

논문 2005-42SP-5-1

내용 기반 영상 검색을 위한 에지 기반의 공간 기술자

(Edge-based spatial descriptor for content-based image retrieval)

김 낙 우*, 김 태 용**, 최 종 수**

(Kim Nac-Woo, Kim Tae-Yong, and Choi Jong-Soo)

요 약

오늘날 급격한 멀티미디어 정보의 증가에 따라 영상에서의 시각적 특성을 이용하여 멀티미디어 데이터를 검색하는 내용 기반 영상 검색 기법에 대한 관심이 크게 늘어나고 있다. 본 논문에서는 효과적인 영상 검색을 위한 새로운 접근으로서 edge correlogram과 color coherence vector를 이용한 에지 기반의 공간 기술자를 제안한다. 우선 color vector angle 기법을 이용하여 주어진 영상을 고주파 성분과 저주파 성분의 영상으로 나눈다. 저주파 성분의 영상에서는 color coherence vector를 이용하여 평탄 화소의 공간적인 색상 분포를 추출함으로써 이를 평탄 영역에서의 특징 정보로서 활용한다. 반면, 고주파 성분의 영상에서는 edge correlogram으로부터 에지 화소들 간의 분포를 추출하여 이를 에지 영역에서의 특징 정보로 이용한다. 제안된 방법은 색상 간의 지역적인 특성과 전체적인 특성을 모두 가지고 있기 때문에, 영상 간의 비교에 있어서 영상의 모양과 크기의 급격한 변화로 인한 오검출 등에 매우 강건하다. 또한, 영상에서의 구조적인 특징을 이용함으로써 복잡한 영상에 대해서도 간단하고 유연한 특징을 제공한다. 실험 결과는 영상 색인 및 검색에 있어서 제안된 알고리즘이 최근의 여러 히스토그램 정밀화 기법에 비하여 더 효과적임을 보여준다. 데이터베이스 내 영상의 색인을 위해서는 R*-tree 구조를 이용하였다.

Abstract

Content-based image retrieval systems are being actively investigated owing to their ability to retrieve images based on the actual visual content rather than by manually associated textual descriptions. In this paper, we propose a novel approach for image retrieval based on edge structural features using edge correlogram and color coherence vector. After color vector angle is applied in the pre-processing stage, an image is divided into two image parts (high frequency image and low frequency image). In low frequency image, the global color distribution of smooth pixels is extracted by color coherence vector, thereby incorporating spatial information into the proposed color descriptor. Meanwhile, in high frequency image, the distribution of the gray pairs at an edge is extracted by edge correlogram. Since the proposed algorithm includes the spatial and edge information between colors, it can robustly reduce the effect of the significant change in appearance and shape in image analysis. The proposed method provides a simple and flexible description for the image with complex scene in terms of structural features of the image contents. Experimental evidence suggests that our algorithm outperforms the recently histogram refinement methods for image indexing and retrieval. To index the multidimensional feature vectors, we use R*-tree structure.

Keywords : Color coherence vector, edge correlogram, color vector angle, image retrieval

I. 서 론

최근의 인터넷 사용 인구의 증가와 멀티미디어 정보의 폭발

적 증가로 인하여 사용자가 보다 효과적이고 유연하게 멀티미디어 자료 검색 서비스를 이용할 수 있도록 하기 위한 영상 검색 및 비디오 브라우저 시스템 등이 크게 요구되고 있다^[1,2,3,4]. 멀티미디어 자료는 기존의 문자 자료에 비해 용량이 클 뿐 아니라 비정형적인 특성을 가지고 있으며, 포함하고 있는 내용이 복잡하여 기존의 문자기반 데이터베이스로는 신속하고 효율적인 검색과 효과적인 관리에 커다란 어려움이 있다. 이러한 이유로 최근 내용 기반 영상 검색 시스템 구축을 위한 많은 연구들이 활발히 이루어지고 있다^[1,2].

* 학생회원, ** 정회원, 중앙대학교 첨단영상대학원 영상공학과

(Dep. of Image Engineering, Graduate School of Advanced Imaging Science, Multimedia, and Film, Chung-Ang University)

※ 본 연구는 과기부의 국가지정연구실(2000-N-NL-01-C-285)과 교육부의 두뇌한국 21 사업(BK21)의 연구지원으로 수행되었습니다.

접수일자: 2005년1월4일, 수정완료일: 2005년3월27일

내용 기반 영상 검색 방법 중에서 가장 대표적인 것은 Swain^[5]에 의한 히스토그램 인터섹션(histogram intersection) 기법이다. 이 방법은 같은 색상을 가진 화소들의 개수를 색상 빈(bin)의 빈도수에 따라 계수(count)하여 색상 히스토그램을 표현한다. Hafner^[6] 등은 영상의 색상분포 뿐 아니라 히스토그램의 각 구간 사이의 거리와 최대거리에 대한 비율을 유사행렬을 이용하여 히스토그램 상의 유사성을 구분하기도 하였다. 이러한 방법은 영상 내의 색상 히스토그램만을 비교하므로 계산이 복잡하지 않다는 장점이 있으나, 포함된 물체의 위치나 공간적인 관계를 나타내지 못하고 색상의 분포만을 고려함으로써 유사한 히스토그램을 가지는 서로 다른 영상을 구별하지 못하는 단점이 있다. 이에, 영상에서의 공간적인 색상 분포를 표현하기 위한 접근 방법이 연구되었다. Huang^[7,8]이 제안한 color correlogram은 효과적인 영상 특징 벡터로서 색상의 공간적 관계를 표현하는 효율적인 영상 검색 기법 중 하나이다. Huang은 이 방법에서 영상에서의 색상 간 전역적인 상관관계 뿐 아니라 부분 영역에서의 상관관계까지 고려하였다. 그러나 color correlogram 방법에서의 많은 계산량은 실제적인 응용에 있어 장애가 되고 있다. Pass^[9]는 영상 내 화소간의 공간적 인접성을 바탕으로 히스토그램 빈을 나누는 방법을 제안하였다. 영상에서 어떤 화소가 일정 크기 이상의 비슷한 색상 영역에 포함될 경우 이 화소를 '인접성 있는 화소'로 분류하고 그렇지 않다면 '인접성 없는 화소'로 분류하게 된다. Color Coherence Vector(CCV)는 영상에 나타난 각 색상마다의 이러한 분류를 통해 영상에서의 색상 응집도를 표현하고, 이를 특징 기술자로서 사용한다. 그러나, 이러한 접근 방법은 영상 상의 화소 위치에 크게 의존하기 때문에, 큰 시점 변화에 민감하다는 단점이 있다.

