 분석함으로써 본 방법의 효율성을 검증하였으며, 제안된 내용 기반 검색 방법은 디지털 박물관(digital library), 위성영상 데이터베이스(satellite image database), 의료영상 데이터베이스(medical image database) 및 기타 상업적인 영상 데이터베이스 검색시스템과 같은 다양한 멀티미디어 정보 검색 시스템에서 효율적이고 중요한 기술로 사용될 수 있다.

## 2. 웨이브릿 변환을 이용한 특징추출

고주파 대역 내 존재하는 웨이브릿 계수의 모멘트를 이용하여 특징벡터를 구성함으로써, 대상물의 이동, 회전, 크기 변화에 영향을 적게 받는 검색을 수행할 수 있다. Mandal 등의 방법[7,14,15]은 대역 내 전체 웨이브릿 계수값을 이용하여 모멘트를 산출하였으나, 이러한 방법은 영상 내 잡음성분이 존재하거나, 배경 부분의 미세한 변화에도 영향을 받는 점과 계산량이 많다는 단점이 존재하게 된다. 따라서 대역 내 웨이브릿 계수값 중 의미 있는 정보를 가지고 있는 계수들만 추출하여 특징벡터를 구성하여야 대상물 위주의 검색을 효율적으로 수행할 수 있게 된다. Mallet 등은 평활화 함수의 1, 2차 미분을 이용한 국소 최대값(local maxima)를 이용하여 윤곽정보를 추출하였다[5]. 이러한 국소 최대값은 대상물의 경계부분에 해당되는 시각적인 윤곽정보를 반영하고 있어, 대상물 위주의 검색에 유용한 정보로 사용될 수 있다.

웨이브릿 변환 영역에서 식 (1)의 조건식을 만족

할 경우, 웨이브릿 계수  $C_{(i,j)}$ 를 modulus maxima로 정의한다.

$$\begin{aligned} & \text{if } (C_{(i,j)} > C_{(i,j-1)} \text{ and } C_{(i,j)} > C_{(i,j+1)}) \text{ or} \\ & (C_{(i,j)} > C_{(i-1,j)} \text{ and } C_{(i,j)} > C_{(i+1,j)}) \\ & \text{then } C_{(i,j)} = C_{\max} \quad (C_{\max} : \text{wavelet modulus maxima}) \end{aligned} \quad (1)$$

윤곽정보의 웨이브릿 계수들의 모멘트를 이용하여 유사한 영상들을 검색함으로써, 이동, 회전, 크기에 영향을 받지 않는 질의를 수행할 수 있게 된다. 식 (2)는 wavelet maxima의 2차 중앙 모멘트를 나타낸다.

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (u-u')^p (v-v')^q \omega(u,v) du dv$$

x  
( $\omega(u,v)$ : 웨이브릿 maxima,  $u,v$ : 주파수영역 신호 값,  $p,q=0,1,2,\dots$ ) (2)

### 3. 칼라와 모멘트를 이용한 전역검색

Jacobs 등이 제안한 방법[4]에서 특징벡터 추출에 처음으로 웨이브릿 변환을 이용하였고, 칼라영상의 R, G, B 각 채널마다 따로 특징벡터를 구성하여 질의를 수행하였다. 이와 유사한 방법으로 Wang 등은 Daubeche-8 웨이브릿을 이용하여 웨이브릿 변환을 수행한 후, level 4의 대역별 웨이브릿 계수들을 특징벡터로 구성하였다[2]. 위의 두 방법은 웨이브릿 계수의 직접적인 비교로 영상간의 유사도를 측정하는 방법이며, 영상의 전체적인 칼라 분포를 이용한 유사도 측정에는 적합하나, 대상물의 형태를 이용한 경우에는 부적합하다. 또한, 특징벡터의 차원이 너무 커서 검색시간이 많이 소요되는 단점이 있다. 이러한 문제점으로 인해 Albuz등이 제안한 방법에서는 특징벡터의 크기를 줄이기 위해 웨이브릿 변환 계수를 직접 이용하지 않고, 대역별 웨이브릿 계수들의 에너지 값을 이용하였다[1]. 이 에너지 값을 Swain이 제안한 히스토그램 방법을 이용하여 나타냄으로써, 한 영상의 특징 벡터 크기를 흑백영상의 경우 10 bytes로 줄여 고속검색이 가능하게 하였으나, 각 부 대역의 특징벡터를 1개의 에너지 값으로만 표현함으로써, 질의능력이 떨어지는 문제점이 있다. Mandal 등이 제안한 방법[7]은 웨이브릿 변환영역에서 히스토그램과 모멘트를 이용하였다. 이 방법에서는 대역 내

전체 웨이브릿 계수를 이용한 모멘트를 이용함으로써, 잡음이나 배경부분의 변화 등에 영향을 받게되는 문제점이 있게된다.

제안한 방법에서는 먼저, 질의 영상의 칼라 공간을 HSI 칼라 공간으로 변환한 후, 각 채널에 대해 2D DWT를 수행한다. 특징벡터의 구성은 기존의 방법에서 문제점인 특징벡터 크기를 줄이기 위해서 웨이브릿 계수들을 직접 이용하지 않고, 에너지 값을 이용한다. 일반적인 사진영상의 경우 영상 내 대상물들은 사진의 중앙에 위치하는 경우가 많고, 카메라를 통해 영상이 획득될 때 대상물에 포커스가 맞추어지는 원리를 이용하여 영상을 5개의 고정영역으로 분할한다. 그림 1(a)는 칼라에 의한 특징추출 및 유사도 측정과정을 나타낸다.

