

논문 2004-41SP-1-3

내용기반 영상검색에서 색과 질감을 나타내는 채널색에너지 (Channel Color Energy Feature Representing Color and Texture in Content-Based Image Retrieval)

정재웅^{*}, 권태완^{**}, 박섭형^{***}

(Jae Woong Jung, Tae Wan Kwon and Seop Hyeong Park)

요약

내용기반 영상검색 분야에서 색, 질감, 모양 등과 같은 영상의 시각적인 내용을 표현하기 위하여 수치화한 특징들이 많이 제안되었다. 이런 특징들은 모두 독립적이라고 가정하기 때문에 한 특징 벡터를 추출할 때는 다른 특징들과의 상관성을 전혀 고려하지 않는다. 이 논문에서는 색과 질감 사이의 관계를 고려하여 새로운 CCE(channel color energy) 특징을 제안한다. 자연 영상을 대상으로 한 실험결과를 분석한 결과 제안하는 방법이 정규 가중거리 비교 방법과 SCFT(sequential chromatic Fourier transform) 기반 색 질감 방법에 비해 우수한 성능을 보이는 것을 확인할 수 있었다.

Abstract

In the field of content-based image retrieval, many numerical features have been proposed for representing visual image content such as color, texture, and shape. Because the features are assumed to be independent, each of them is extracted without any consideration of the others. In this paper, we consider the relationship between color and texture and propose a new feature called CCE(channel color energy). Simulation results with natural images show that the proposed method outperforms the conventional regular weighted comparison method and SCFT(sequential chromatic Fourier transform)-based color texture method.

Keywords: 내용기반 영상검색, 색특징, 질감특징, Gabor, 채널색에너지

I. 서 론

컴퓨터와 인터넷의 발전으로 멀티미디어 데이터의 양이 급속도로 증가하고 있다. 대량의 멀티미디어 데

이터를 효과적으로 활용하기 위해서는 데이터의 저장 및 검색을 위한 효율적인 방법이 개발되어야 할 것이다. 이런 요구에 따라 여러 방면의 연구가 이루어지고 있으며, 그 중 영상의 검색, 특히 내용기반 영상검색에 관한 연구가 활발히 진행되고 있다^[1].

내용기반 영상검색이란 영상이 가지고 있는 저수준의 특징, 예를 들면 색(color), 질감(texture), 모양(shape), 움직임(motion) 등을 수치화하여 이를 검색에 이용하는 방법이다. 이 방법은 컴퓨터를 이용해 데이터베이스를 자동으로 구축할 수 있으며, 의미가 불분명한 영상의 검색을 가능하게 하는 등 텍스트기반 영상검색의 단점을 극복할 수 있는 방법이다.

* 정회원, (주)와이더엔닷컴 (WiderThan.com Co., Ltd.)

** 학생회원, 한림대학교 대학원 전자공학과

(Department of Electronics Engineering, Graduate School, Hallym University)

***정회원, 한림대학교 정보통신공학부

(Division of Information and Telecommunications Engineering, Hallym University)

※ 이 논문은 2002년도 한림대학교 교비연구비(HRF-2002-42)에 의하여 연구되었음.

접수일자 : 2003년9월5일, 수정완료일 : 2003년12월26일

내용기반 영상검색에 사용할 수 있는 특징으로는 색, 질감, 모양 등이 있으며, 대부분의 영상이 여러 가지 특징을 동시에 포함하므로 영상 검색에 다중 특징을 사용하는 것이 바람직하다. 내용기반 영상검색에 가장 널리 사용되는 특징에는 색과 질감특징이 있으며, 이들을 수치화할 때 색 히스토그램과 가보 필터가 각각 가장 많이 사용되고 있다. 다중 특징을 사용하여 영상간의 유사도를 측정하는 가장 일반적인 방법은 개별적으로 추출된 각 특징들의 중요도에 따라 서로 다른 가중치를 부여하는 정규 가중거리 비교방법(RWC: regular weighted comparison)^[2]이다.