시점 변환된 영상 검색 시의 오검출 문제나 많은 계산량 등의 문제를 처리하기 위하여, 본 논문에서는 전처리된 예지 영상을 이용하여 색상 및 모양 특징을 추출하는 방법을 제안하고 있다. 먼저, color vector angle^[10] 기법을 이용하여 영상에서의 각 화소를 평탄 영역에서의 화소(이하 평탄 화소: smooth pixel)와 예지 영역에서의 화소(이하 예지 화소: edge pixel)로 분류한다. Color vector angle 기법은 영상에서의 밝기 변화에 둔감하고, 색조 및 채도의 변화에는 민감한 특징을 가지고 있다. 분류된 영상 중 예지 화소 영역에서의 각 그레이 쌍(gray pair)에 대한 분포는 edge correlogram을 이용하여 표현하였고, 평탄 화소 영역에서의 색상 분포

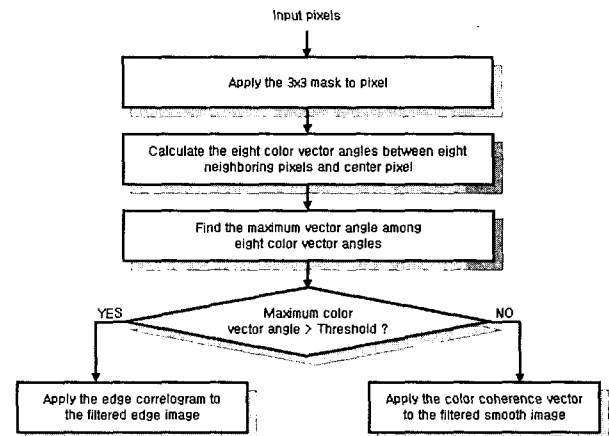


그림 1. 제안된 알고리즘의 블록도

Fig. 1. Block diagram of the proposed algorithm.

는 CCV로 나타내었다. 이로부터 영상의 색상과 모양 정보에 대한 특징을 기술하고 효과적인 특징 기반의 영상 검색 시스템을 구현하였다.

이 논문의 구성은 다음과 같다. II장에서는 color vector angle을 이용한 영상 분할과 CCV와 edge correlogram을 이용한 특징 추출 기법에 대해 제안하고, III장과 IV장에서는 제안된 방법에 대한 실험 결과와 결론을 도출한다. 그림 1은 본 연구의 대략적인 블록도를 나타낸 것이다.

II. 예지 기반의 공간 기술자

1. Proposed edge-based spatial descriptor

색상 히스토그램^[5]과 같은 전통적인 색상 기술자에서의 문제는 영상에서의 공간 정보가 부족하다는 데에 있다. 이러한 공간 정보 부재 문제의 해결을 위해 본 논문에서는 영상 객체의 예지 정보를 이용한다. 먼저, color vector angle 기법에 기초하여 각 화소를 평탄 화소나 예지 화소로 분류한다. 영상에 3 x 3 윈도우 마스크를 적용시킨 후, 최대의 color vector angle을 만들어내는 중심화소와 이에 이웃한 화소가 에지를 검출하는 데에 이용된다. 추출된 예지 화소 집합에서의 그레이 쌍의 분포는 RGB 색상 영역 상에서 양자화된 색상 간의 edge correlogram에 의해 표현된다. 역으로, 평탄 화소로 분류된 집합군의 경우 색상의 분포는 CCV에 의해 표현된다. 이렇게 추출된 특징맵으로부터 예지 기반의 공간 기술자를 취득할 수 있다.

2. Color edge detection using color vector angle

Color vector angle은 광도 변화에는 둔감하지만 색

도나 채도에는 민감한 특징을 가지고 있기 때문에 일반적으로 영상에서 조명 효과를 최소화하여 특징을 추출하기 위해 주로 사용된다^[10]. 본 논문에서는 color vector angle의 이러한 특징을 영상에서의 색상 에지를 식별하는 데에 이용하였다.

가장 간단한 색상 거리 측정법은 RGB 공간영역에서의 유클리디안 거리 측정이지만, RGB 공간에서의 유클리디안 거리 측정법은 광도 변화에 민감하고, 색도나 채도 변화에는 둔감하기 때문에 색상의 인지적인 차이를 구별 짓는 데에 사용되기는 힘들다. 그림 2에서는 같은 유클리디안 거리를 갖는 두 개의 칼라 쌍에 대한 시각적 차이에 대해 보여주고 있다. 비록 인지적으로는 색상 쌍 (C1, C2)가 (C3, C4)의 색상 쌍보다 더 유사한 것으로 보이지만, 두 색상 쌍은 같은 유클리디안 거리를 가지고 있다. 보이는 바와 같이 유클리디안 거리가 광도의 변화에 극히 민감하기 때문에, 거의 같은 색상을 가진 두 개의 화소가 큰 거리 값을 갖게 될 수도 있다. 그러나, color vector angle을 이용한다면 채도 차에 대한 이 방법의 민감성으로 인하여 색상 쌍 (C3, C4)의 color vector angle 값이 색상 쌍 (C1, C2)에서의 color vector angle 값보다 훨씬 크게 나타나게 된다. 결과적으로 color vector angle을 이용하면 두 색상 쌍 간의 인지적인 차이를 더 적절하게 표현할 수 있다는 장점을 갖게 된다. Color vector angle을 구하는 수식은 다음과 같다. 먼저, 각 θ 에 의해 나뉜 색상들의 sine값을 수식 (1)에 의해 표현한다.

$$\sin(\theta)_{v_1, v_2} = \left(1 - \frac{(v_1^T v_2)^2}{v_1^T v_1 v_2^T v_2} \right)^{1/2} \quad (1)$$

이로부터, 팔근방 연결을 기초로 하여 색상 에지를 검출한다. 3x3 마스크를 각 픽셀에 적용시키고, 이웃한 여덟 개의 픽셀들과 중심 픽셀 사이에 여덟 개의 color vector angle 중 최대값을 수식 (2)로부터 계산한다.

$$\sin(\theta)_{\max} = \text{MAX}[\sin(\theta)_{v_c v_1}, \sin(\theta)_{v_c v_2}, \sin(\theta)_{v_c v_3}, \Lambda, \sin(\theta)_{v_c v_8}] \quad (2)$$

여기에서, V_c 와 V_i 는 각각 3 x 3 마스크에서의 중심 화소와 이웃한 화소를 나타낸다.