각 영역의 웨이브릿 계수의 에너지를 HSI 세 채널에 대해 산출한다. 이때 특징벡터의 크기는 5(영역 수)×3(채널 수)×4bytes = 60bytes가 된다. 모멘트에 의한 질의를 위해 웨이브릿 변환영역에서 3개의 고주파 대역(LH, HL, HH)에서 wavelet maxima를 이용한 2차 중앙 모멘트 값을 HSI 각 채널에 대해

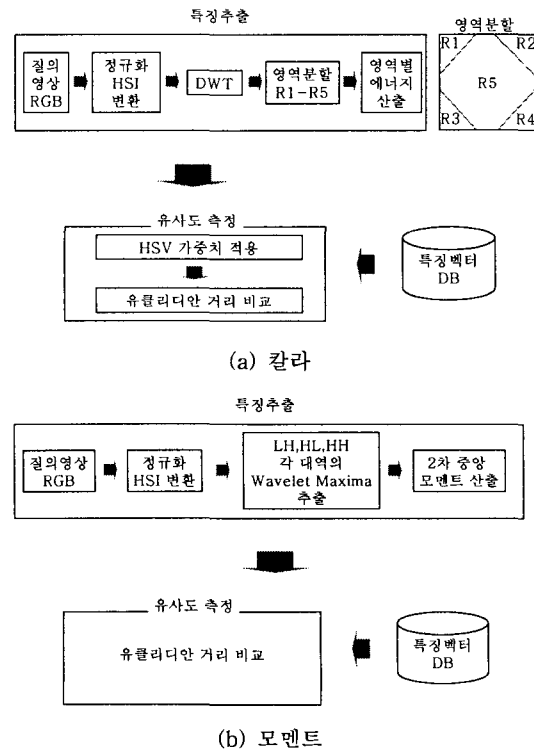


그림 1. 칼라와 모멘트에 의한 특징추출 및 유사도 측정

산출하여 특징 벡터로 이용한다. 이때 특징벡터의 크기는 3(대역 수)×4 bytes = 12bytes가 된다. 그림 2-1 (b)는 모멘트에 의한 특징추출 및 유사도 측정과정을 나타낸다.

#### 4. 영역기반 영상검색

##### 4.1 칼라영상의 잡음제거 및 평활화

영상 분할시 잡음(noise)에 의한 영향을 줄이고 양자화 과정에서 부호화 비트(bit)수를 감소시키기 위해 원 영상에 대해 잡음 제거와 평활화(smoothing)를 수행한다. 영상 내 존재하는 임펄스성 잡음(impulse noise)을 제거하기 위한 방법으로 선형 평활화(linear smoothing) 방법과 비선형 평활화(non-linear smoothing) 방법이 사용되고 있다. 선형 평활화 방법은 알고리즘상의 연산자가 지역 윈도우 내 모든 화소들에게 균등(uniform)하게 적용됨으로 해서 잡음에 의해 손상되지 않은 화소들을 변형시키는 문제점을 가진다. 이러한 단점을 보완하기 위해서 비선형 평활화 방법들이 사용되고 있다. 대표적인 비선형 평활화 방법으로는 중간값 필터링(median filtering)과 방향성 중간값 필터링(directional median filtering) 등이 사용되고 있다.

Yining 등은 Fisher의 discriminant estimation을 응용한 peer group filtering을 이용하여 지역 윈도우 내 화소들을 두 그룹으로 분류한 후 유사도가 높은 그룹의 화소들에 연산자를 적용하는 방법을 이용하여 영상의 국부적인 특성에 따라 잡음제거와 평활화를 수행하는 방법을 제안하였다[18].

그러나 이러한 방법은 원 영상의 RGB 값을 256 단계로 양자화한 후, 필터링을 수행하기 때문에 원 영상의 칼라 정보가 변형될 수 있는 문제점이 있다. 이러한 칼라정보의 변형을 줄이기 위해서는 원영상의 RGB 값을 이용한 필터링 방법이 필요하나, 이 경우 R, G, B 각각의 채널에 대해 연산을 수행해야 하기 때문에 계산량이 증가하는 문제점이 있게 된다.

칼라는 특징벡터 구성에서 중요한 정보로 사용되며, 이러한 칼라정보의 변형을 줄이기 위해 양자화 전 단계에서 RGB값의 칼라 거리를 이용한 peer group filtering을 방법을 사용한다.

$x_0(k)$ 을 영상 내 위치  $n$ 에서  $w \times w$  윈도우의 중심 화소로 둘 때, 윈도우 내 모든 화소들은  $x_0(k)$ 과의 칼라 거리에 따라 오름차순으로 정렬되며, 식 (3)과 같이 정의된다.

$$x_i(k), i=0, \dots, n, n = w^2 - 1 \quad (3)$$

거리 값은 각 화소간에 RGB 값의 유클리디안 거리를 사용하며 식 (4)와 같다.

$$d_i(k) = \|x_{R_0}(k) - x_{R_i}(k)\| + \|x_{G_0}(k) - x_{G_i}(k)\| + \|x_{B_0}(k) - x_{B_i}(k)\|$$

$$d_0(k) \leq d_1(k) \leq \dots \leq d_n(k), i=0, \dots, n \quad (4)$$

크기  $m(n)$ 인  $x_0(n)$ 의 peer group  $P(n)$ 은 식 (5)와 같다.

$$P(n) = \{ x_i(n), i=0, \dots, m(n)-1 \} \quad (5)$$

$P(n)$ 은  $x_0(n)$ 과 칼라 성분이 유사한 이웃하는 화소들의 집합으로 정의할 수 있다. Peer group에 해당

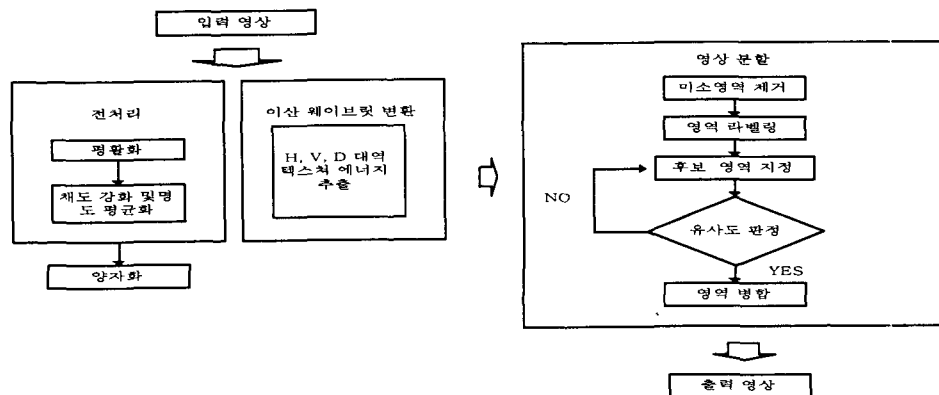


그림 2. 영역기반검색 구성도

되는 화소들은  $w \times w$  윈도우 내 화소들에 의해 지역적으로 결정되며, 그룹의 크기  $m$ 은 식 (6)에 의해서 자동적으로 결정된다.

