

논문 2006-43SP-1-7

웨이브릿 변환 영역의 칼라 및 질감 특징을 이용한 영상검색

(Image Retrieval Using Multiresolution Color and Texture Features in Wavelet Transform Domain)

천 영 덕*, 성 중 기**, 김 남 철*

(Young-Deok Chun, Joong-Ki Sung, and Nam-Chul Kim)

요 약

본 논문에서는 웨이브릿 변환된 영역에서 추출된 다해상도 칼라 및 질감 특징의 효율적인 결합을 이용한 점진적 영상검색 기법을 제안한다. 칼라 특징으로 칼라 영상의 H(Hue)와 S(Saturation) 성분의 칼라 오토코렐로그램을 선택하였고, 질감 특징으로는 V(value) 성분의 BDIP와 BVLC 모멘트를 선택하였다. 선택된 특징들에 대하여 웨이브릿 변환 영역의 각 분해 레벨로부터 다해상도 특징벡터들을 얻었다. 칼라와 질감 특징의 다해상도 특징벡터들은 특징들의 차원들과 표준 편차 벡터들에 의해 정규화되어 효율적으로 결합되었고, 저장 공간을 고려하여 각 대상 영상들의 특징벡터들은 효율적으로 양자화 되었으며, 점진적 검색 기법을 적용하여 유사도 계산시 계산량을 줄였다. 제안한 방법은 칼라 히스토그램, 칼라 오토코렐로그램, SCD, CSD, 웨이브릿 모멘트, EHD, BDIPBVLC, 칼라 히스토그램과 웨이브릿 모멘트의 결합을 이용한 방법들보다 정확도 대 재현율 평가에서는 평균 15%, ANMRR 평가에서는 평균 0.2 향상된 성능을 나타내었다. 특히, 제안한 방법은 다양한 해상도를 가지는 영상 DB에서 더욱 우수한 성능을 나타내었다

Abstract

We propose a progressive image retrieval method based on an efficient combination of multiresolution color and texture features in wavelet transform domain. As a color feature, color autocorrelogram of the hue and saturation components is chosen. As texture features, BDIP and BVLC moments of the value component are chosen. For the selected features, we obtain multiresolution feature vectors which are extracted from all decomposition levels in wavelet domain. The multiresolution feature vectors of the color and texture features are efficiently combined by the normalization depending on their dimensions and standard deviation vector, respectively, vector components of the features are efficiently quantized in consideration of their storage space, and computational complexity in similarity computation is reduced by using progressive retrieval strategy. Experimental results show that the proposed method yields average 15% better performance in precision vs. recall and average 0.2 in ANMRR than the methods using color histogram, color autocorrelogram, SCD, CSD, wavelet moments, EHD, BDIP and BVLC moments, and combination of color histogram and wavelet moments, respectively. Specially, the proposed method shows an excellent performance over the other methods in image DBs contained images of various resolutions.

Keywords : CBIR, multiresolution feature, combination, progressive retrieval

I. 서 론

영상검색 방법 중의 하나인 내용기반 영상검색(content-based image retrieval) 기법은 질의 영상으로부터 칼라(color), 질감(texture), 모양(shape) 등과 같은 영상의 내용과 관련된 시각적 특징을 추출한 다음, 영상 데이터베이스(DB)로부터 미리 저장된 각각의 대상 영상 특징과 질의 영상의 특징과의 유사도를 측정하여, 질의 영상과 유사한 영상을 검색하는 기법으로, 최근 다양한 방법들이 활발하게 연구되고 있다^{[1]~[6]}.

* 정희원, 경북대학교 전자공학과
(Department of Electronic Engineering, KyungPook National University)

** 회원, LG.PHILIPS LCD (LG.PHILIPS LCD)

※ 본 논문은 정보통신부 정보통신연구진흥원의 정보통신기초기술연구지원사업(04-기초-007)의 지원을 받아 수행되었습니다.

접수일자 : 2005년7월12일, 수정완료일 : 2005년11월1일

하나의 영상을 대표할 수 있는 특징의 추출은 내용기반 영상검색에서 가장 중요한 것 중의 하나이다. 모양은 영상 내의 객체의 윤곽선을 나타내는 특징으로 인간이 시각적으로 이미지를 분류하거나 인식하는 데 있어서 가장 중요한 정보이다. 그러나 객체의 윤곽선을 정확하게 추출하는 것이 쉽지 않기 때문에 영상검색을 위한 모양 특징은 일반적으로 상표 영상과 같이 객체와 배경의 구분이 명확한 영상에 주로 사용됨을 볼 수 있다^[5]. 칼라는 내용기반 영상검색에서 가장 널리 사용되는 특징으로 영상의 크기와 방향에 독립적이며 특징 추출 방법이 간단하지만 조명이나 밝기 변화에 민감한 단점을 가진다^[6]. 일반적으로 사용하는 칼라 특징으로는 칼라 히스토그램(color histogram)^[1], 칼라 코렐로그램(color correlogram)^[2], 그리고 MPEG-7 서술자인 CSD(color structure descriptor) 및 SCD(scalable color descriptor)^[3] 등이 있다. 특히, 칼라 코렐로그램은 거리에 따른 칼라 쌍의 공간적인 상관관계를 나타낸 것으로, 칼라 히스토그램이 표현하지 못하는 영상의 공간 정보를 포함하기 때문에 일반적으로 칼라 히스토그램에 비해 검색 성능이 우수하다고 알려져 있다^[2]. 칼라 코렐로그램에서의 칼라 쌍 중 같은 칼라 쌍의 공간적인 상관관계를 나타내는 칼라 특징을 칼라 오토코렐로그램(autocorrelogram)이라 한다. 이는 칼라 코렐로그램에 비해 특징의 크기가 작고 계산량이 적기 때문에 일반적으로 칼라 코렐로그램보다 더 많이 사용된다^[5].

질감은 구름, 나무, 벽돌, 직물 등과 같은 실세계 객체들의 표면특성과 관련된 중요한 특징으로 표면의 구조적인 배열과 주위 화소와의 관계에 대한 정보를 나타낸다^[4]. 내용기반 영상검색에 사용되는 질감 특징으로는 웨이브릿 모멘트^[7], MPEG-7 서술자인 EHD(edge histogram descriptor)^[3] 등이 있다. 최근에는 영상내의 국부 밝기변화를 잘 표현하는 BDIP(block difference inverse probabilities) 서술자와 영상이 가지는 질감의 부드러움과 거친 정도를 잘 표현하는 BVLC(block variance of local correlation coefficient) 서술자를 이용한 검색 기법^[8]이 제안된 바 있으며 차원이 작은 특징의 경우에도 비교적 성능이 우수한 검색 결과를 나타내는 효율적인 특징으로 알려져 있다. 그러나 이들 방법의 특징은 각 서술자를 구하는 데 있어 영상의 국부적 특성을 측정하기 위해 블록으로 나눌 때 영상을 고정된 크기의 블록을 사용한다.

하나의 영상 획득 장치를 이용하여 어떤 피사체를 촬영할 때, 원거리에서 촬영된 피사체는 작은 수의 화소

로 표현되지만 근거리에서 촬영된 피사체는 보다 많은 수의 화소로 표현된다. 그 결과로서, 같은 종류의 영상들이라도 서로 다른 해상도를 갖는 영상들이 한 영상 DB로 수집되어 저장될 수 있다. 따라서 해상도에 관계없이 고정된 거리의 이웃 화소와의 관계를 이용하는 특징은 다양한 해상도의 영상들이 포함된 DB에 대해서는 저조한 성능을 나타낼 수 있다. 따라서 다양한 해상도를 갖는 영상 DB에 대해 강인한 특성을 나타낼 수 있도록 특징들을 다해상도로 표현하는 연구가 요구된다.

일반적으로 하나의 영상은 여러 가지 시각적 특징을 포함하고 있으므로 하나의 특징만으로 영상을 표현하는 데에는 다소 한계가 있다. 이에 따라 최근에는 여러 가지 특징을 결합하여 영상을 검색하는 다중 특징을 이용한 검색기법에 관한 연구가 활발히 수행되고 있다^{[9]-[11]}. [9]에서는 칼라 특징으로는 CIE Lab 색좌표에서의 히스토그램이 사용되었고 질감 특징으로는 웨이브릿 프레임(frames)에서 추출한 분산이 사용되었다. [10]에서는 칼라 특징으로 칼라 히스토그램이 사용되었고 질감 특징으로 Haar and Daubechies 웨이브릿 모멘트가 사용되었다. [11]에서는 RGB값의 공분산(covariance)과 DCT 계수를 추출하여 각각 칼라 특징과 질감 특징으로 사용하였다. 한편 이와 같이 다양한 결합의 예를 볼 수 있으나 아무 특징들이나 결합한다고 해서 항상 특징의 크기를 증가시키지 않고도 검색 성능이 향상되는 것을 기대하기는 어렵다^[12]. 각각의 특징이 우수한 성능을 나타내면서 상호 보완되는 특징들을 선택하여 효율적으로 결합한다면 보다 향상된 성능을 얻을 수 있으리라 기대된다.