여덟 개의 color vector angle 중 최대의 값을 α 라 할 때, 그림 4에서는 α 값의 변화에 의한 다양한 에지 영상들을 보여주고 있다. α 값에 대한 임계치를 지나치게 낮게 잡을 경우, 그림 4의 3번째 행(그림3-(c))에 대

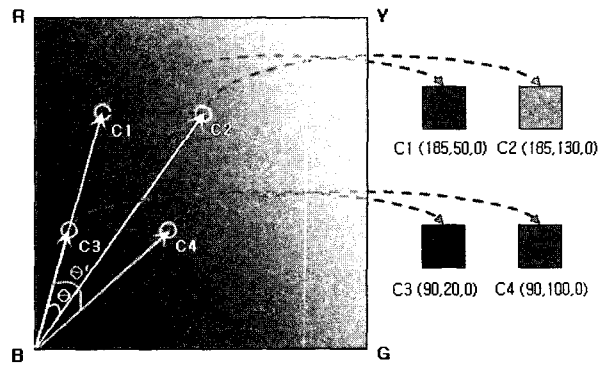


그림 2. 같은 유클리디안 거리를 갖는 두 개의 칼라 쌍에 대한 시각적 차이

Fig. 2. Visual differences of two color pairs with the same Euclidean distance in the RGB space.

한 에지 영상)에서의 그림 4-(a), 그림 4-(b)와 같이 에지 화소의 오검출이 늘게 되고, 반대로 α 값에 대한 임계치를 지나치게 높게 잡을 경우 그림 4의 2번째 행(그림 3-(b)에 대한 에지 영상)에서의 그림 4-(d)와 같이 에지 화소의 미검출이 발생한다는 것을 알 수 있다. 이러한 실험으로부터 α 값이 0.08~0.09 이상의 값을 가질 때 좋은 결과를 보인다는 것을 확인하였다. 이를 바탕으로 하여 본 논문에서는 α 값이 0.09 이상의 값을 가질 때의 V_c 를 에지 화소로, 그렇지 않을 때의 V_c 를 평탄 화소로 나눈다.

3. Color coherence vector for spatial representation

Color vector angle에 의하여 분류된 평탄 화소에 대하여, CCV라 하는 히스토그램 세밀화기법을 적용시킨다. CCV는 화소의 공간적인 인접성에 의해 히스토그램의 빈을 나누는 방법이다. 영상에서 임의의 한 화소가 같은 색상과 충분한 크기를 갖는 어느 한 영역에 속해 있을 때 '인접성 있는 화소'로 분류한다. 색상은 같으나 그 속한 영역이 충분히 크지 않을 때, 이러한 화소를 '인접성 없는 화소'로 분류한다. CCV는 영상의 각 색상 모두에 대하여 이러한 화소 구분을 실행한다. CCV를 계산하기 위한 첫 단계는 각 화소의 값을 인접한 화소들과의 평균값으로 대체함으로써 영상을 조금 흐리게 하는 것이다. 이는 영상의 양자화 시에 비슷한 위치의 영역으로부터 과도한 색상 분할을 발생시키지 않게 하기 위함이다. 다음 단계로서 흐려진 영상에 대하여 n 개의 색상 값을 갖도록 양자화한다. 양자화된 영상은 다시 각 화소별로 이웃한 화소들과의 색상 및 크기 비교를 통하여 '인접성 있는 화소'의 히스토그램과 '인접

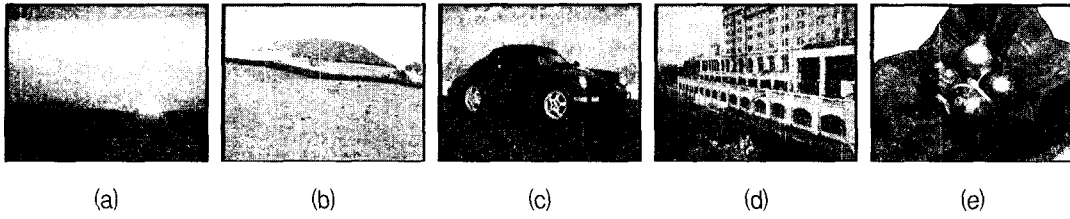


그림 3. 실험영상
Fig. 3. Test images.

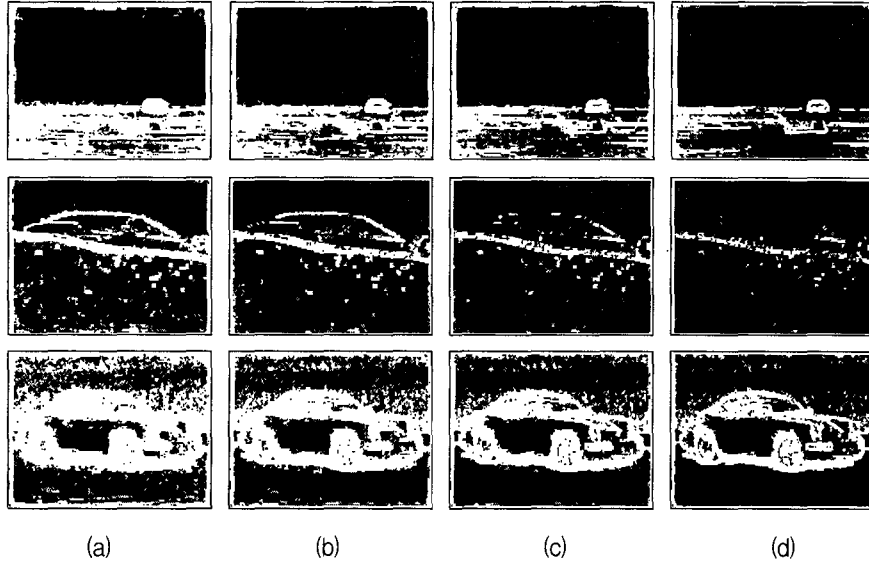


그림 4. 그림 3-(a)~그림3-(c)에서의 color vector angle에 대한 에지 영상:
(a) $\alpha=0.05$, (b) $\alpha=0.07$, (c) $\alpha=0.09$, (d) $\alpha=0.11$