중심 화소  $x_0(n)$ 을 지역 윈도우 내 모든 화소들의 평균값으로 대체하지 않고 peer group에 소속되는 화소들의 평균값을 이용함으로써, 윤곽성분이 블러링(blurring)되는 것을 최소화할 수 있다. 만약, 그룹 크기  $m$ 이 주어진 임계치보다 작을 경우,  $x_0(n)$ 은 임펄스성 잡음으로 간주되며, 지역 윈도우 내 모든 화소들의 평균값으로 대체된다.

$$\begin{aligned}
 J(i) &= \frac{|a_1(i) - a_2(i)|^2}{s_1^2(i) + s_2^2(i)} \quad i = 1, \dots, k \\
 a_1(i) &= \frac{1}{i} \sum_{j=0}^{i-1} d_j(n), \quad a_2(i) = \frac{1}{k+1-i} \sum_{j=i}^k d_j(n), \\
 s_1^2(i) &= \sum_{j=0}^{i-1} |d_j(n) - a_1(i)|^2, \\
 s_2^2(i) &= \sum_{j=i}^k |d_j(n) - a_2(i)|^2 \\
 m(n) &= \max(J(i)) \quad (6) \\
 (J(i) : \text{peer group 크기}, a_1(i), a_2(i) : \text{지역평균}, \\
 s_1^2(i), s_2^2(i) : \text{지역평균과의 차}, m(n) : \text{peer group 크기의 최대값})
 \end{aligned}$$

#### 4.2 HSI 좌표계를 이용한 칼라 근사화

제한된 방법에서의 칼라 근사화는 칼라 양자화 단계에서 양자화 비트수를 최소한으로 줄이기 위해 수행된다. 이를 위해서는 영상 내 존재하는 비균질성을 줄여야 한다. 일반적으로 물체의 색은 빛의 반사, 투과 및 흡수 등에 의해 인식되며, 물체색 중에서도 빛의 반사에 의해 나타나는 색을 표면색(surface color)라 정의한다. 이러한 표면색은 물체를 비추는 빛의 종류에 따라 반사되는 빛의 성질에 의하여 크게 달라진다.

인간은 물체에서 반사되는 빛 에너지에 기반을 두어 주위 광원의 효과를 배제하여 물체의 고유색을 인지하며, 이를 칼라 항상성(color constancy)라 정의한다. 인간시각은 주관적인 색채 지각 능력으로 인한 칼라 항상성을 가지고 있어서 조명의 종류나 조명과 관련된 주변환경이 달라져도 물체색의 변화를 느끼지 못한다.

이와 같은 성질을 이용하면 조명이나 불규칙한 광원이 조사된 영상이라도 원래의 고유색으로 근사화

된 영상을 얻을 수 있게 된다. RGB 좌표계는 색의 세 가지 속성 중에 색상정보만을 가지고 있어 채도나 명도 해석이 힘들다. 따라서 본 논문에서는 HSI 좌표계에서 채도 강화 및 명도 평균화를 통해 칼라 근사화를 수행하였다.

채도강화는 HSI 좌표계를 이용한 칼라 항상 방법에서 칼라영상 개선을 위해 사용되었다. 채도  $S$ 값을  $\alpha$  배하여 채도 강화를 수행하며 식 (7)과 같다.

$$S = \sqrt{R(R-G) + G(G-B) + B(B-R)} \times \alpha \quad (7)$$

HSI 3차원 모델에서 중앙부분은 다른 부분에 비해 넓은 영역을 가지며, 이는 매우 어둡거나 밝은 영역보다는 인간 시각이 중간 명도값 영역에서의 채도 변화를 민감하게 느끼는 것을 의미한다.

따라서 본 논문에서는 채도 강화를 기존의 방법에서와 같이 선형적으로 수행하지 않고, 중간 명도값을 가지는 영역에서는 넓은 범위의 채도값을 허용하고, 반면에 어둡거나 밝은 영역에서는 채도변화의 폭을 감소시키는 채도 강화 방법을 사용한다.

$$S = S \times \alpha, \quad \alpha = [w_1 \ w_2 \ w_3] (\omega_1, \omega_2, \omega_3: \text{가중치}) \quad (8)$$

명도는 R, G, B 세 광원의 명도를 의미하므로 세 값의 평균으로 정의된다. 불규칙한 광원이나 조명에 의한 영향을 줄이기 위해 유사한 색상값을 가지는 영역에서는 명도값을 평균화하여 고유의 색으로 근사화시킨다.

명도가 평균화되는 색상 구간은 색상 히스토그램 상에서 색상값의 분포에 따라 구간간격이 조정된다. 명도값이 너무 크거나, 작을 경우는 HSI 좌표계에서 색상이나 채도와는 상관없이 흰색이나 검정색을 띄는 영역이므로 명도 평균화 과정에서 제외시킨다.