그리나 다중 특징을 이용한 기존의 방법들은 각 특징을 개별적으로 추출하여 검색에 사용하기 때문에 영상에서 특징간의 상관관계를 고려하지 못하는 한계가 있다. 예를 들면 특정 질감 영역에 위치한 색 정보 등은 기존의 방법으로는 표현할 수 없다. 이런 단점을 보완하기 위하여 영상을 여러 개의 영역으로 분할하고 각 영역에 대한 특징을 추출하는 방법을 사용하기도 한다^[3]. 그러나 일반적인 영상의 경우 영역을 분할하는 것이 매우 어렵기 때문에 이 방법을 사용하는 것이 쉽지만은 않다.

이 논문에서는 색과 질감 사이의 관계를 고려하여 새로운 특징 벡터인 CCE(channel color energy)를 제안한다. 제안하는 특징은 영상을 국소 주파수대역으로 분할하여 각 채널마다 특정 색이 포함된 정도를 표현한다.

이 논문의 구성은 다음과 같다. II장에서는 색과 질감 특징을 이용한 기존의 연구들을 살펴보고, III장에서는 CCE를 사용한 내용기반 영상검색 방법에 대해 알아보겠다. IV장에서는 실험 및 검토, 마지막으로 V장에서는 결론을 맺는다.

II. 색과 질감 특징을 이용한 RWC

색은 사람이 영상을 구별할 수 있게 하는 가장 중요한 요소 중 하나이다. 그러므로 색은 영상의 기본적인 내용 특징이 될 수 있으며, 내용기반 영상검색에서 가장 많이 사용되는 특징 중 하나이다. 내용기반 영상검색에서 색 특징을 표현하는 방법에는 평균 색, 색 히스토그램(color histogram)^[4], 색 모멘트(color moments), 국부 색 히스토그램(local color histogram), 색 코렐로그램(color correlogram)^[5], 누적 색 히스토그램(cumulative color histogram) 등이 있다.

질감은 색과 더불어 사람이 영상을 인식할 때 중요하게 작용하는 요소이다. 질감은 재질의 표면 차이에서 받는 느낌으로 거칠, 부드러움, 규칙성에 관한 척도라고 할 수 있다. 색이 영상의 화소 기반 특성이라 하면 질감은 공간적인 지역의 개념을 포함한다. 질감 특징에 자주 사용되는 방법으로는 이웃 화소(pixel neighborhood) 특성을 이용한 방법, 히스토그램 기반(histogram based) 방법, 동시 발생 행렬(co-occurrence matrix) 방법^[6], 타무라(Tamura) 질감 표현법^[7], 가보 필터(Gabor filter)^[8]를 사용한 방법 등이 있다.

다중 특징을 사용하여 영상 검색을 수행할 경우 일반적으로 RWC 방법^[2]을 사용한다.

$A = \{a_1, a_2, \dots, a_n\}$ 를 데이터베이스 내 영상의 특징, $Q = \{q_1, q_2, \dots, q_n\}$ 를 질의 영상의 특징이라 하자. 그리고 $d_i(a_i, q_i)$ 를 a_i 와 q_i 사이의 거리라고 정의하자. 그리고 두 영상 A 와 Q 의 정규 가중 거리를 식 (1)과 같이 정의한다.

$$d(A, Q) = \sum_{i=1}^n w_i d_i(a_i, q_i) \quad (1)$$

여기에서 w_i 는 $d_i(a_i, q_i)$ 에 대한 가중치로 $\sum_{i=1}^n w_i = 1$ 이다.

만약 RWC 방법에 색과 질감 특징을 사용할 경우에 최종 거리는 식 (2)와 같다.

$$d(A, Q) = ad_c + (1 - a)d_t \quad (2)$$

여기에서 a 는 색 특징에 대한 가중치이며 d_c 는 영상 A 와 Q 의 색 특징의 거리이며, d_t 는 영상 A 와 Q 의 질감 특징의 거리이다.