Fig. 4. Edge images by color vector angle in Fig. 3-(a)~Fig.3-(c): (a) $\alpha=0.05$, (b) $\alpha=0.07$, (c) $\alpha=0.09$, (d) $\alpha=0.11$.

성 없는 화소'의 히스토그램으로 나뉜다. 이러한 CCV 기법을 통해 저주파 성분에서의 영상에 대한 색상 기술자를 얻는다.

4. Edge correlogram for shape representation

Color vector angle을 통해 에지 화소로서 분류된 화소에 대해서는 edge correlogram을 적용시킨다. 일반적으로 correlogram은 각 색상 쌍에서의 공간적인 상관관계가 거리에 따라 어떻게 변화하는 지를 표현하는 기법이다. 우리는 color correlogram 기법에서의 많은 저장 공간 요구와 큰 계산량 등을 고려하여 영상 내의 에지를 기반으로 하는 edge correlogram 기법을 이용한다. 우선, $n_1 \times n_2$ 크기의 영상 I로부터 color vector angle을 통해 분류된 에지 화소 영역을 추출하고, 이 영역을 m개의 그레이 색상으로 양자화한다. 화소 $p=(x, y) \in I$ 에 대하여, $I(p)$ 는 화소 p가 가지는 그레이 색상을 의미한다고 할 때, $I_c \approx \{p \mid I(p) = c\}$ 를 정의하고, 색상 쌍간의 거리를 나타내는 $d \in [\min(n_1, n_2)]$ 를 묘사한다. 본

논문에서는 효과적인 연산을 위하여 $d \in \{1,3,5,7\}$ 을 이용하였다. 최종적으로 edge correlogram은 $ij \in [m]$, $k \in [d]$ 에 대하여 다음과 같이 계산된다^[7,11].

$$\gamma_{c_i, c_j}^{(k)}(I) \triangleq \Pr_{p_1 \in I_{c_i}, p_2 \in I} [p_2 \in I_{c_j} \mid |p_1 - p_2| = k] \quad (3)$$

$\gamma_{c_i, c_j}^{(k)}(I)$ 는 에지 영상에서 그레이 화소 c_i 가 주어졌을 때, 이 화소로부터 k만큼 떨어진 에지 화소가 그레이 화소 c_j 일 확률을 나타내고 있다.

5. Histogram comparison between spatial area and edge area

그림 5는 그림 3에 대한 평탄 영역과 에지 영역에서의 히스토그램을 보여주고 있다. 그림 5에서 보는 바와 같이, 보통의 자연 영상 내에서 평탄 영역에서의 히스토그램은 일반적으로 에지 영역에서의 히스토그램과 비교할 때 서로 간에 큰 색상 불일치를 보인다. 이러한 영역 사이의 색상분포 차이를 고려하여, 영상의 평탄 영

역이나 에지 영역으로부터의 특징 추출 방법을 각각 분리함으로써 최종적으로 영상의 검색 성능을 향상시키고자 한다.

6. Similarity measure

영상 검색 시스템은 주어진 질의 영상과 색인된 자료 사이의 유사도 검색에 기반한다. 영상 검색 시스템은 주어진 질의 영상과 색인된 자료사이의 유사도 검색에 기반한다. 영상 간 특징값 비교를 통하여 두 영상의 유사도를 계산한 후 유사도가 높은 영상부터 차례대로 사용자에게 결과 영상을 보여준다.

가. Similarity measure for color coherence vector

질의 영상(Q)에서 j번째 색상에 대한 '인접성 있는 화소' 성분과 '인접성 없는 화소' 성분의 집합을 (α_j, β_j) 라 하고, 같은 색상에 대한 색인 영상(R)에서의 화소 성분 집합을 (α'_j, β'_j) 라 할 때, L_1 거리 측정법을 이용하여 Q와 R 영상 간 평탄 영역에서의 유사도를 다음과 같이 비교한다.

$$\Delta simC_{cev} = \sum_{j=1}^n |(\alpha_j - \alpha'_j) + (\beta_j - \beta'_j)| \quad (4)$$

임의의 두 영상 쌍 사이의 비교에서 색상 j에서의 히스토그램 차분이 비록 같은 값을 갖더라도, 그 색상 bin에서의 히스토그램 값의 크기에 의해 그 의미가 크게 달라지므로 일반적으로 다음과 같은 표준화된 거리를 이용한다. Q와 R 사이의 표준화된 거리는 다음과 같이 표현된다.

$$simC = \sum_{j=1}^n \left| \frac{\alpha_j - \alpha'_j}{\alpha_j + \alpha'_j + 1} + \frac{\beta_j - \beta'_j}{\beta_j + \beta'_j + 1} \right| \quad (5)$$

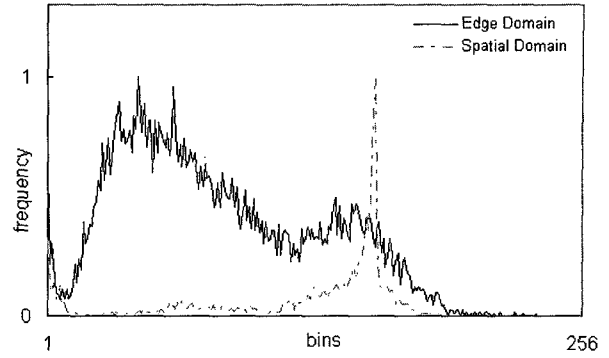
나. Similarity measure for edge correlogram

Edge correlogram을 위한 에지 영역에서의 유사도 비교 측정을 위하여 L_1 거리 측정법을 사용하고, 이를 통해 Q와 R사이의 유사도는 다음과 같이 측정된다.