이때 임계치  $\beta_{white}$ ,  $\beta_{black}$ 는 명도 히스토그램 분포에 따라 설정되며 식 (9)와 같다.

$$\begin{aligned}
 \beta_{white} &= \begin{cases} 0.8 & \text{if } \text{mean}(N(v)) > m \\ 0.9 & \text{if } \text{mean}(N(v)) \leq m \end{cases} \\
 \beta_{black} &= \begin{cases} 0.2 & \text{if } \text{mean}(N(v)) > m \\ 0.1 & \text{if } \text{mean}(N(v)) \leq m \end{cases} \\
 N(v) &= \text{mean}(v_i) \\
 (m : \text{명도 빈도수의 평균값}, v_i : \text{명도 빈도수가 } m \text{ 이상인 경우}) \quad (9)
 \end{aligned}$$

#### 4.3 영역 병합에 의한 영상분할

그림 2는 제안한 방법의 영역기반검색의 전체 구

성도를 나타낸다. 칼라 양자화 후, 칼라 영역의 면적이 전체영상에 비해 0.1% 이내인 영역을 미소 영역으로 간주하고 영역확장 전단계에서 제거한다. 제거된 영역은 인접영역 중 RGB 값의 유클리디안 거리가 최소가 되는 영역으로 병합한다.

제안된 방법에서의 영역병합 방법은 화소간 병합 방법을 사용하지 않고, 분할된 칼라 영역간 병합 방법을 사용한다. 미소 영역이 제거된 초기영역들은 영역 크기순으로 라벨링(labeling)되고, 병합 후보영역을 선정한다. 후보영역 선정은 입력영상의 웨이브릿 변환 계수를 이용한다.

웨이브릿 변환의 부대역에서 LL 대역은 원 영상에 비해 해상도가 낮아진 저해상도 영상이며, LH, HL, HH의 고주파 대역은 수평, 수직방향으로 고역 통과 필터링되어 얻어진 대역으로 원영상 내 존재하는 윤곽정보가 고역 필터링 되는 방향에 따라 나뉘어서 존재하게 된다.

필터링 방향과 동일한 윤곽성분은 제거되고 직교하는 방향의 윤곽정보가 나타나게 되는데, LH 대역에서는 수평방향, HL 대역에서는 수직방향, HH 대역에서는 대각방향의 윤곽성분이 나타나게 된다. 웨이브릿 변환의 고주파 영역에서 원영상에 대한 정보를 포함하고 있는 척도로 웨이브릿 계수들의 자승합(sum of square)으로 표현되는 에너지량이 이용되며 식 (10)과 같다.

$$E = \sum_i \sum_j C(i, j) \cdot C^*(i, j) = \sum_i \sum_j C(i, j)^2$$

(E : 에너지, C(i,j) : 웨이브릿 계수값) (10)

영역 병합의 대상이 되는 후보 영역은 식 (8)에 나타난 영역 내 존재하는 웨이브릿 계수의 에너지량에 의해 선정이 되며, 이 에너지량에 의해 윤곽성분의 강도를 측정할 수 있다. 윤곽성분의 강도가 식 (11)의 주어진 조건을 만족할 경우, 병합 후보영역으로 지정이 되고 유사도 판정 기준에 의해 영역 병합이 이루어진다.

$$Energy_{H, V, D} = \sum_i \sum_j C_H(i, j)^2 + \sum_i \sum_j C_V(i, j)^2 + \sum_i \sum_j C_D(i, j)^2$$

(CH(i,j), CV(i,j), CD(i,j) : H, V, D 대역 웨이브릿 계수값) (11)

식 (12)는 후보영역 선정 조건을 나타낸다.

$$\text{if } \text{mean}(Energy_{Re_{H, V, D}}) < \text{mean}(Energy_{H, V, D}) \text{ or } \text{max}(Energy_{Re_{H, V, D}}) < \gamma \text{ then } Re_{cur} \in Re_{caa}$$

$$\gamma = \frac{\text{max}(Energy_{H, V, D})}{2}$$

( $Re_{cur}$  : 현재영역,  $Re_{caa}$  : 후보영역) (12)

영역병합은 후보영역과 인접한 영역간에 유사도 측정에 의해 이루어진다. 영역간 유사도 판정에 사용되는 파라메타는 칼라 유사도를 사용하며, 칼라 유사도는 인접 영역간 칼라 값의 유클리디안 거리를 이용한다.

#### 4.4 영역기반 질의

영역별 질의를 위한 특징벡터의 구성은 칼라, 질감, 형태를 이용하며, 그림 3은 특징벡터 구성과 영역기반 질의과정을 나타낸다.

칼라 특징추출은 영역 내 화소들의 칼라 히스토그램 상에서 빈도수가 가장 높게 나타나는 칼라를 선택함으로써 수행된다. 이때 영역간 칼라 유사도는 두 칼라간 R, G, B 성분의 유클리디안 거리를 측정하여 결정된다.

질감 특징추출은 웨이브릿 변환 계수의 국소 최대치를 나타내는 wavelet maxima의 모멘트를 이용하여 이동, 회전, 크기의 변화에 불변인 특성을 가지는 특징벡터를 구성하며, XY 평면에서 좌표쌍으로 표현한 대상물의 경계 화소들을 식 (13)의 이산푸리에 변환으로 표현한 푸리에 서술자를 형태 특징벡터로 이용한다. 유사도 측정은 질의영상의 특징벡터와 데이터베이스 내 영상들간의 특징벡터와의 비교로서 수행되며, 두 특징벡터간의 유클리디안 거리를 이용한다. 질의에 사용되는 특징은 사용자로부터 입력이 되

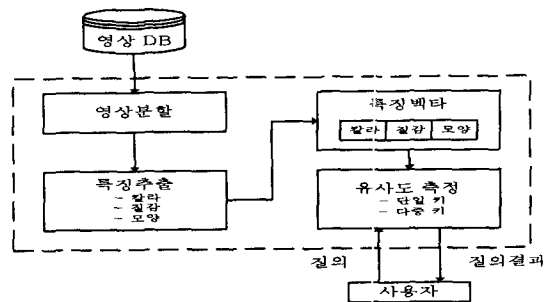


그림 3. 특징벡터 구성과 질의 수행 과정

며, 단일키 방식과 다중키 방식을 제공한다. 단일 키 방식의 경우 질의영역과의 특징벡터 크기의 차를 구하여 유사도를 판정하며, 다중키 방식의 경우 선택된 각각의 특징 벡터 크기의 차를 [0~1] 사의 구간으로 정규화하여 유사도를 측정한다.