다중 특징을 영상검색에 사용하려면 각각의 특징 벡터에 적당한 가중치를 부여할 필요가 있으며 영상의 특징에 맞는 가중치의 선택은 매우 중요한 문제이다. 그러나 대부분의 경우 영상마다 뚜렷하게 구분이 되는 특징이 다르기 때문에 모든 영상에 적용할 수 있는 가중치를 찾는 것은 불가능하다. 최적 가중치를 구하기 위하여 신경망, 적합성 피드백(relevance feedback) 등을 이용하기도 하지만 이러한 방법들을 사용한 기존의 방법들 중 대부분은 개별적인 저 수준 특징 벡터 각각에 의존하여 가중치를 조절하게 되므로 특징간의 상관관계를 표현하기 어렵다는 한계가 있다.

III. 채널 색 에너지

영상 분석이나 영상 분류 분야에서 색과 질감은 대단히 중요한 특징으로 인식되어 왔지만 대부분의 연구자들은 이 둘이 완전히 독립적이라고 가정하였다. 따라서 질감 특징은 컬러 영상이 아닌 흑백 영상에서 추출하는 것이 일반적이었다. 그러나 최근에 영상 분류에 색질감(color texture) 특징을 이용하면 흑백 질감 특징만을 사용하는 것보다 우수한 성능을 얻을 수 있다는 연구 결과가 발표되는 등^[10], 색과 질감 사이의 상관성을 활용하기 위한 연구가 진행되고 있다.^[10]에서는 기존에 흑백 영상에만 적용했던 가보 필터를 색 영상에 적용하기 위하여 푸리에 변환을 확장한 SCFT(sequential chromatic Fourier transform)를 사용하였고, 각 채널 신호의 크기 에너지와 위상 에너지를 색질감 특징으로 사용하였다. 이 경우에 푸리에 변환을 할 영상의 차원이 흑백 영상에 비해서 3 배 증가하기 때문에 가보 필터링에 필요한 계산량도 3 배 증가하는 단점이 있다.

이 논문에서 제안한 CCE는 흑백 영상에만 가보 필터링을 적용하여 채널 별 영상을 구한 후에, 원 색 영상에서 동일한 색을 가진 화소들을 조사하여 그 화소들과 같은 위치에 있는 채널 영상의 화소들의 에너지를 색질감 특징으로 정의하였다. 즉, CCE는 가보 필터를 사용하여 추출한 흑백 영상의 질감 특징과 색 히스토그램으로 표현하는 색 특징 사이의 관계를 이용한 것이다. CCE를 구하는 구체적인 방법은 다음과 같다.

먼저, 입력 색 영상에서 흑백 영상을 추출한 후, 이를 다시 24 개의 가보 필터와 하나의 저대역 통과 필터를 사용하여 25 채널 신호로 분리한다. 원본 색 영상을 RGB 색 공간에서 각 색 채널 별로 4 개의 대표값으로 균일 양자화 한 후에, 25 개의 채널 영상 모두에서 원본 영상의 양자화된 색 화소와 같은 위치에 있는 채널 영상의 화소들의 절대값의 합을 구한다. 이 값을 채널 색 에너지라 부른다. 이 채널 색 에너지를 특징 벡터로 정의하고, 특징 벡터 사이의 거리는 L_1 놈(norm)을 사용하여 계산한다.

가보 필터에 기반 하는 질감특징을 얻기 위한 2차원 가보 필터 $g(x, y)$ 는 다음과 같다^[9].

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp -\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j W_x, \quad (3)$$

여기에서 $\sigma_u = \frac{1}{2\pi\sigma_x}$, $\sigma_v = \frac{1}{2\pi\sigma_y}$ 이다.

가보 함수의 비직교성(nonorthogonality)으로 인해 필터링 된 영상에는 중복된 정보가 존재한다. 따라서 중복된 정보를 줄이기 위해^[9]에서 제안한 4개의 척도(scale)와 6개의 방위(orientation)를 갖는 가보 필터를 사용하였다.