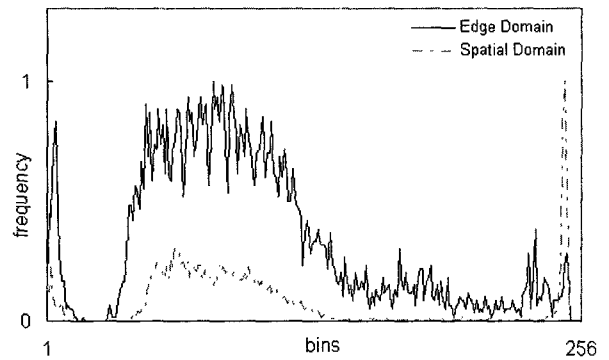
$$\Delta simS_{correl} = \sum_{i,j \in [m], k \in [d]} |\gamma_{c_i, c_j}^{(k)}(Q) - \gamma_{c_i, c_j}^{(k)}(R)| \quad (6)$$

또한, Q와 R 사이의 표준화된 거리는 다음과 같다.

$$simS = \sum_{i,j \in [m], k \in [d]} \frac{|\gamma_{c_i, c_j}^{(k)}(Q) - \gamma_{c_i, c_j}^{(k)}(R)|}{1 + \gamma_{c_i, c_j}^{(k)}(Q) + \gamma_{c_i, c_j}^{(k)}(R)} \quad (7)$$



(a)



(b)

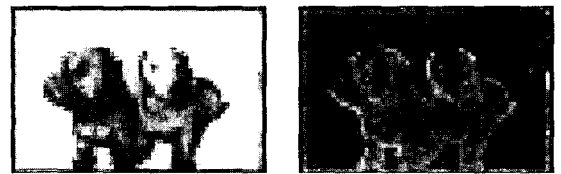
그림 5. 평탄 영역과 에지 영역에서의 히스토그램 비교: (a) 그림 3-(d)에 대한 히스토그램 (b) 그림 3-(e)에 대한 히스토그램

Fig. 5. Histogram comparison for edge and spatial area in image: (a) Histogram in Fig. 3-(d) (b) Histogram in Fig. 3-(e).



(a)

(b)



(c)

(d)

그림 6. 영상 분할: (a) 원영상 (b) 에지 영상 (c) CCV를 위한 평탄 영상 (d) Edge correlogram을 위한 에지 영상

Fig. 6. Processed images: (a) Original image (b) Image after color vector angle (c) Image to be operated by CCV (d) Image to be operated by edge correlogram.

제안한 방법의 유사도 측정은 평탄 영역과 에지 영역에서 각각 추출된 특징을 $simC(R, Q)$ 와 $simS(R, Q)$ 의 유사도 함수를 통해 계산함으로써 이루어진다. 최종적인 유사도 함수는 선형 조합으로써 다음과 같이 계산할 수 있다.

$$sim(R, Q) = \alpha \times simS(R, Q) + \beta \times simC(R, Q) \quad (8)$$

여기에서 α 와 β 는 가중치로서 일반적으로 각각 0.5의 값이 주어진다. 그림 6은 제안한 방법을 통해 처리된 중간 단계의 영상들을 보여주고 있다.

7. Indexing structure

본 논문은 R*-tree 색인 기법^[12,13]을 이용해서 제안하는 알고리즘의 특징 기술자를 색인 한다. R*-tree 기법은 현재 주로 사용되고 있는 여러 색인 기법의 기본이 되는 방법이다. R*-tree는 다차원 공간상에 위치나 크기의 관련된 정보를 원활히 관리하기 위한 동적 색인 구조이다. R*-tree의 특성을 이해하기 위해 가장 핵심적인 요소는 바로 MBR(Minimum Bounding Rectangle)인데, MBR은 표현하려는 객체 또는 하위 MBR들을 포함하는 가장 작은 구역을 말하며, 여기서 구역이란 2차원의 경우에는 직사각형, 3차원의 경우에는 직육면체, n차원의 경우에는 2 x n의 지표로서 정의될 수 있는 공간을 말한다. MBR이라는 요소로서 R*-tree에서는 공간상의 객체를 표현하고 정의할 수 있다.

III. 실험 결과

본 논문에서는 JPEG 압축된 2,920여장의 영상으로 DB를 구성하고, 제안된 검색 기법을 평가 하였다. 실험 DB는 실내영상, 식물, 동물, 풍경, 드라마, 뉴스 등의 다양한 영상으로 구성하였다. 그림 7과 그림 8은 실험에 사용된 기본 영상들과 검색 시스템을 보여주고 있다.

본 논문에서 제안한 영상 검색의 정확성은 recall, precision과 ANMRR(Average of the Normalized Modified Retrieval Rank)을 통해 이루어진다. 먼저, 각 영상 질의 q에 대한 DB에서의 유사 영상의 개수를 NG(q)라 한다. 질의에 대한 검색 시 DB에서 찾은 최초 M개의 검색 결과 중에 유사 영상으로 기(既) 지정된 표준 기저 영상(ground truth images)과 비교하여 정확하게 찾은 영상의 수를 n_c , 놓친 영상의 수를 n_m , 잘못 찾은 영상의 수를 n_f 라 정의할 때, 질의 영상 q에 대한



그림 7. DB 영상
Fig. 7. DB images.