$$d(u) = \frac{1}{N} \sum_{k=0}^{N-1} s(k) e^{-\frac{R_{\text{max}}k}{N}} \quad (13)$$

### 5. 실험 및 결과 고찰

#### 5.1 전역검색

제안한 방법의 검증을 위해 총 650개의 칼라 영상을 대상으로 Wang의 방법, Albus의 방법과 비교하였다. 실험은 Pentium II 600MHz dual CPU를 장착한 워크스테이션상에서 수행되었다. 표 1은 각 방법의 특징벡터 크기를 나타낸다. Albus의 방법은 흑백 영상을 대상으로 실험을 하였으나, 칼라영상의 경우는 세 개의 신호 채널을 고려하여 특징벡터의 크기를 결정하였다. 웨이브릿 계수의 직접 비교방식인 Wang의 방법이 Albus나 본 논문에서 제안한 방법에 비해 특징벡터의 크기가 상대적으로 크게 나타났다.

650개 영상들에 대해서 각각 질의를 수행하여, 유클리디안 거리 값에 의해 유사도가 높은 순으로 10개 영상을 추출하였다. 유사 영상 추출 결과에서 영상범주 내 동일한 그룹의 영상이 추출된 비율을 히트율(hit ratio)로 정의하며 식 (13)과 같다.

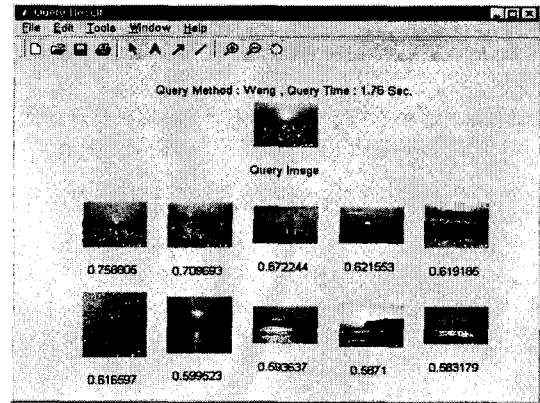
$$H = \frac{N}{K} * 100\%$$

(H: 히트율, N: 질의 결과 추출된 영상 수, K: 추출된 영상수) (13)

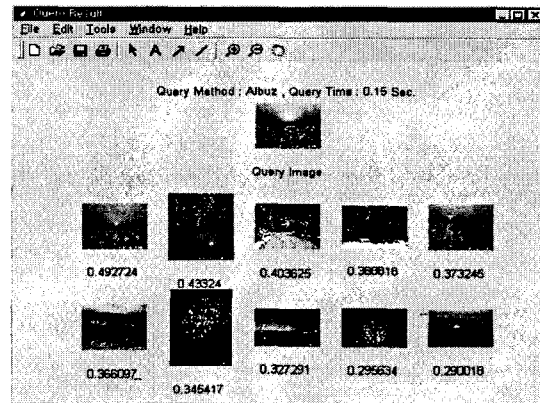
표 1. 각 방법의 비교

구분	영상 크기	특징벡터 크기	평균 히트율	질의 응답시간
Wang의 방법 WBIIS	128×128	768 Byte	77.54 %	1.25~3 (초)
Albus의 방법 UDel	512×512	30 Byte	53.35 %	0.1~0.2 (초)
제안한 방법	128×128 256×256	72 Byte	75.49 %	0.2 ~0.3 (초)

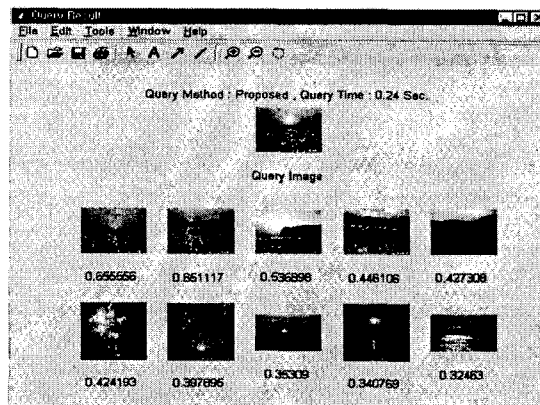
그림 4는 칼라를 이용하여 질의를 수행한 결과를 나타내고 그림 5는 히트율을 나타낸다. Wang의 방법과 제안한 방법이 Albus의 방법에 비해 우수한 결



(a) Wang의 방법



(b) Albus의 방법



(c) 제안한 방법

그림 4. 칼라에 의한 질의 결과

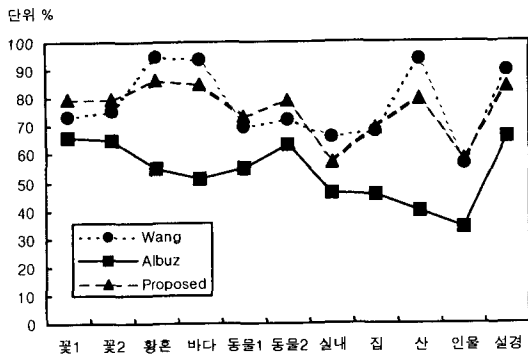


그림 5. 히트율

과를 얻었으며, 웨이브릿 계수를 직접 비교하여 유사도를 측정하는 Wang의 방법은 영상의 전체적인 칼라분포가 일정한 경치영상 등에서는 제안한 방법보다 우수한 결과를 나타내었으나, 동물영상이나 인물영상과 같은 대상물이 두르러진 영상에서는 제안한 방법보다 낮은 히트율을 나타내었다. 제안한 방법은 특징벡터의 크기를 Wang의 방법에 비해 10% 정도로 축소하면서도 질의 결과는 크게 차이가 없음을 실험을 통해 검증할 수 있었다. 그림 6은 웨이브릿 계수의 모멘트를 추출한 결과를 나타낸다. 고해상도에서는 매디안 필터링을 이용하여 잡음을 제거한 영상과 전처리 과정을 거치지 않은 영상간에 모멘트 추출 결과가 차이점을 나타내고 있으나, 저해상도 영역으로 갈수록 모멘트 추출 결과가 근접하게 나타나는 특성을 지닌다. 따라서 모멘트에 의한 검색에서는 원 영상에 대한 별도의 전처리 과정이 필요 없게 된다. 그림 7은 모멘트를 이용한 질의 결과를 나타낸다. 2차 중앙 모멘트를 이용한 결과에 비해 2차 중앙모멘트를 이용하여 산출된 wavelet maxima moment를 이용한 경우 질의결과가 우수하게 나타났다.