실용적인 내용기반 영상검색을 위해서는 검색 성능 뿐만 아니라 질의 영상의 특징 추출 시의 계산량, 특징의 차원, 저장 공간, 그리고 유사도 계산시 필요한 계산량도 중요하게 고려되어야 한다. 특히, 유사도 계산시 필요한 계산량은 특징의 차원과 DB 내의 영상 수에 의존하게 된다. 방대한 양의 영상들을 대상으로 검색할 경우 많은 계산량이 필요하게 되어 검색 속도를 저하시키게 된다. 그러나 이러한 문제는 [13], [14]에서 적용된 것과 같은 점진적인 검색 방법을 적용한다면 어느 정도 해결할 수 있다.

본 논문에서는 다해상도 칼라 및 질감 특징의 효율적인 결합을 이용한 점진적 영상검색 기법을 제안한다. 칼라 특징으로 칼라 영상의 H(Hue)와 S(Saturation) 성분의 칼라 오토코렐로그램을 선택하였고, 질감 특징으로는 V(value) 성분의 BDIP와 BVLC 모멘트를 선택하

였다. 선택된 특징에 대하여 웨이브릿 변환 영역의 각 분해 레벨로부터 다해상도 특징벡터들을 얻었다. 효율적인 특징벡터의 결합을 위해 추출된 칼라와 질감 특징의 다해상도 특징벡터들은 특징들의 차원들과 표준 편차 벡터들에 의해 정규화되었고, 저장공간을 고려하여 각 대상 영상들의 특징벡터들은 효율적으로 양자화되었다. 그리고 점진적 검색 기법을 적용하여 유사도 계산시 계산량을 줄였다.

II. 내용기반 영상검색을 위한 칼라 및 질감 특징

본장에서는 제안된 영상검색 기법에서 사용될 기준의 특징들, 칼라 특징으로 칼라 오토코렐로그램, 질감 특징으로 BDIP 및 BVLC에 대해 살펴본다.

1. 칼라 오토코렐로그램

칼라 오토코렐로그램은 주어진 화소의 칼라와 임의의 거리에 존재하는 그 주위 화소의 칼라가 같을 확률로 정의되며 다음과 같이 표현된다^[2].

$$\alpha^k(l) = \Pr [p' \in I \mid p' - p = k \text{ and } p' \in I(l) \text{ for } p \in I(l) \subset I, l \in \{0, 1, \dots, L-1\}] \quad (1)$$

여기서, $\Pr[\cdot]$ 은 내부조건을 만족하는 확률이다. I 는 영상내의 모든 화소들의 집합을 나타내고 $I(l)$ 은 칼라가 l 인 화소들의 집합을 나타낸다. 영상 내의 화소 p 와 p' 의 위치가 각각 (x, y) , (x', y') 일 때 두 화소 사이의 거리는 다음과 같이 정의된다.

$$|p' - p| = \max\{|x' - x|, |y' - y|\}. \quad (2)$$

그 결과로 (1)식에서 $\alpha^k(l)$ 는 칼라가 l 인 임의의 화소 p 가 주어질 때 주어진 화소로부터 거리 k 만큼 떨어진 화소 p' 의 칼라가 l 일 확률을 나타낸다. 따라서 (2)식을 사용하여 이웃 화소와의 거리 k 에 따른 영상의 칼라 특성을 구할 수 있다. 칼라 오토코렐로그램은 같은 칼라를 가진 영역이 영상 전체에 퍼져있을 때 보다, 모여 있을 때 더 크게 나타나며, 카메라의 위치, 줌 변화에 의한 모양 변화에 강인한 특징을 보인다^[15]. 칼라 오토코렐로그램을 추출하는 데 있어서, RGB 칼라 공간보다 인간의 직관적 시각표현에 바탕을 둔 색상(hue), 채도(saturation), 명도(value)로 구성된 HSV 칼라 공간상에서 추출한 오토코렐로그램으로 영상을 검색하는 것이 우수한 성능을 나타내는 것으로 알려져 있다^[16].

2. BDIP(Block Difference Inverse Probabilities)

화소의 밝기가 급격히 변화하여 물체와 물체간의 경계를 나타내는 에지와 주위 화소들에 비해서 상대적으로 밝기가 어두운 벨리는 영상에 있어서 매우 중요한 특징들이다. 특히, 벨리는 시각적인 물체의 형태 인식에 있어서 에지보다 더욱 기본적이고 중요한 특징이다^[17]. BDIP는 영상 내의 국부적 밝기가 고려된 에지뿐만 아니라 벨리를 잘 측정하는 특징으로 다음과 같다^[8].

$$\beta^k(l) = \frac{\frac{1}{|B_l^k|} \sum_{(x,y) \in B_l^k} (\max_{(x,y) \in B_l^k} I(x,y) - I(x,y))}{\max_{(x,y) \in B_l^k} I(x,y)} \quad (3)$$

여기서, $I(x, y)$ 는 (x, y) 위치의 화소 밝기를 나타내며, $l = (u, v)$ 은 래스터 주사(raster scanning)에 따른 블록의 순서를 나타내며 (u, v) 는 영상 내의 블록의 수직 및 수평의 위치를 나타낸다. 그리고 B_l^k 는 블록 내 최대거리가 k 인 l 위치의 블록을 나타낸다. 이 때, 블록의 크기는 $B_l^k = (k+1)^2$ 로 주어진다. (3)식에서 분자는 비선형적으로 특징된 블록 내의 밝기의 평균적 변화도를 나타낸 것이다.

BDIP의 특성을 보다 깊이 있게 이해하기 위하여 다음과 같이 정의되는 두 연산자를 생각하기로 하자.

$$\text{연산자 1: } \frac{|\nabla I|}{I} \quad (4)$$

$$\text{연산자 2: } \frac{|\nabla I|}{I} \quad (5)$$

여기서 $|\nabla I|$ 는 주어진 영역 내의 I 의 그래디언트 크기의 평균을 나타내며, I 는 그 영역 내의 I 의 평균을 나타낸다. (4)식의 첫 번째 연산자는 Sobel이나 Roberts 연산자와 같은 일종의 에지 연산자라고 할 수 있다. 또한, (5)식의 두 번째 연산자는 (4)식의 에지 연산자 결과를 추가로 밝기값의 평균으로 나누어 줌으로써 얻을 수 있음을 알 수 있다. 이 연산자에서 각각 선형적으로 주어지는 밝기의 국부적 평균 변화도와 분모의 국부적 평균 밝기를 비선형적으로 주어지는 양들로 대치한 것이 BDIP 연산자라 할 수 있다.

그림 1은 원영상과 이 영상에 앞서 언급된 연산자들을 가하여 얻은 결과 영상을 나타낸다. 이때 결과 영상 그림 1(b)-(d)는 연산자들에 의한 출력값들을 반전(negative)하여 나타내었다. 그리고 (3)식의 $k=1$ 을 선택하였으며 $|\nabla I|$ 는 Roberts 교차 기울기(crossing gradient)^[18] 연산자를 사용하였고 I 는 크기가 2×2 블록 내의 화소들의 평균값을 사용하였다. 그림 1(b)는 화

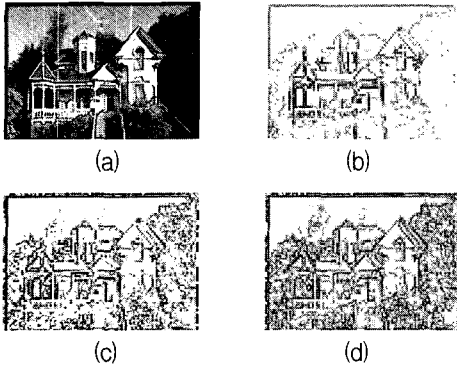


그림 1. 원영상과 에지연산자에 의해 처리된 영상들: (a) 원영상, (b) ∇ , (c) $\nabla T/I$, (d) $\beta^1(I)$

Fig. 1. Original image and its processed images:: (a) original image, (b) ∇ , (c) $\nabla T/I$, (d) $\beta^1(I)$.

표 1. 에지 연산자에 따른 정확도 성능
Table 1. Precision performance according to edge operators.