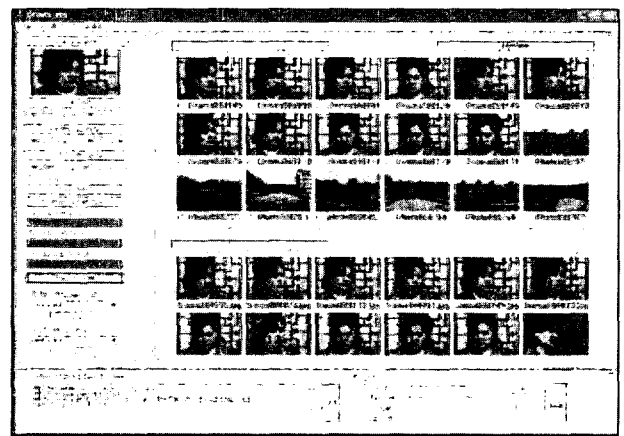


그림 8. 검색 시스템
Fig. 8. Retrieval system.

precision과 recall은 각각 다음과 같은 수식으로 얻는다.

$$Precision = \frac{n_c}{n_c + n_f} = \frac{n_c}{M} \quad (9)$$

$$Recall = \frac{n_c}{n_c + n_m} = \frac{n_c}{NG(q)} \quad (10)$$

수식에서 알 수 있듯이 precision은 사용자의 질의를 통해 실제로 검색된 결과 영상과 사용자의 질의와 관계된 DB에서의 표준 기저 영상과의 비율을 뜻하며, recall은 표준 기저 영상과 실제 검색을 통해 사용자에게 보여진 영상과의 비율을 나타내고 있다. 즉, precision은 검색 정확성을 측정하는 기준이며, recall은 적합한 자료를 얼마나 많이 검색했느냐를 보여주는 지표이다. Precision과 recall은 모두 [0.0~1.0]의 값을 가지며 높은 값을 가질수록 좋은 검색 성능을 나타낸다.

또한, 다른 성능 평가 기준으로서 MPEG-7에 정의된 ANMRR을 이용한다. ANMRR는 모든 질의에 대하여 표준화된 수정 검색 순위의 평균을 나타낸다^[14]. 먼저, 질의 q에 대한 평균 순위를 수식 11로 표현한다.

$$AVR(q) = \sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)} \quad (11)$$

여기서 Rank(k)는 k번째 질의에 대한 검색 순위를 의미한다. 이러한 평균 순위 AVR은 다음의 수식 12로 수정된다.

$$MRR(q) = AVR - 0.5 - \frac{NG(q)}{2} \quad (12)$$

표준화된 수정 검색 순위는 수식 13으로 정의한다.

$$NMRR(q) = \frac{MRR(q)}{\max\{4 \times NG(q), \max_{q=Q} (NG(q))\} + 0.5 - 0.5 * NG(q)} \quad (13)$$

여기에서, Q는 모든 질의 영상의 수를 표현한다. 표준화된 수정 검색 순위는 [0.0~1.0]의 값을 가지며 recall, precision과는 반대로 낮은 값을 가질수록 높은 검색 성능을 나타낸다. 최종적으로, ANMRR은 모든 질의 영상에 대한 NMRR의 평균값을 의미하며 다음과 같이 계산한다.

$$ANMRR(q) = \frac{1}{Q} \sum_{q=1}^Q NMRR(q) \quad (14)$$

우리는 제안된 방법의 성능 비교를 위해 CCV와 correlogram, 그리고 CCV와 correlogram의 조합 (AC-CCV) 등의 방법을 이용한다. 일반적인 correlogram을 이용하면 많은 계산량이 필요하기 때문에 본 실험에서는 {1,3,5,7}의 공간 거리에 대하여 auto-correlogram을 사용하였다. CCV와 correlogram은 각각 128개의 bin을 이용하며, AC-CCV 방법과 제안된 ESD 방법은 각각 256개의 bin을 이용한다. CCV에서의 bin은 64개의 양자화된 색상과 색상의 인접성 여부를 고려하여 이용된다. 또한, correlogram에서의 bin은 32개의 양자화된 색상과 {1,3,5,7}의 색상 거리에 대한 표현으로서 이용된다. AC-CCV는 CCV와 correlogram으로부터 각각 특징을 추출하여, 선형적인 유사도 측정기에 의해 두 특징값을 조합하는 방법이다. 이러한 방법들에 대한 공정한 비교를 위하여, 본 실험에서는 모든 질의에 대하여 같은 표준 기저 영상을 이용한다.

표 1과 표 2에서는 제안한 알고리즘과 다른 영상 검색 기법들과의 검색 성능 및 소비 시간 등을 비교하고 있다. 우선 표 1에서는 ANMRR, recall, precision 등의 검색 성능 측정 비교를 통해 제안한 방법이 타 알고리즘들과 비교하여 뛰어난 검색 성능을 가지고 있다는 것

표 1. 여러 검색 기법과의 성능 비교

Table 1. Comparison for retrieval performance with other retrieval methods.

Method	ANMRR	Recall	Precision
Correlogram	0.425	0.10	0.38
CCV	0.382	0.18	0.47
AC-CCV	0.279	0.24	0.47
Proposed (ESD)	0.223	0.28	0.47

표 2. 여러 검색 기법과의 검색 시간 비교 (ms)

Table 2. Comparison of the elapsed time for retrieval with other retrieval methods (ms).

Method	Off-line		On-line	
	Feature extraction	Indexing	Feature reading	Indexing
Correlogram	43	0.07	0.12	0.05
CCV	27	0.06	0.15	0.06
AC-CCV	63	-	-	-
Proposed (ESD)	52	0.14	0.17	0.13

을 보여준다. Correlogram 방법은 일정 정도 균일한 패턴을 가지는 영상에서는 비교적 좋은 성능을 보이나, 일반적인 자연 영상에서는 그다지 좋은 결과를 나타내지 못한다. CCV 방법의 경우 전통적인 히스토그램 기법에 의한 영상 검색 기법보다는 훨씬 좋은 성능을 가지지만, 영상에 포함되는 각 색상에서의 영역이 대부분 일정 크기 이상으로 그룹화 되어있는 경우에 있어서는, 전통적인 히스토그램 기법과 성능에 있어 큰 차이를 보이지 않게 된다. AC-CCV와 제안한 ESD방법의 경우는 모두 좋은 성능을 보이지만, AC-CCV의 경우 전역적인 영상에 대해서 CCV와 correlogram을 각각 적용하여 특징을 추출하는 반면에, 제안하는 방법은 영상을 화소의 주파수에 따라 분할하여 각 분할된 영상의 특징에 의해 CCV와 correlogram을 따로 적용시킴으로써 특징 추출 시간을 줄임과 동시에 향상된 검색 성능을 나타내고 있다. 전체 영상에 CCV와 correlogram을 적용하여 특징을 추출할 때 걸리는 시간을 각각 ft_{cv} 와 ft_{ac} 라고 할 때, AC-CCV의 특징 추출 시간은 $O(ft_{cv} + ft_{ac})$ 인 반면, 제안된 ESD 방법에서는 $O(ft_{cv} + (1-\sigma)ft_{ac})$ 의 시간만이 소요된다. 여기서 σ 는 [0.0~1.0]의 값을 가지며 영상의 전체 화소 중 평탄 화소로서 분류된 화소들의 비율을 나타낸다.