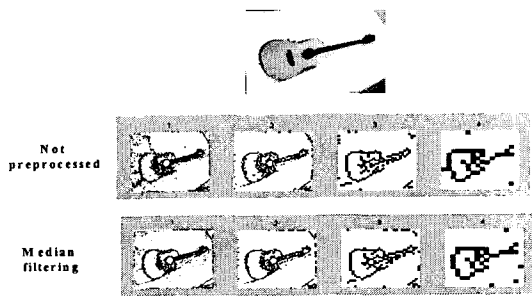
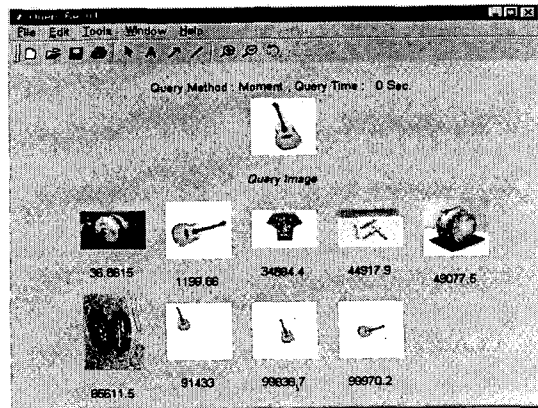
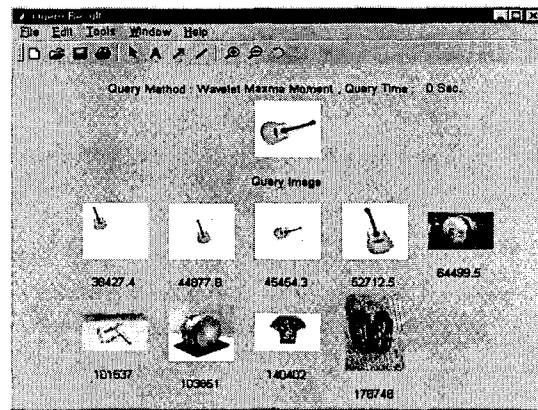


그림 6. Wavelet Maxima Moment 추출결과



(a) 중앙 모멘트를 이용한 질의결과



(b) Wavelet Maxima Moment를 이용한 질의결과

그림 7. 모멘트를 이용한 질의결과

### 5.2 영역분할

그림 8(b)는 전처리 단계에서 채도 강화와 명도 평균화를 수행한 결과를 나타낸다. 채도 강화를 통해 대상물의 대비가 커지는 효과를 얻을 수 있었으나, 데이터베이스 내 모든 영상들에 대해 동일한 가중치 적용함으로써, 원영상의 칼라정보가 소실될 수 있으며, 이를 해결하기 위한 최적의 가중치 설정이 필요하다.

채도강화 다음단계로 명도 평균화를 통해 불규칙한 광원에 의해 영향을 줄일 수가 있었다.

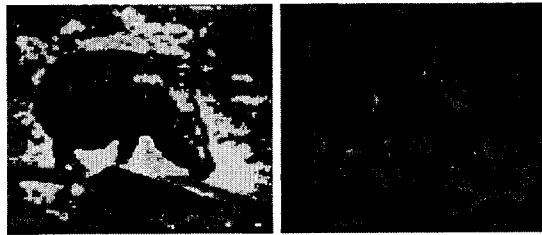
그림 8(c), (d)는 양자화 결과를 나타낸다. (c)는 칼라 근사화를 수행하지 않고 4 칼라로 양자화한 결과를 나타내며, (d)는 칼라 근사화를 수행한 후 양자화한 결과를 나타낸다.

그림 9는 웨이브릿 변환의 고주파 대역 계수의 에



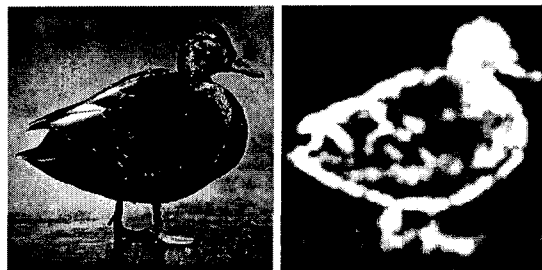


(a) 원영상 (b) 칼라 근사화

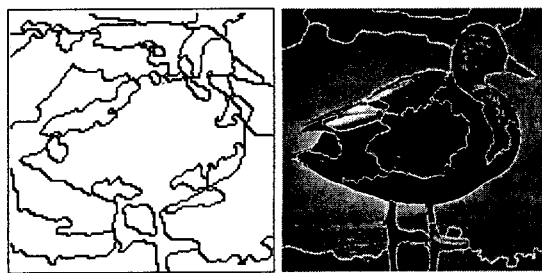


(c) 4 칼라 양자화 결과 (d) 4 칼라 양자화 결과

그림 8. 칼라 근사화 및 양자화 결과



(a) 원영상 (b) 에너지 분포



(c) 초기영역 (d) 최종영역

그림 9. 에너지 분포와 영상 분할 결과

너지 분포와 초기분할영역 및 최종분할영역을 나타낸다.