	∇	$\nabla T/I$	$\beta^1(I)$
Corel DB	59.8%	84.0%	85.2%
VisTex DB	61.9%	84.5%	86.1%

소의 밝기 변화도가 큰 물체의 경계부분을 잘 추출함을 볼 수 있다. 그림 1(c)는 그림 1(b)에서의 물체의 경계 부분의 추출뿐만 아니라 어두운 부분 내의 밝기 변화량도 잘 추출함을 볼 수 있다. 그림 1(d)는 그림 1(c)와 유사한 결과를 나타내지만 어두운 부분 내의 밝기 변화량을 더욱 강조하여 나타냄을 볼 수 있다. 이와 같은 결과는 ∇ 를 I 로 정규화함으로써 어두운 부분의 밝기 변화를 강조함을 볼 수 있으며 또한 I 를 블록 내의 최대값으로 대체함으로써 어두운 부분의 밝기 변화를 더욱 더 강조함을 볼 수 있다. 그리고 최대값에 의한 정규화는 잡음을 완화하는 효과가 있다.

이러한 특성을 갖는 연산자들의 영상 검색 성능을 조사한 결과는 표 1과 같다. 각 연산자들의 특징벡터 추출방법은 [8]에서의 방법과 같이 적용하였고 영상 DB로는 Corel DB와 VisTex DB^[8]를 사용하였다. $\nabla T/I$ 을 이용한 방법은 ∇ 을 이용한 방법보다 평균 23% 우수한 성능을 나타내었으며 BDIP를 이용한 방법은 $\nabla T/I$ 을 이용한 방법보다 평균 1.4% 우수한 성능을 나타냄을 볼 수 있다.

이와 같이 BDIP는 밝은 영역에서의 밝기 변화뿐만 아니라 어두운 영역에서의 밝기 변화를 잘 나타내며 이를 이용한 검색 방법도 우수한 성능을 나타내었다. 이러한 BDIP는 Weber 법칙^[18]으로 알려진 바와 같이 밝은 영역에서의 밝기 변화보다 어두운 영역에서 동일한

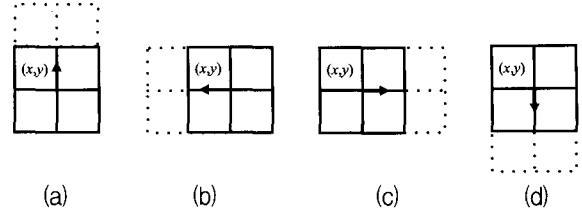


그림 2. 2×2블록과 네 방향으로 이동된 블록의 형상: (a) 90°, (b) 180°, (c) 0°, (d) -90°

Fig. 2. Pixel configuration in 2×2 windows and their corresponding windows shifted in each of four directions, which are required to compute: (a) 90°, (b) 180°, (c) 0°, (d) -90°.

밝기 변화를 더 잘 인지하는 인간의 비선형적인 시각적 특성이 잘 반영된 특징이라 할 수 있다. 또한 BDIP는 (3)식의 블록 내 화소간의 최대 거리 k 에 따른 영상의 질감 특성을 구할 수 있다.

3. BVLC(Block Variance of Local Correlation Coefficient)

BVLC^[8]는 이웃 화소들과의 국부상관계수들 변화량을 블록 단위로 구하여 영상내의 부드러움과 거친 정도를 측정하는 질감 특징이다. 국부상관계수(local correlation coefficients)는 국부 분산(local variance)으로 정규화된 국부 공분산(local covariance)을 의미하며 다음과 같이 표현된다.

$$\rho^k(I) = \frac{\frac{1}{|B_l^k|} \sum_{(x,y) \in B_l^k} I(x,y)I(x+\Delta_x(k),x+\Delta_y(k)) - \mu_l \mu_{l+\Delta(k)}}{\sigma_l \sigma_{l+\Delta(k)}} \quad (6)$$

여기서, $l=(u,v)$ 은 래스터 주사에 따른 블록의 순서를 나타내며 (u,v) 는 영상 내의 블록의 수직 및 수평의 위치를 나타낸다. μ_l 과 σ_l 은 각각 블록 l 의 밝기 평균과 표준편차를 나타낸다. $\Delta(k)=(\Delta_x(k),\Delta_y(k))$ 는 각 네 방향(90°, 180°, 0°, -90°) 중의 하나로 k 만큼 이동함을 나타내며 $\mu_{l+\Delta(k)}$ 와 $\sigma_{l+\Delta(k)}$ 는 블록 l 로부터 k 만큼 이동된 블록의 밝기 평균과 표준편차를 나타낸다. 그림 2는 $k=1$ 일 때 각각 네 방향으로 이동된 블록의 형상을 나타낸다. 따라서, BVLC는 화소 위치가 (x,y) 일 때, 그림 2의 네 방향의 국부상관계수에서 최대값에서 최소값을 뺀 값으로 정의되며 다음과 같이 표현된다.

$$\gamma^k(I) = \max_{\Delta(k) \in O_4} [\rho^k(I)] - \min_{\Delta(k) \in O_4} [\rho^k(I)], \quad (7)$$

$$O_4 = \{(-k,0), (0,-k), (0,k), (k,0)\}.$$

이 식으로부터 BVLC는 블록 내에 거침이 강하면 큰

값을 가지고 부드러움이 강하면 작은 값을 가짐을 볼 수 있다. 또한 BVLC는 (7)식의 블록 내 화소간의 최대 거리 k 에 따른 영상의 질감 특성을 구할 수 있다.

III. 제안한 영상검색 기법

그림 3은 제안한 검색 기법의 블록도를 나타낸다. RGB 칼라 질의 영상이 검색 시스템으로 입력되면, HSV 칼라 변환을 거친 다음 웨이브릿 변환으로 분해된다. 분해된 H, S 영상으로부터는 칼라 오토코렐로그램의 칼라 특징을 추출하고 V 영상으로부터는 BDIP 및 BVLC 모멘트의 질감 특징을 추출한다. 시스템은 추출된 특징들을 결합한 다음 영상 DB에서 미리 추출된 각각의 대상 영상의 특징들과 질의 영상의 특징들과의 유사도를 점진적으로 측정하고 질의 영상과 유사도가 높은 일정 개수의 대상 영상들을 출력한다.

1. 웨이브릿 분해

HSV로 변환된 질의 영상 $I(x, y)$ 의 각 칼라 성분 영상 $I_H(x, y), I_S(x, y), I_V(x, y)$ 는 웨이브릿 변환으로 분해되어 $W_m^H(x, y), W_m^S(x, y), W_m^V(x, y)$ 영상이 된다. 여기서 n 은 웨이브릿 분해 레벨을 나타내며 $m(n)$ 은 n 레벨에서의 부대역(subband)을 나타낸다. 그림 4는 $n=1$ 레벨에서의 웨이브릿 분해를 보여준다. 그림 4(a)에서 칼라 성분영상 $I_c(x, y)$ 에 수평 방향으로 각각 고역통과필터 $g(x)$ 와 저역통과필터 $h(x)$ 를 통과시키고, 변환된 계수를 간축(decimation)한다. 다음으로 이렇게 처리된 계수들에 수직방향으로 각각 고역통과필터 $g(y)$ 와 저역통과필터 $h(y)$ 를 통과시키고, 변환된 계수를 간축한다. 그 결과로 1 레벨 분해된 저역밴드 영상 $W_{LL(1)}^c(x, y)$ 와 3개의 고역밴드 영상, 즉 수직, 수평 및 대각선 방향의 영상 $W_{HL(1)}^c(x, y), W_{LH(1)}^c(x, y), W_{HH(1)}^c(x, y)$ 를 얻게 된다. $n=2$ 레벨에서의 웨이브릿 분해 영상은 그림 4(a)에서와 같이 $I_c(x, y)$ 대신 저역밴드 영상 $W_{LL(1)}^c(x, y)$ 을 입력영상으로 하여 그림 4(a)와 같이 웨이브릿 분해하면 $W_{LL(2)}^c(x, y), W_{HL(2)}^c(x, y), W_{LH(2)}^c(x, y), W_{HH(2)}^c(x, y)$ 을 얻을 수 있다. 제안한 검색 방법에서 사용되는 웨이브릿 필터로는 실용적인 검색을 위해 계산량이 비교적 적게 소요되는 Haar 필터^[19]를 선택하였고 검색성능과 특징의 차원을 고려하여 분해 레벨 $n=2$ 를 선택하였다.

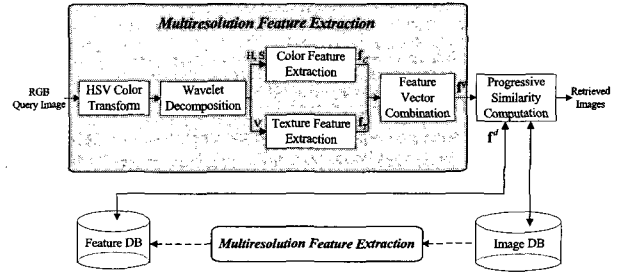


그림 3. 제안한 검색 기법의 블록도
Fig. 3. Block diagram of the proposed retrieval method.