표 2에서는 오프라인과 온라인 상으로부터 여러 방법에 의한 특징 추출 및 인덱스 시간 등을 비교하고 있

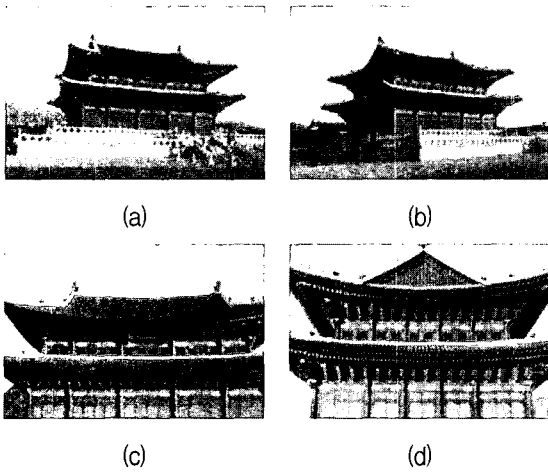


그림 9. 검색 결과: (a) Correlogram (CG): 1, CCV (CV): 1, AC-CCV (AC): 1, Proposed (P): 1 (b) CG: 10, CV: 7, AC: 4, P: 6 (c) CG: 13, CV: 5, AC: 7, P: 5 (d) CG: 23, CV: 14, AC: 13, P: 11

Fig. 9. Retrieval results with rank.

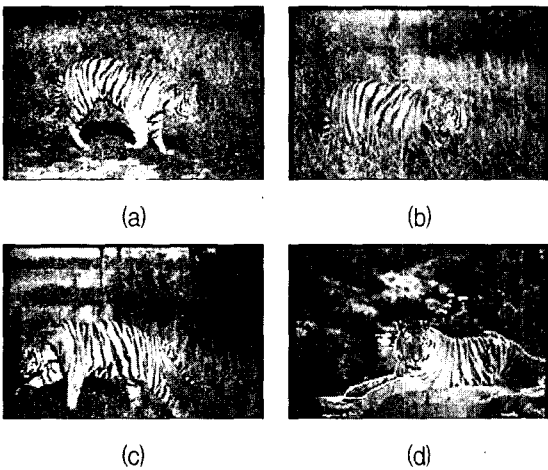


그림 10. 다른 검색 결과: (a) Correlogram (CG): 1, CCV (CV): 1, AC-CCV (AC): 1, Proposed (P): 1 (b) CG: 4, CV: 2, AC: 2, P: 2 (c) CG: 11, CV: 7, AC: 8, P: 6 (d) CG: > 30, CV: 17, AC: 15, P: 12

Fig. 10. Additional retrieval results with rank.

다. 오프라인 상에서 하나의 영상에 대한 특징 추출 시간을 비교할 때, 제안한 기법이 CCV나 correlogram기 법보다는 약간 느린 처리 시간을 가지지만, AC-CCV와 비교할 때는 15% 정도의 빠른 처리 시간을 보인다. 특히 온라인 상에서 비교할 경우, 처음 추출된 특징 정보가 이미 R*-tree에 의해 정보 색인이 되어 있기 때문에, 결과 검색된 최초의 M개의 영상에 대한 특징 추출 시간은 저장된 자료를 읽기 위한 시간만이 필요하게 된다. 결과적으로, 모든 검색 기법에 대해서 1ms 이하의 값을 가지므로 오프라인 상에서의 CCV나 correlogram

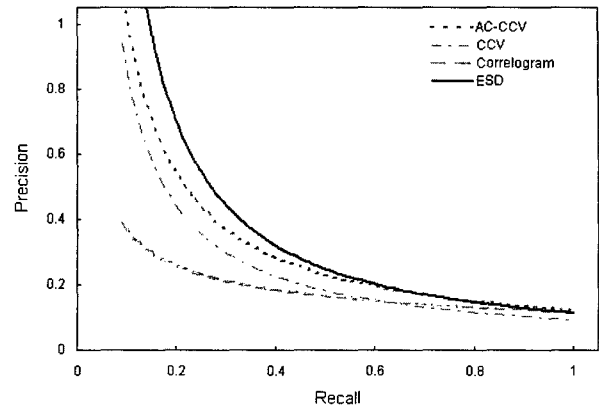
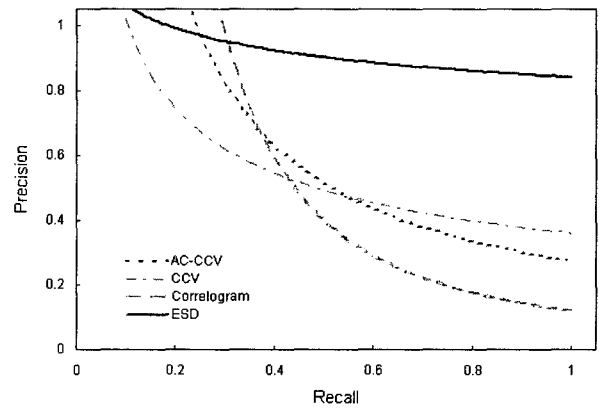
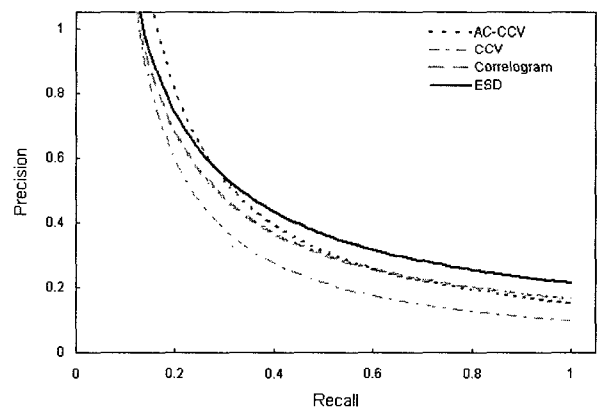


그림 11. Recall과 precision

Fig. 11. Recall and precision.