그림 10은 최종적인 영상 분할 결과를 나타낸다.

그림 11은 영역별 질의 결과를 나타낸다. (a)는 유사한 칼라 성분을 가지는 영역을 검색한 결과를 나타

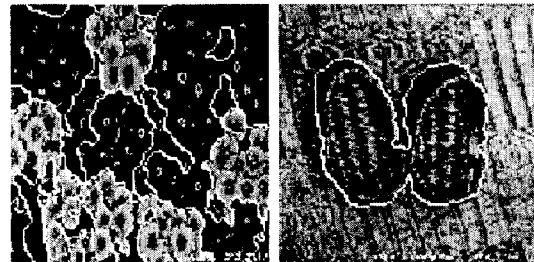
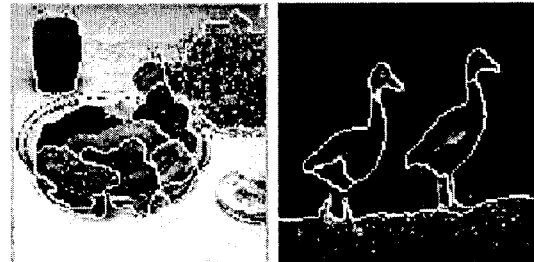
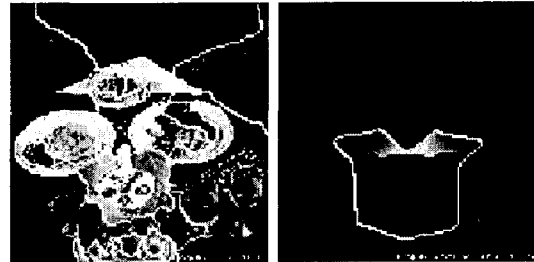
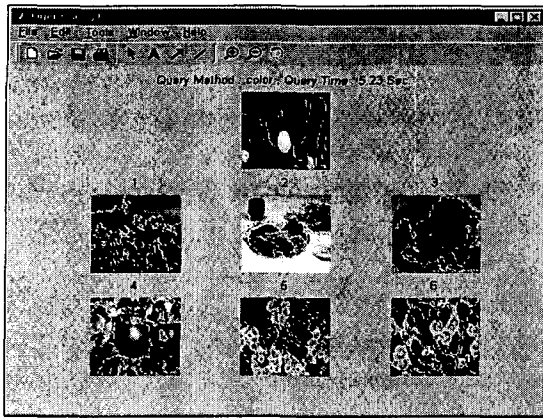


그림 10. 영상 분할 결과

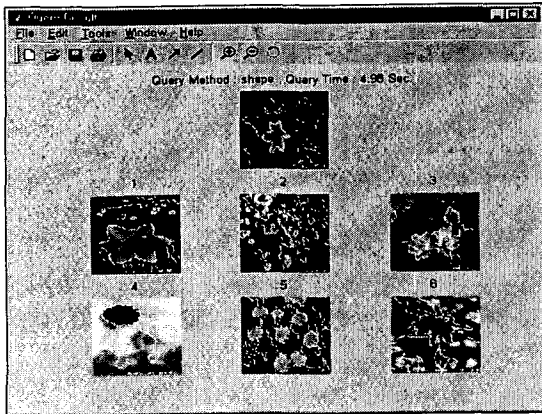
내며, (b)는 형태 특징을 이용한 질의결과를 나타낸다. (c)는 다중키 방식으로 칼라와 질감을 선택하여 질의를 수행한 결과를 나타낸다.

## 6. 결 론

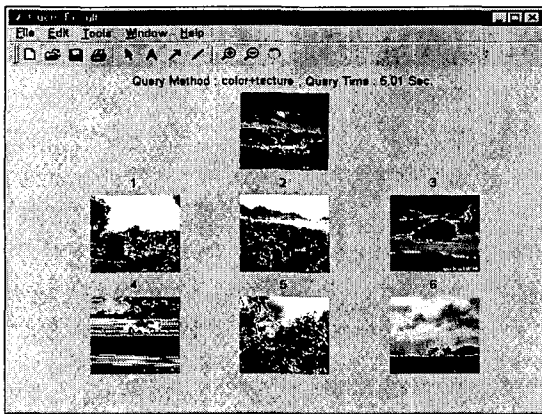
기존의 공간영역에서의 내용기반 영상검색 방법은 알고리즘의 복잡도가 높고, 계산량이 많다는 단점이 있다. 이러한 문제점을 해결하기 위해 웨이브릿 변환 영역에서 특징 추출 방법을 이용한 내용기반 영상검색 방법을 제안하였다. 질의 방법은 검색을 수행하는 사용자의 의도에 따라 영상의 전체적인 특징을 이용한 전역검색 방법과 부분적인 특징을 이용한 영역기반 검색 방법이 제공되어야 하며, 영상 데이터로부터 자동으로 추출할 수 있는 낮은 수준의 특징과 사람이 영상의 내용을 추상적으로 표현하는 논리적 특징의 효과적인 접목이 중요하다. 제안한 질의방법은 전역검색 및 영역기반 검색기능을 제공하며, 키워드를 활용한 영상검색 방법을 통해 질의의 정확도와



(a) 칼라



(b) 형태



(c) 칼라와 질감

그림 11. 질의 수행 결과

융통성을 부여할 수 있는 장점이 있다.

특징벡터를 이용한 내용기반 영상 검색에서는 검색 효율을 현격히 저하시키지 않는 범위 내에서 특징

벡터의 크기를 줄여 고속 검색이 가능하게 하여야 한다. 제안한 전역검색 방법은 웨이블릿 변환 영역에서 웨이블릿 계수들을 특징벡터로 이용하지 않고 계수들의 영역별 에너지 값을 이용함으로써, 기존의 방법 보다 특징벡터의 크기를 줄여 검색효율을 높였으며, 이동, 회전, 크기 변화에 영향을 받지 않는 wavelet maxima를 이용한 모멘트를 특징벡터로 사용하여 대상물 위주의 질의를 가능하게 하였다.