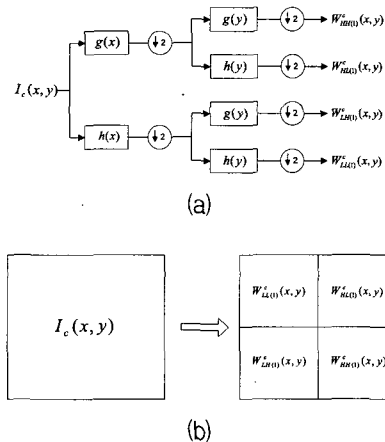


그림 4. 웨이브릿 분해: (a) 4대역 웨이브릿 분해과정, (b) 1레벨 웨이브릿 분해된 형상
Fig. 4. Wavelet decomposition: (a) Block diagram for 4-band wavelet decomposition, (b) Configuration of 1-level wavelet decomposed images.

2. 칼라 특징 추출

웨이브릿 분해된 각 H, S 칼라 성분 영상들로부터 다해상도 칼라 오토코렐로그램을 추출하는 과정은 그림 5와 같다. 먼저 H와 S 성분으로부터 웨이브릿 분해된 각 칼라 영상 $W_m^H(x, y)$ 와 $W_m^S(x, y)$ 는 양자화 Q 를 거쳐 양자화된 영상 $Q(W_m^H(x, y))$ 와 $Q(W_m^S(x, y))$ 가 된다. 이 때 양자화 방법은 각 부대역의 칼라 분포 특성에 따라 양자화 방법을 다르게 적용한다. 그림 6은 100개의 시험 영상들의 H 칼라 성분에 대해 각각 1 레벨 분해한 다음, 이들 분해된 영상들에 대해 각 대역별 평균 칼라 분포를 나타낸 것이다. 그림 6(a)의 LL 밴드의 칼라 값은 원 칼라 성분 영상에서와 같이 전 범위에 걸쳐 분포하지만 그림 6(b)-(d)의 HL, LH, HH 밴드들에서는 0 주위에 몰려 있음을 볼 수 있다. 이러한 통계적 특성에 맞춰 LL 밴드는 균등 양자화하였고 나머지 대역들은 일반화된 로이드 알고리즘(generalized Lolyd algorithm)을 이용하여 비균등 양자화 하였다.

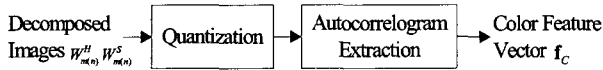


그림 5. 칼라 특징 추출 과정

Fig. 5. Procedure of color feature extraction.

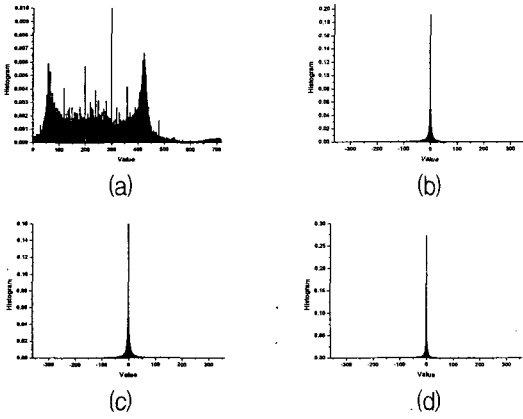


그림 6. 시험영상에 대한 웨이브릿 분해된 hue 영상의 평균 칼라 분포: (a) LL band. (b) HL band. (c) LH band. (d) HH band.

Fig. 6. Average color distribution for wavelet decomposed hue images obtained from test images.

칼라 특징의 양자화 레벨 수 L 또는 차원은 다음과 같이 결정하였다. 주어진 칼라 특징의 차원이 L 로 주어질 때 각 부밴드의 차원 L_i 와는 다음과 같은 관계를 가진다.

$$L = \sum_{i=0}^{K-1} L_i = \text{constant} \quad (8)$$

여기서 i 는 부밴드 영상 $W_{m(n)}^c(x, y)$ 을 나타내며 K 는 부밴드 영상의 수를 나타낸다. (8)식의 부밴드의 차원 배정에 따른 왜곡을 최소화하는 L_i 는 다음과 같이 주어진다^[21].

$$\log_2 L_i = \frac{\log_2 L}{K} + \frac{1}{2} \log_2 \frac{\sigma_i^2}{\left[\prod_{j=0}^{K-1} \sigma_j^2 \right]^{1/K}} \quad (9)$$

여기서 σ_i^2 는 i 번째 부밴드 영상의 분산을 나타낸다. 즉 부밴드의 차원은 이들의 분산에 따라 결정된다. 그 결과로, 양자화된 영상 $Q(W_{m(n)}^H(x, y))$ 와 $Q(W_{m(n)}^S(x, y))$ 에 대해 각각 1 레벨 분해된 영상의 부밴드에서 LL 밴드는 8, 나머지 HL, LH, HH 밴드들은 각각 4, 그리고 2 레벨 분해된 영상의 부밴드에서 LL 밴드는 4, 나머지 HL, LH, HH 밴드들은 각각 2로 결정되었다.

각 부밴드에 대해 양자화 레벨이 결정된 후 다음과

같이 칼라 오토코렐로그램을 추출한다.

$$\alpha_{m(n)}^k(c, l) = \Pr[p' \in Q(W_{m(n)}^c) \mid |p' - p| = k \text{ and } p' \in Q(W_{m(n)}^c(l)) \text{ for } p \in Q(W_{m(n)}^c(l)) \subset Q(W_{m(n)}^c)] \quad (10)$$

여기서 $\Pr[\cdot]$ 은 내부조건을 만족하는 확률이며 $Q(W_{m(n)}^c)$ 은 양자화된 분해 영상을 나타내며 $Q(W_{m(n)}^c(l))$ 은 화소 값이 l 인 화소들의 집합을 나타낸다. p 는 영상의 화소를 나타내고 p' 는 p 와의 거리가 k 인 이웃 화소를 나타낸다. 따라서 (10)식에서 $\alpha_{m(n)}^k(c, l)$ 는 다해상도로 분해된 양자화 영상 $Q(W_{m(n)}^c)$ 에서 화소 값이 l 인 임의의 화소 p 가 주어질 때 주어진 화소로부터 거리 k 만큼 떨어진 화소 p' 의 화소 값이 l 일 확률을 나타낸다. 본 연구에서는 칼라 특징의 차원과 검색성을 고려하여 $k=1$ 을 선택하였으며, 그 결과로 칼라 오토코렐로그램의 특징벡터 f_c 는 다음과 같이 주어진다.

$$f_c = [\alpha_{m(n)}^1(c, i)] \quad (11)$$

여기서 $[\alpha_{m(n)}^1(c, i)]$ 는 각 칼라성분 $c \in \{H, S\}$, 부밴드 $m \in \{LL, HL, LH, HH\}$, 분해레벨 $n \in \{1, 2\}$, 그리고 양자화 레벨 $i \in \{0, 1, \dots, L_{m(n)}\}$ 를 만족하는 성분들로 구성된 칼라 오토코렐로그램 벡터를 나타낸다.

3. 질감 특징 추출

선택된 질감 특징인 BDIP 및 BVLC 모멘트의 다해상도 추출 과정은 그림 7과 같다. 먼저, 질의 영상의 V 성분으로부터 웨이브릿 분해된 $W_{m(n)}^V(x, y)$ 영상을 얻은 다음, 각각의 부대역 영상을 일정한 크기의 블록으로 겹치지 않게 나눈다. 이어서 각각의 블록에 대해 BDIP와 BVLC를 계산한다. 그러나 BDIP 계산시 음수가 존재하지 않는 공간영역에서만 추출이 가능하도록 정의되어 있는 (3)식의 기존의 BDIP 연산자는 적합하지 않다. (3)식의 분모는 웨이브릿 영역에서는 음수값을 가질 수 있으므로 이러한 방식의 밝기 변화에 대한 측정은 부적절해질 수 있다. 따라서 BDIP의 정의를 다음과 같이 수정하였다.

$$\beta_{m(n)}^k(l) = \frac{1}{|B_l^k|} \sum_{(x,y) \in B_l^k} \frac{\left(\max_{(x,y) \in B_l^k} W_{m(n)}^V(x, y) - W_{m(n)}^V(x, y) \right)}{\max_{(x,y) \in B_l^k} W_{LL(n)}^V(x, y)} \quad (12)$$

$m \in \{LL, HL, LH, HH\}$

여기서, $W_{m(n)}^V(x, y)$ 는 n 레벨 웨이브릿 변환된 V 칼라 성분 영상의 m 방향의 부밴드에서의 (x, y) 위치의 화소

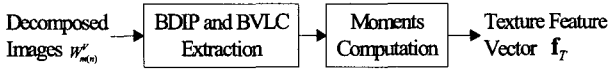


그림 7. 질감 특징 추출 과정
Fig. 7. Procedure of texture feature extraction.