(a)



(a)

그림 12. Recall과 precision: (a) for 그림7, (b) for 그림8

Fig. 12. Recall and precision:

(a) for Fig. 7, (b) for Fig. 8.

에 대한 제안 기법에서의 특징 추출 시간의 증가는 사실상 큰 의미를 갖지 않는다. 추출된 특징에 대한 인덱스 시간은 온라인이나 오프라인 상에서 같은 값을 갖고, 약 1ms 이하의 작은 시간만이 소요된다.

그림 9는 카메라 줌과 시점 변화 등을 포함하고 있는

관련 영상에 대한 검색 결과를 보여준다. 그림 9-(a)는 질의 영상이고, 그림 9-(b)~그림 9-(d)는 검색된 결과이다. 제안된 방법과 여러 다른 검색 기법들과의 비교를 위해 각 영상마다 검색된 순위를 숫자로 표기하였다. 질의 영상에 대한 검색 결과에서 보이는 바와 같이, 제안된 방법은 카메라 줌에 더 강건한 결과를 보여주는 반면에, 다른 알고리즘들은 그러한 변화에 대하여 상대적인 민감함을 가지고 있다. 그림 10은 주변 공간에서의 새로운 변화를 가지는 다른 질의 영상에 대한 결과를 보인다.

그림 11과 그림 12에서는 recall과 precision에 기초한 영상 검색 성능 비교 그래프를 표현하고 있다. 그림 11에서의 각각의 값은 전체적인 성능 비교를 위한 평균값을 나타낸다. 그래프는 제안한 방법이 CCV나 correlogram, AC-CCV 등의 방법과 비교하여 좀 더 우월한 성능을 가진다는 것을 보여주고 있으며, 그림 11에서 보듯이, 영상 검색 결과는 제안된 방법, AC-CCV, CCV, 그리고 correlogram의 순서로 우수한 성능을 보인다. 그림 12-(a)와 그림 12-(b)는 각각 그림 9과 그림 10에 대한 recall과 precision 그래프를 보여주고 있다.

VI. 결 론

본 논문에서, 우리는 color vector angle 방법을 통해 영상에서의 에지를 추출하고, 이를 이용하여 영상의 평탄 영역과 에지 영역 각각에 CCV와 correlogram 기법을 적용시키는 새로운 ESD 방법을 제안하였다. 제안하는 방법에서의 correlogram은 에지 픽셀만을 이용하기 때문에, 영상 에지 상의 색상 상관관계를 효과적으로 표현할 수 있고 영상에서의 많은 색상 변화 등에 대해서도 강건함을 제공할 수 있다. CCV 방법은 영상에서 추출된 평탄 영역에서의 전역적인 색상 분포에 대하여 효과적으로 표현하여 준다. 제안된 방법은 영상의 많은 부분에서 색 변화에 일어나더라도 만족할 만한 검색 성능을 제공하며, 에지 히스토그램 등을 이용하여 부분적인 영상의 검색에도 적용 가능하다. 향후 제안된 방법에서 사용된 색상 양자화 등에 대한 최적화 연구가 필요하며, 이에 따른 성능 향상도 기대된다.

참 고 문 헌

[1] M. Flickner et al., "Query by image and video

- content: The QBIC system," *IEEE computer*, vol. 28, no. 9, pp. 23-32, 1995.
- [2] V. Ogle and M. Stonebraker, "Chabot: Retrieval from a relational database of images," *IEEE computer*, vol. 28, no. 9, pp. 40-48, 1995.
- [3] J. R. Smith and S.-F. Chang, "VisualSEEK: A fully automated content-based image query system," in *ACM Multimedia Conf.*, 1996.
- [4] A. Pentland, R. Picard, and S. Sclaroff, "Photobook: Content-based manipulation of image databases," *IJCV*, vol. 18, no. 3, pp. 233-254, 1996.
- [5] M. Swain and D. Ballard, "Color indexing," *Int. J. Comput. Vis.*, vol. 7, no. 1, pp. 11-32, 1991.
- [6] J. Hafner, H.S. Sawhney, W. Equitz, M. Flickner, and W. Niblack, "Efficient color histogram indexing for quadratic form distance functions," *IEEE Transactions on PAMI*, vol. 17, num. 7, pp. 729-736, 1995.
- [7] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlograms," *CVPR*, pp. 762-768, 1997.
- [8] J. Huang, S. R. Kumar, and M. Mitra, "Combining supervised learning with color correlograms for content-based image retrieval," in *Proc. 5th ACM Multimedia Conf.*, pp. 325-334, 1997.
- [9] G. Pass and R. Zabih, "Histogram refinement for content-based image retrieval," *IEEE WACV*, pp. 96-102, 1996.
- [10] R.D. Dony and S. Wesolkowski, "Edge detection on color images using RGB vector angle," in *Proc. Conf. Signals, Systems & Computers*, pp. 687-692, 1998.
- [11] J. Huang, S. R. Kumar, M. Mitra, and W. J. Zhu, "Spatial color indexing and applications," *ICCV*, pp. 602-607, 1998.
- [12] A. Guttman, "R-trees: a dynamic index structure for spatial searching," *Proc. ACM SIGMOD*, pp. 47-57, 1984.
- [13] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger, "The R*-tree: An efficient and robust access method for points and rectangles," *Proc. ACM SIGMOD*, pp. 322-331, 1990.
- [14] "MPEG vancouver meeting," *ISO/IEC JTC1/SC29/WG11, Experimentation Model Ver.2.0, Doc. N2822*, 1999.

— 저 자 소 개 —

김 낙 우(학생회원)

전자공학회논문지
제 41권 SP편 제 5호 참조

김 태 용(정회원)

전자공학회논문지
제 41권 SP편 제 5호 참조

최 종 수(정회원)

전자공학회논문지
제 41권 SP편 제 5호 참조