영역기반의 영상 검색을 위해서 칼라 영상 분할 방법에 대해 고찰하고, 분할된 영역정보를 이용한 내용기반 영상 검색 방법에 대해 연구를 하였다. 영상 분할 전단계에서 비선형 필터링 방법을 이용한 평활화와 채도 강화 및 명도 평균화를 수행하여, 영상 내 존재하는 비균질성을 줄일 수 있었고, 칼라 히스토그램의 zero-crossing 정보를 이용한 비균일 양자화를 수행하여 유사한 칼라성분을 가지는 영역들을 분할하였다. 영역병합을 위한 후보영역 선정은 웨이블릿 변환의 고주파 대역 에너지 값을 이용함으로써, 불필요한 영역이 분할되는 것을 줄일 수 있었다.

칼라와 질감, 형태 정보를 특징벡터로 구성하여 영역기반 질의를 수행하였고, 단일키 방식 및 다중키 방식의 질의를 제공하여 사용자가 다양한 형태로 질의를 수행할 수 있는 기능을 제공하였다.

제안된 영상 분할에 소요되는 시간은 초기 영역의 수에 따라 결정되며, 복잡한 영상일수록 처리시간이 많이 소요되는 문제점이 있다. 대상물의 형태를 이용한 질의는 영상분할 과정에서 대상물의 외곽선 검출의 정확도에 그 성능이 좌우된다. 이러한 문제점을 보완하기 위한 방안과 영역간 유사도 판정기준에 대한 고찰이 필요하며, 데이터베이스 측면에서 고속검색을 위한 색인방법 등이 향후 연구과제로 남는다.

### 참 고 문 헌

- [1] Elif Albuz, "Scalable Image Indexing and Retrieval using Wavelets", Technical Report, University of Delaware, 1998. 11.
- [2] James Ze Wang, "Content-based Image Indexing and Searching using Daubechies' Wavelet", Journal of Digital Library, 1998.
- [3] John R. Smith, "a fully automated content-based image query system", ACM Multimedia 96 Boston, 1996.

[4] Charles E. Jacobs, "Fast Multiresolution Image Query", Proceedings of the 1995 ACM SIGGRAPH, New York, 1995.

[5] S. Mallat, "Characterization of signals from multiscale edges", IEEE Trans. Pattern Anal. Machine Intell., vol. 14, pp. 710~732, 1992.

[6] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets", IEEE Trans. Info. Theory, vol. 38, pp. 617~643, 1992.

[7] M.K. Mandal, S. Panchanathan and T. Aboulnasr, "Illumination Invariant Image Index using Moment and Wavelet", Journal of Electronic Imaging, 1998. 4.

[8] Micheal J. Swain, Dana H. Ballard, "Color Indexing", IJCV, vol. 7, no. 1, pp. 11~32, 1991.

[9] Christos Faloutsos, "Fast Searching by Content in Multimedia Databases", Data Engineering, vol. 18, no. 4, 1995.

[10] W.Y. Ma and B.S. Manjunath, "NeTra: a toolbox for navigating large image databases", IEEE International Conference on Image Processing(ICIP), vol. 1, pp. 568~571, 1997.

[11] H. Kauppinen, et. al. "An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classification", IEEE Trans. Pattern Anal. Machine Intell., vol. 17, no. 2, pp. 201~207, 1995.

[12] H. Sawhney and J. Hafner, "Efficient Color Histogram Indexing", IEEE International Conference on Image Processing(ICIP), vol. II of III, pp. 66~70, 1994.

[13] Greg Pass and Ramin Zabih, "Histogram Refinement for Content-Based Image Retrieval", Third IEEE WACV(Workshop on Applications of Computer Vision), pp. 96~102, 1996. 12.

[14] M.K. Mandal, T. Aboulnasr and S. Panchanathan, "Image Indexing using Moments and Wavelets", IEEE Transactions on Consumer Electronics, vol. 42, No. 3, pp. 557~565, 1996. 8.

[15] M.K. Mandal, T. Aboulnasr and S. Panchanathan, "Image Indexing using Translation and Scale-

invariant Moments and Wavelets", Storage and Retrieval for Image and Video Databases V, SPIE Proceedings Series vol I. 3022, pp. 380~389, 1997. 2.

[16] M. Stricker and M. Orengo, "Similarity of Color Images", Storage and Retrieval for Image and Video Databases III, SPIE Proceedings Series vol I. 2420, pp. 381~392, 1995.

[17] W. Niblack, R. Barber, et. al., "The QBIC Project: Querying Images by Content Using Color, Texture, and Shape", SPIE, vol. 1908, pp. 173~187, 1993.

[18] Yining Deng, et. al., "Peer group filtering and perceptual color image quantization", Proc. of ISCAS, 1999.

[19] Virginia E. Ogle. Michael Stonebraker, "Chabot: Retrieval from a Relational Database of Images", IEEE Computer vol. 28, no. 9, pp. 40~48, 1995.

[20] John R. Smith. Shih-Fu Chang, "VisualSEEK: a fully automated content-based image query system", ACM Multimedia' 96, 1996. 11.



최 인 호

1990년 2월 경희대학교 전자공학과 졸업(공학사)  
 1992년 2월 경희대학교 전자공학과 대학원(공학석사)  
 2000년 2월 경희대학교 전자공학과 대학원(공학박사)  
 1996년 3월~현재 김포대학 전자정

보계열 조교수

관심분야 : 영상처리 및 영상통신, 영상 및 비디오 검색 및 색인, 코딩



이 상 훈

1989년 2월 경희대학교 전자공학과 졸업(공학사)  
 1991년 2월 경희대학교 전자공학과 대학원(공학석사)  
 2000년 8월 경희대학교 전자공학과 대학원(공학박사)  
 1997년 3월~현재 부산정보대학 정

보통신계열 조교수

관심분야 : 디지털영상처리, 데이터베이스 응용