밝기를 나타낸다. B_l^k 는 블록 내 최대거리가 k 인 l 위치의 블록을 나타낸다. 여기서, l 은 래스터 주사에 따른 블록의 순서를 나타낸다. 이때 블록의 크기는 $B_l^k = (k + 1)^2$ 로 주어진다. (12)식의 분모 식이 1보다 작을 경우, 이로 인하여 불안정하게 분자가 크게 증폭되는 것을 피하기 위해 강제로 분모를 1의 값을 갖도록 하였다.

웨이브릿 분해된 영상에서 BVLC는 다음과 같이 추출된다. 먼저 다해상도로 분해된 영상의 국부상관계수는 다음과 같다.

$$\rho_{m(n)}^k(l) = \frac{\frac{1}{|B_l^k|} \sum_{(x,y) \in B_l^k} W_{m(n)}^V(x,y) W_{m(n)}^V(x+\Delta_x(k), y+\Delta_y(k)) - \mu_l \mu_{l+\Delta(k)}}{\sigma_l \sigma_{l+\Delta(k)}} \quad (13)$$

BVLC는 네 방향의 국부상관계수에서 최대값에서 최소 값을 뺀 값으로 정의되며 다음과 같이 표현된다.

$$\gamma_{m(n)}^k(l) = \max_{\Delta(k) \in O_4} [\rho_{m(n)}^k(l)] - \min_{\Delta(k) \in O_4} [\rho_{m(n)}^k(l)], \quad (14)$$

$$O_4 = \{(-k,0), (0,-k), (0,k), (k,0)\}.$$

칼라 특징과 마찬가지로 다해상도로 분해된 영상 $W_{m(n)}^V(x,y)$ 은 (12)식과 (14)식의 주어진 k 에 의해 분해 레벨 n 에 따라 이웃 화소의 거리가 다르게 표현되므로 분해 레벨에 따라 서로 다른 영상의 질감 특성을 얻을 수 있다. 본 연구에서는 질감 특징의 차원과 검색성을 고려하여 $k=1$ 을 선택하였다.

이와 같이 블록 단위로 BDIP와 BVLC를 계산한 다음, 각 분해 레벨에 따른 각 밴드별로 다음과 같이 BDIP와 BVLC의 1차 및 2차 모멘트들을 추출한다.

$$\mu(\beta_{m(n)}^1) = \langle \beta_{m(n)}^1(l) \rangle \quad (15)$$

$$\sigma^2(\beta_{m(n)}^1) = \langle (\beta_{m(n)}^1(l) - \mu(\beta_{m(n)}^1))^2 \rangle \quad (16)$$

$$\mu(\gamma_{m(n)}^1) = \langle \gamma_{m(n)}^1(l) \rangle \quad (17)$$

$$\sigma^2(\gamma_{m(n)}^1) = \langle (\gamma_{m(n)}^1(l) - \mu(\gamma_{m(n)}^1))^2 \rangle \quad (18)$$

여기서 $\mu(\cdot)$ 와 $\sigma(\cdot)$ 는 각각 평균과 표준편차를 나타낸다. 이로부터 질감 특징벡터 f_T 는 다음과 같이 주어진다.

$$f_T = [[\mu(\beta_{m(n)}^1)], [\sigma(\beta_{m(n)}^1)], [\mu(\gamma_{m(n)}^1)], [\sigma(\gamma_{m(n)}^1)]] \quad (19)$$

여기서 $[\mu(\beta_{m(n)}^1)]$ 와 $[\sigma(\beta_{m(n)}^1)]$ 는 각각 BDIP의 평균 벡

터와 표준편차 벡터를 나타내며 $[\mu(\gamma_{m(n)}^1)]$ 와 $[\sigma(\gamma_{m(n)}^1)]$ 는 각각 BVLC의 평균 벡터와 표준편차 벡터를 나타낸다. 이들은 각각 분해 레벨 $n \in \{1, 2\}$ 와 부밴드 $m \in \{LL, HL, LH, HH\}$ 을 만족하는 성분들로 구성된다.

4. 특징벡터의 결합 및 점진적 유사도 측정

질의 영상의 특징벡터가 구성되면 시스템은 이 특징벡터와 영상 DB에서 미리 추출된 각각의 대상 영상의 특징벡터와의 유사도를 계산하여 질의 영상과 유사도가 높은 일정 개수의 대상 영상들이 검색된다. 본 연구에서는 유사도 계산시 계산량을 줄이기 위해 3 단계 점진적인 검색을 적용하였다. 이 때, (11)식의 칼라 특징벡터 f_C 와 (19)식의 질감 특징벡터 f_T 는 다음과 같이 결합되어 각 단계별로 적용된다.

$$f_j = \left[\frac{f_{C(j)}}{N_{C(j)} \cdot \sigma_{C(j)}}, \frac{f_{T(j)}}{N_{T(j)} \cdot \sigma_{T(j)}} \right] \quad (20)$$

$$f_{C(j)} = [[\alpha_{m(n)}^1(H, i)], [\alpha_{m(n)}^1(S, i)]]$$

$$f_{T(j)} = [[\mu(\beta_{m(n)}^1)], [\sigma(\beta_{m(n)}^1)], [\mu(\gamma_{m(n)}^1)], [\sigma(\gamma_{m(n)}^1)]]$$

여기서 $f_{C(j)}$, $f_{T(j)}$, 그리고 f_j 는 각각 j 단계 검색에서 사용되는 칼라 특징벡터, 질감 특징벡터, 그리고 이들을 결합한 특징벡터를 나타낸다. $N_{C(j)}$ 와 $N_{T(j)}$ 는 각각 j 단계 검색에서 사용되는 칼라와 질감 특징벡터들의 차원을 나타낸다. $\sigma_{C(j)}$ 와 $\sigma_{T(j)}$ 는 각각 전체 DB에서 얻은 칼라 특징벡터와 질감 특징벡터에 대한 표준편차벡터를 나타낸다. 여기서 벡터 연산은 벡터 성분간의 연산을 의미한다. 각 단계에서 적용되는 분해레벨 n , 부밴드 m , 그리고 양자화 레벨 i 는 다음을 만족한다.

$$n \in X_j \subset \{1, 2\}, \quad i \in Z \subset \{0, 1, \dots, L_{m(n)}\}, \quad (21)$$

$$m \in Y_j \subset \{LL, HL, LH, HH\}$$

$$\text{where } A_u \subset A_v \text{ for } u < v, \quad A \in \{X, Y, Z\}.$$

(21)식의 각 특징벡터들은 $N_{C(j)}$ 와 $N_{T(j)}$ 에 의해 특징벡터간에 정규화되고 $\sigma_{C(j)}$ 와 $\sigma_{T(j)}$ 에 의해 특징벡터의 각 성분간에 정규화 되었다. 전자는 특징벡터의 각 차원이 다를 경우 차원이 큰 특징벡터에 의해 검색 성능이 좌우되는 것을 보완해주며 후자는 특징벡터 각 성분별 분포 편차가 다를 경우 분포 편차가 큰 성분에 의해 검색 성능이 좌우되는 것을 보완해준다.

이와 같이 결합된 특징벡터를 이용하여 다음과 같이 점진적 검색을 수행한다. 먼저 1 단계 검색에서는 비교적 낮은 특징벡터의 차원으로 구성된 1 레벨 분해된 LL 부밴드에서 추출한 특징벡터들의 성분을 이용하여 DB

내의 모든 영상들을 대상으로 유사도를 계산하여 일정 수의 영상을 검색한다. 2단계 검색에서는 좀 더 높은 특징벡터의 차원으로 1 레벨 분해된 영상의 모든 부밴드들에서 추출한 특징벡터들의 성분을 이용하여 1 단계에서 검색된 영상들만 대상으로 하여 검색한다. 마지막으로 3단계 검색에서는 모든 특징벡터의 성분들을 사용하여 2 단계에서 검색된 영상들만 대상으로 최종 검색한다. 이 때, 각 단계에서 적용한 질의 영상의 특징벡터와 대상 영상의 특징벡터의 유사도 측정을 위한 거리는 다음과 같다.

$$D_j(\mathbf{f}_j^q, \mathbf{f}_j^d) = \sum_{i=1}^{N_j} |f_j^q(i) - f_j^d(i)| \quad (22)$$

여기서 q 와 d 는 각각 질의 영상과 대상 영상을 나타내며 $f_j^q(i)$ 와 $f_j^d(i)$ 는 j 단계에서 사용되는 각각 질의 특징벡터 \mathbf{f}_j^q , 대상 특징벡터 \mathbf{f}_j^d 의 i 번째 성분들을 나타내고, N_j 는 j 단계에서 적용된 특징벡터의 차원을 나타낸다.

점진적 유사도 측정을 위한 j 단계에서의 검색되는 영상 수 RN_j 는 최종 J 단계에서 검색되는 영상 수에 의해 결정되며 다음과 같이 주어진다.

$$RN_j = RN_J \times e^{(J-j)}. \quad (23)$$

이때, 점진적인 검색에 의한 유사도 측정시의 계산복잡도는 (40)식에 의해 덧셈만 사용되며 그 수는 다음과 같이 주어진다.

$$S \times (2N_1 - 1) + \sum_{j=2}^J RN_{j-1} \times (2N_j - 1) \quad (24)$$

여기서 S 는 DB 내의 영상 수를 나타낸다. 한편, (22)식의 유사도 측정시 (20)식의 특징벡터 \mathbf{f} 를 점진적으로 적용하지 않고 모든 특징 벡터를 한 번에 적용하는 비점진적인 검색에 의한 유사도 측정시 사용되는 덧셈 수는 다음과 같이 주어진다.

$$S \times (2N_j - 1) \quad (25)$$

따라서 $S \gg RN_j$ 이면 유사도 측정시 점진적 검색에서 사용되는 덧셈 수는 비점진적인 검색에서 사용되는 덧셈 수보다 약 N_1/N_j 배 만큼 적게 사용됨을 볼 수 있다.

5. 특징벡터의 저장

효율적인 영상검색 시스템을 구현하기 위해서는 검색 성능, 특징벡터의 차원, 계산 복잡도 뿐만 아니라 특징 DB의 저장공간도 중요하게 고려되어야 한다. DB의 저장공간을 결정하는 것은 DB내 영상의 개수, 특징벡

터의 차원, 그리고 각 벡터성분이 가지는 비트 수(bit size)이다. 따라서 본 논문에서는 특징 DB의 저장공간을 줄이기 위한 방법으로 각 칼라 및 질감 특징벡터들의 성분들을 검색 성능이 크게 떨어지지 않는 범위 내에서 양자화하여 비교적 적은 비트 수로 표현하였다. 칼라 오토코렐로그램의 특징벡터 성분들은 5 bits로 표현하였고 BDIP 및 BVLC 특징벡터 성분들은 각각 6 bits로 표현하였다. 따라서 특징벡터가 하나의 영상을 저장하는데 필요한 비트 수는 $5N_c + 6N_t$ 가 된다.

IV. 실험결과 및 검토

1. 영상 DB 및 성능 평가

실험을 위한 영상 DB는 칼라 영상으로서 Corel DB^[8], VisTex DB^[8], Corel_MR DB, VisTex_MR DB를 사용하였다. 뒤의 두 가지 DB들은 앞의 두 가지 DB들로부터 각각 파생된(derived) DB로서 각 부류 내의 영상의 해상도가 여러 가지로 구성되어 있다. Corel DB는 일반적인 실세계 영상과 질감 영상이 포함되어 있고, VisTex DB는 질감 영상으로만 구성되어 있다.

Corel DB는 11 부류 각 90장씩 총 990장의 192×128 해상도를 가지는 영상들로 구성되어 있으며 VisTex DB는 75 부류 각 16장씩 총 1200장의 128×128 해상도의 질감 영상들로 구성되어 있다. Corel_MR DB에는 Corel DB에서 각 부류당 90개의 영상들 중에서 3분의 1은 그대로 포함되었고 3분의 2는 각각 $(1.5:1)^2$ 와 $(2:1)^2$ 의 비율로 간축(decimation)되어 포함되었다. VisTex_MR DB에는 VisTex DB에서 각 부류당 16개의 영상들 중에서 4분의 1은 그대로 포함되었고 4분의 3 각각 $(1.5:1)^2$, $(1.75:1)^2$, $(2:1)^2$ 의 비율로 간축되어 포함되었다. 그림 8은 DB들에 포함된 영상의 예를 나타낸 것이다.

검색 성능을 평가하는 방법으로는 정확도(precision) 및 정확도 대 재현율(precision vs. recall)^[21], 그리고 ANMRR(average normalized modified retrieval rank)^[22]을 사용하였다. DB영상들의 집합에서 질의영상을 q , 질의영상과 유사한 영상들의 집합을 $S(q)$, 그리고 검색 결과로 출력되는 영상들의 집합을 $A(q)$ 라 했을 때, 정확율 $P(q)$ 와 재현율 $R(q)$ 는 다음과 같이 표현된다.

$$P(q) = \frac{|A(q) \cap S(q)|}{|A(q)|}, R(q) = \frac{|A(q) \cap S(q)|}{|S(q)|} \quad (26)$$

여기서, \cdot 는 집합의 크기를 나타낸다. 정확도는 검색된 영상의 수에 대한 검색된 영상 중에 질의영상과

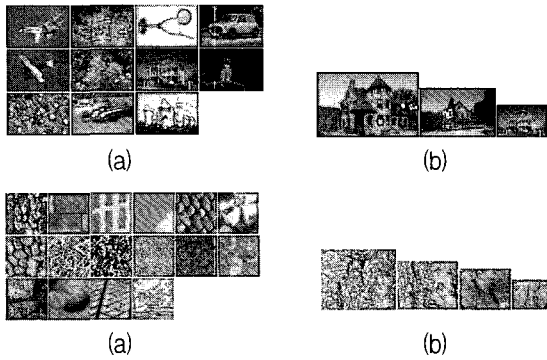


그림 8. 실험 DB 영상들의 예: (a) Corel DB, (b) Corel_MR DB, (c) VisTex DB, (d) VisTex_MR DB

Fig. 8. Example of test DB images: (a) Corel DB, (b) Corel_MR DB, (c) VisTex DB, (d) VisTex_MR DB,

유사한 영상의 수의 비를 나타내고, 재현율은 DB내에서 질의 영상과 유사한 영상의 수에 대한 검색된 영상들 중에서 질의 영상과 유사한 영상의 수의 비를 나타낸다. 질의 영상으로는 각각의 DB 내의 각각의 영상을 모두 사용하였고, 이들을 이용한 검색 성능은 각 DB내의 사용된 각 질의 영상에 대해 평가된 성능들의 평균으로 정해진다. ANMRR은 MPEG-7에서 표준으로 정해진 검색 성능 평가 방법으로 적은 값을 가질수록 우수한 성능을 나타낸다.

2. 실험 결과

표 2는 제안된 점진적인 검색법과 (22)식에서 $j = 3$, 즉 모든 특징벡터들을 적용하여 DB 내의 모든 영상을 대상으로 하여 유사도를 측정하는 비점진적 검색법에 대한 정확도 성능을 나타낸다. 여기서 점진적인 검색에서의 최종 3단계에서 추출되는 영상 수 RN_3 와 비점진적인 검색에서 추출되는 영상 수, 즉 (26)식의 $A(q)$ 의 크기는 각각 10을 선택하였다. 점진적 검색의 각 단계에서 사용되는 특징벡터의 차원 N_1, N_2, N_3 는 각각 20, 56, 92를 가지며, 비점진적인 검색에서의 특징벡터의 차원 N 은 92가 된다.

표 2에서 점진적 검색의 검색 성능은 비점진적 검색의 검색성능에 비해 Corel DB에서는 0.9%, VisTex DB

표 2. 비점진적 검색과 점진적 검색의 정확도 성능
Table 2. Precision performance of nonprogressive and progressive retrieval.

	Non-progressive	Progressive
Corel DB	94.7%	93.8%
VisTex DB	93.6%	93.1%

표 3. 각 검색방법에서 사용된 칼라공간과 특징의 차원
Table 3. Color space and dimension of retrieval methods.

Method	Color space	Dimension
Color histogram	RGB	128
Color autocorrelogram	HSV	128
SCD	HSV	128
CSD	HMMD	128
EHD	RGB	240
Wavelet moments	RGB	96
BDIPBVLC	RGB	96
Color histogram + Wavelet moments	RGB	124
Proposed	HS, V	92

에서는 0.7%만큼 떨어짐을 볼 수 있다. 그러나 비점진적 검색과 점진적 검색의 계산 복잡도는 (24)식과 (25)식으로부터 각각 $183S, 39S + 13155$ 의 덧셈 수를 가진다. 여기서 S 는 DB 내의 영상 수를 나타낸다. Corel DB와 VisTex DB 내의 영상 수는 각각 990, 1200 이므로 점진적 검색의 덧셈 수는 비점진적 검색의 덧셈 수에 비해 각각 3.5배, 3.7배 적게 나타냄을 볼 수 있다. 따라서 점진적 검색은 검색 성능에 대한 계산 복잡도의 효율이 다소 큼을 볼 수 있다. 이러한 점진적 검색은 DB내의 영상 수가 클수록 더욱 효과적이라고 할 수 있다.

그림 9는 칼라 히스토그램, 칼라 오토코렐로그램, SCD, CSD, EHD, 웨이브릿 모멘트, BDIP 및 BVLC 모멘트, 칼라히스토그램과 웨이브릿 모멘트의 결합을 이용한 방법, 그리고 제안한 방법에 대한 검색 성능을 나타낸다. 각 방법들에서 사용된 칼라 공간과 차원은 표 3과 같다. 각 검색방법들에 대해 차원을 선택하는데 있어서 정해진 차원을 사용하는 방법(EHD)을 제외하고는 각 검색방법들의 차원이 큰 차이가 나지 않는 범위에서 차원에 따른 검색성능을 고려하여 효율적인 성능을 나타내는 차원으로 결정하였다. 칼라 히스토그램에서 사용된 각각 RGB 성분들에 대한 양자화 레벨은 각각 8:4:4를 선택하였다. 칼라 오토코렐로그램에서 HSV 성분들에 대한 양자화 레벨은 각각 8:4:4로 선택하였으며 (1)식에서의 거리 $k = 1$ 을 선택하였다. 웨이브릿 모멘트에서 웨이브릿 분해를 위해 웨이브릿 필터 중 가장 우수한 검색 성능을 나타내는 Daubechies biorthogonal 9/7 tap filters^[20]를 선택하였고, 분해레벨 $n = 5$ 를 선택하였다. 칼라 히스토그램과 웨이브릿 모멘트를 결합을 이용한 방법에서 RGB 칼라 히스토그램의 각 성분의 양자화 레벨은 각각 4로 선택하였고 웨이브릿 모멘트에서 사용된 분해 레벨 $n = 3$ 을 선택하였다. 그리고 두 특징은 (20)식과 같이 제안한 방법과 같은 방식으로 결합되

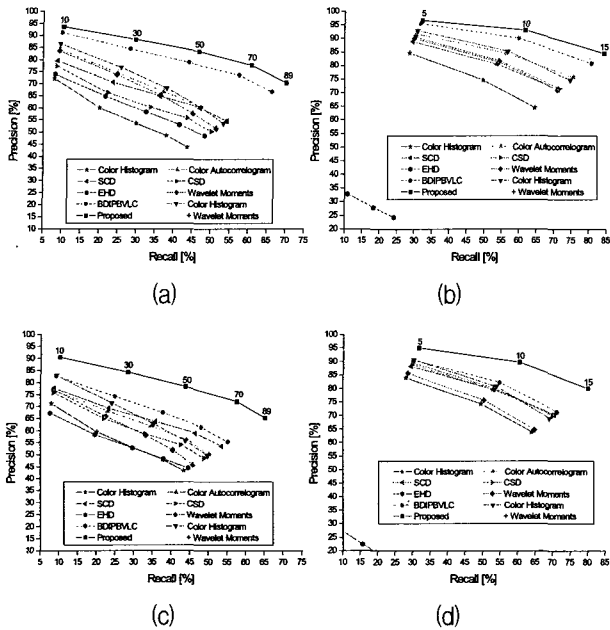


그림 9. 검색 방법들의 정확도 대 재현율: (a) Corel DB, (b) VisTex DB, (c) Corel_MR DB, (d) VisTex_MR DB

Fig. 9. Precision vs. recall of retrieval methods: (a) Corel DB, (b) VisTex DB, (c) Corel_MR DB, (d) VisTex_MR DB.

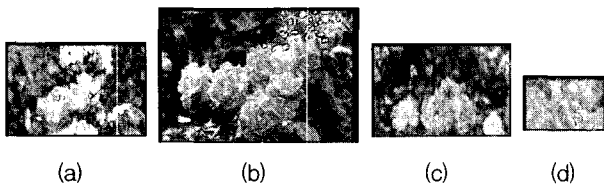


그림 10. 질의 영상 (a)과 유사한 영상들 (b) (c), (d)
Fig. 10. Query image (a) and its relevant images (b), (c), (d).

었다. 그림 9에서 제안한 방법의 그래프 위의 숫자는 검색된 영상의 수를 나타낸다.

제안한 방법은 평균 정확도에 있어서 나머지 방법들보다 Corel DB(그림 9(a))에서는 3.8%~27.1%, VisTex DB(그림 9(b))에서는 3%~63.1%, Corel_MR DB(그림 9(c))에서는 9.8%~23.8%, VisTex_MR DB(그림 9(d))에서는 6.9%~65.1% 향상된 성능을 나타내었다. 또한 제안한 방법은 BDIP 및 BVLC 모멘트 방법과 칼라 오토코렐로그램 방법보다 Corel DB에서는 각각 3.8%, 14.7%, VisTex DB에서는 각각 3%, 7.4%, Corel_MR DB에서는 각각 9.8%, 15.4%, VisTex_MR DB에서는 6.9%, 8.8% 향상된 성능을 보여 두 특징의 결합의 효과가 다소 크게 나타났다.

각 검색 방법들의 다해상도 영상 DB와 비다해상도 영상 DB에서의 성능을 비교하기 위해 Corel DB(그림

표 4. 그림 10에서의 질의 영상에 대한 유사 영상들의 검색 순위

Table 4. Retrieval rank of it relevant images for a query image in Fig. 10.

Method \ Image	Fig. 10(b)	Fig. 10(c)	Fig. 10(d)
Color autocorrelation	54	2	15
BDIPBVLC	17	1	43
Proposed	7	2	3

표 5. 검색 방법들의 ANMRR 성능

Table 5. ANMRR performance of retrieval methods.

Method \ DB	Corel	VisTex	Corel_M_R	VisTex_M_R
Color Histogram	0.4622	0.2637	0.4657	0.2680
Autocorrelation	0.3480	0.1705	0.4054	0.2216
SCD	0.3574	0.2047	0.3649	0.2132
CSD	0.4040	0.1949	0.4310	0.2168
EHD	0.3849	0.6831	0.4272	0.7647
Wavelet Moments	0.3669	0.1995	0.4264	0.2507
BDIP and BVLC	0.2213	0.1193	0.3220	0.1930
Color Histogram + Wavelet Moments	0.3564	0.1704	0.3944	0.2158
Proposed	0.1837	0.0840	0.2165	0.1131

9(a)와 Corel_MR DB(그림 9(c))에서의 성능 차이와 VisTex DB(그림 9(b))와 VisTex_MR DB(그림 9(d))에서의 성능 차이를 살펴보자. 칼라 히스토그램은 평균 1.3%, SCD 평균 1.6%, CSD 평균 1.9%, 칼라 오토코렐로그램 평균 6.2%, BDIPBVLC 평균 9.8%, 그리고 제안한 방법은 평균 3.3%로 Corel DB와 VisTex DB에서 우수하게 나타내었다. 이로부터 각각 고정된 이웃 화소와의 거리와 고정된 블록의 크기를 사용한 칼라 오토코렐로그램 방법과 BDIP 및 BVLC 모멘트 방법은 다해상도로 구성된 DB들에서 성능 저하가 다른 방법에 비해 다소 큼을 볼 수 있다. 그러나 이들을 다해상도에서 추출하여 결합을 이용한 제안한 방법은 다해상도 DB들에서도 성능 차이가 비교적 크지 않음을 볼 수 있다.

그림 10은 질의 영상과 이것의 유사 영상들의 예를 나타내며, 표 4는 그림 10의 질의 영상에 대한 BDIP 및 BVLC 모멘트 방법, 칼라 오토코렐로그램 방법, 그리고 제안한 방법에 의해 검색된 이들 유사 영상들의 검색 순위를 나타낸다. 그림 10(b)의 영상은 질의 영상보다 높은 해상도를 가지는 영상이며, 그림 10(c)는 질의 영상과 같은 해상도를 가지는 영상이며, 그리고 그림 10(d) 질의 영상보다 낮은 해상도를 가지는 영상이다. 질의 영상과 해상도가 같은 그림 10(c) 영상의 검색 순위는 세 가지 검색방법에서 비슷하게 나타나지만 질의

영상과 해상도가 다른 그림 10(b)와 그림 10(d)의 영상들은 BDIP 및 BVLC 모멘트 방법과 칼라 오토코렐로그래프 방법보다 제안한 방법에서의 검색 순위들이 훨씬 높음을 볼 수 있다. 이러한 결과로부터 제안한 방법은 질의 영상과 다른 해상도를 가지는 영상들에 대해서도 강한 특성을 보임을 알 수 있다.

표 5는 각 방법들에 대한 ANMRR 성능을 나타낸다. 제안한 방법은 다른 방법들에 비해 Corel DB에서 0.2785~0.0376, VisTex DB에서 0.0353~0.5991, Corel_MR DB에서 0.2492~0.1055, VisTex_MR DB에서 0.0799~0.6515 향상된 성능을 나타내었다. ANMRR 평가에서도 제안한 방법의 다해상도 영상 DB와 비다해상도 영상 DB에서의 성능 차이는 BDIP 및 BVLC 모멘트 방법과 칼라 오토코렐로그래프 방법의 성능 차이에 비해 작음을 볼 수 있다. 또한 제안한 방법은 다양한 해상도를 가지는 영상 DB에서 다른 방법들과의 성능 차이가 더 크게 나타남을 볼 수 있다.

V. 결 론

본 논문에서는 다해상도 칼라 및 질감 특징의 효율적인 결합을 이용한 점진적 영상검색 기법을 제안하였다. 칼라 특징으로 칼라 영상의 H와 S 성분의 칼라 오토코렐로그래프를 선택하였고, 질감 특징으로는 V성분의 BDIP와 BVLC 모멘트를 선택하였다. 선택된 특징에 대하여 웨이브릿 변환 영역의 각 분해 레벨로부터 다해상도 특징벡터들을 얻었다. 효율적인 특징벡터의 결합을 위해 추출된 칼라와 질감 특징의 다해상도 특징벡터들은 특징들의 차원들과 표준 편차 벡터들에 의해 정규화되었고, 저장공간을 고려하여 각 대상 영상들의 특징벡터들은 효율적으로 양자화되었다. 그리고 점진적 검색 기법을 적용하여 유사도 계산시 계산량을 줄였다.

실험 결과, 제안된 점진적 검색은 비점진적 검색에 비해 소폭의 검색 성능의 저하를 나타내었으나 유사도 계산시 계산 복잡도를 크게 감소시켜 검색 성능에 대한 계산 복잡도의 효율이 큼을 보였다. 또한 제안한 방법은 비교한 검색 방법들보다 향상된 성능을 나타내었으며, 칼라 오토코렐로그래프 방법과 BDIP 및 BVLC 모멘트 방법들을 단독으로 사용하였을 때보다 높은 성능을 나타내어 결합 효과가 큼을 확인하였다. 특히, 제안한 방법은 다양한 해상도를 가지는 영상 DB에서 다른 방법들과의 성능 차이가 크게 나타내어 이들 영상 DB에서 더욱 효과적임을 확인하였다.

참 고 문 헌

- [1] M. J. Swain and D. H. Ballard, "Color indexing," *Int. J. Computer Vision*, vol. 7, pp. 11-32, 1991.
- [2] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlograms," *IEEE Proceedings of Computer Vision and Pattern Recognition*, pp. 762-768, 1997.
- [3] "ISO/IEC 15938-3/FDIS Information technology Multimedia content description interface-part 3 visual," ISO/IEC/JTC1/SC29/WG11, Doc. N4358, July 2001.
- [4] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Texture features for image classification," *IEEE Trans. Syst. Man Cybern.*, vol. 8, pp. 610-621, Nov. 1973.
- [5] D. Feng, W. C. Siu, and H. J. Zhang, *Fundamentals of Content-based Image retrieval, in Multimedia Information Retrieval and Management-Technological Fundamentals and Applications*, New York, NY, Springer, 2003.
- [6] Y. Rui and T. S. Huang, "Image retrieval: current techniques, promising directions, and open issues," *J. Visual Communication and Image Representation*, vol. 10, pp. 39-62, Oct. 1999.
- [7] J. R. Smith and S.-F. Chang, "Transform features for texture classification and discrimination in large image databases," in *Proc. IEEE Int. Conf. Image Processing*, vol. 3, pp. 407-411, Nov. 1994.
- [8] Y. D. Chun, S. Y. Seo, and N. C. Kim, "Image retrieval using BDIP and BVLC moments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 951-957, Sep. 2003.
- [9] S. Liapis and G. Tziritas, "Color and texture image retrieval using chromaticity histograms and wavelet frames," *IEEE Trans. Multimedia*, vol. 6, pp. 676-686, Oct. 2004.
- [10] A. Vadivel, A. K. Majumdar, and S. Sural, "Characteristics of weighted feature vector in content-based image retrieval applications," in *Proc. IEEE Int. Conf. Intelligent Sensing and Information processing*, Chennai, India, pp. 127-132, Jan. 2004.
- [11] H. Permuter, J. Francos, and I. H. Jermyn, "Gaussian mixture models of texture and colour for image database retrieval," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal processing*, vol. 3, Hong Kong, pp. 569-572, Apr. 2003.
- [12] 성중기, 칼라의 공간적 상관관계 및 국부 질감특성

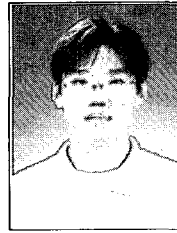
- 을 이용한 영상검색, 경북대학교 석사학위논문, 2004년 12월.
- [13] M. Ankerst, H. P. Kriegel, and T. Seidl, "A multistep approach for shape similarity search in image databases," *IEEE Trans. Knowledge and Data Engineering*, vol. 10, pp. 996-1004, Nov.-Dec. 1998.
- [14] B. C. Song, M. J. Kim, and J. B. Ra, "A fast multiresolution feature matching algorithm for exhaustive search in large image databases," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11, pp. 673-678, May 2001.
- [15] Jing Huang, S. R. Kumar, M. Mitra, and W. J. Zhu, "Spatial color indexing and applications," *Computer Vision, Sixth International Conference*, pp. 602-607, 1998.
- [16] T. Ojala, M. Rautiainen, E. Matinmikko, and M. Aittola, "Semantic image retrieval with HSV correlogram," *Proc. 12th Scandinavian Conf. On Image Analysis*, Bergen, Norway, pp. 621-627, 2001.
- [17] D. E. Pearson and J. A. Robinson, "Visual communication at very low data rates," *Proc. IEEE*, vol. 73, pp. 795-812, Apr. 1985.
- [18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing 2nd Edition*, Prentice Hall, Upper Saddle River, NJ, 2002.
- [19] E. J. Stollnitz, T. D. DeRose, and D. H. Salesin, *Wavelets for Computer Graphics: Theory and Applications*, Morgan Kaufmann, 1996.
- [20] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, 1992.
- [21] S. F. Chang, W. C. Horace J. Meng, H. Sundaram, and D. Zhong, "A fully automated content-based video search engine supporting spatiotemporal queries," *IEEE Trans. Circuits Sys. Video Technol.*, vol. 8, no. 5, pp. 602-615, Sep. 1998.
- [22] P. Ndjiki-Nya, J. Restat, T. Meiers, J. R. Ohm, A. Seyferth, and R. Sniehotta, "Subjective evaluation of the MPEG-7 retrieval accuracy measure (ANMRR)," ISO/WG11 MPEG Meeting, Geneva, Switzerland, May 2000, Doc. M6029.

 저 자 소 개



천 영 덕 (정회원)
 2000년 인제대학교 전자공학과
 학사
 2002년 경북대학교 전자공학과
 석사
 2002년~현재 경북대학교 전자공
 학과 박사과정

<주관심분야 : 영상처리, 영상통신, 컴퓨터비전>



성 중 기 (정회원)
 2003년 경북대학교 전자공학과
 학사
 2005년 경북대학교 전자공학과
 석사
 2005년~현재 LG.PHILIPS LCD

<주관심분야 : 영상처리, 영상검색>



김 남 철 (정회원)
 1978년 서울대학교 전자공학과 학사
 1980년 한국과학기술원 전기 및 전자공학과 석사
 1984년 한국과학기술원 전기 및 전자공학과 박사
 1984년~현재 경북대학교 전자전기공학부 교수
 1994년 1월~1996년 12월 한국통신학회 및 대한전자공학회 편집위원
 1997년 1월~1998년 12월 한국통신학회 신호처리연구회 위원장

2002년 9월~2004년 8월 경북대학교 전자전기컴퓨터 학부장
 2002년 9월~2004년 8월 경북대학교 BK21 정보기술인력양성사업단 단장
 2005년 1월~현재 대한전자공학회 협동이사, 경북대 직장대표
 2004년 12월~현재 모바일 단말 상용화센터 인력양성사업 경북대학교 사업책임자
 2005년 9월~현재 경북대학교 정보전산원장
 <관심분야 : 영상처리, 영상압축, 영상복원, 컴퓨터 비전>