제안하는 특징벡터 CCE는 가보 필터에 의해 원본 영상을 다중 채널로 분리한 후 공간 영역(spatial domain)에서 특징 벡터를 추출한다. 만약 원본 영상의 특정 영역에 DC가 집중되어있는 경우 척도와 방위에 의해 만들어진 필터만으로는 해당 영역의 특징을 잘 추출할 수가 없다. 따라서 이런 특징을 추출하기 위하여 식 (4)과 같이 정의한 별도의 저대역 통과 필터를 사용하였다.

$$g_{dc}(x, y) = \left(\frac{w}{2\pi\sigma_x\sigma_y} \right) \exp -\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right). \quad (4)$$

$$w = \frac{\sigma_{ul}}{\sigma_{ud}}, \quad (5)$$

$$\sigma_{ud} = -\sigma_{ul} + \frac{U_l}{\sqrt{2\ln 2}}, \quad (6)$$

$$\sigma_{vd} = \sigma_{ud}, \quad (7)$$

여기에서 σ_{ud} 와 σ_{vd} 는 저대역 통과 필터의 u, v 축 표준편차이며, σ_{ul} 은 가보 필터 중에서 중심 주파수가 U_l 에 해당하는 필터의 방사상의 표준편차이다. w 는 필터의 에너지를 가보 필터와 동일하게 하기 위한 척도율(scale factor)이며, g_{dc} 는 에너지가 정규화된 2차원 가우시안 함수이다. 그리고 $\sigma_x = \sigma_y = \frac{1}{2\pi\sigma_{ud}}$ 이다.

색 영상으로부터 질감 특징을 추출하기 위해 원본 영상의 색 공간을 HSV(Hue Saturation Value) 색 공간으로 변환한다. 그리고 변환된 색 공간의 색 채널 중 V 채널 영상만을 이용하여 가보 필터링 및 저대역 통과 필터링을 수행한다. 계산량을 줄이기 위해 주파수 공간에서 필터링을 수행하였다.

색 특징을 표현하기 위해서는 먼저 색 공간을 정의 해야 한다. 이 논문에서 제안하는 색 특징은 색 히스토그램 분포이다. 이 때, 특징 벡터의 차원을 줄이기 위해 영상 검색 성능이 크게 저하되지 않는 범위에서 색의 개수를 줄일 필요가 있다. 이 논문에서는 각 색 채널 신호를 4 개의 값으로 균일 양자화 하였다.

이제 각 색이 특정 가보 채널에서 가지는 에너지를 계산한다. I 는 모든 색 채널을 가지는 $n_1 n_2$ 크기의 원본 색 영상이며, $\mathbf{p} = (x, y)$ 는 원본 영상에서 좌표, $i(\mathbf{p})$ 는 좌표 \mathbf{p} 에서의 화소 값이라 하자. I 의 색 공간이 m 개의 색 (c_1, c_2, \dots, c_m)으로 양자화 되었다고 가정하면 척도 s , 방위 o 를 가지는 가보 채널의 색 에너지는 다음과 같이 정의 된다.

$$\mathbf{e}_{so} = [e_{so}(c_1), e_{so}(c_2), \dots, e_{so}(c_m)], \quad (8)$$

$$e_{so}(c_i) = \sum_{\{\mathbf{p} | i_q(\mathbf{p}) = c_i\}} |f_{so}(\mathbf{p})|, \quad (9)$$

여기에서 i_q 는 화소 i 를 양자화한 값을 의미한다. F_{so} 를 $n_1 n_2$ 크기의 가보 필터링된 공간영역의 영상이라고 하면, $f_{so}(\mathbf{p})$ 는 영상 F_{so} 에서 좌표 \mathbf{p} 의 화소 값을 의미한다. $e_{so}(c_i)$ 는 척도 s , 방위 o 에 해당하는 필터에 의해 필터링 된 영상의 색 c_i 에 대한 에너지를 의미한다.

그리고 저대역 통과 영상의 색 에너지는 다음과 같이 정의된다.

$$\mathbf{e}_{lp} = [e_{lp}(c_1), e_{lp}(c_2), \dots, e_{lp}(c_m)], \quad (10)$$

$$e_{lp}(c_i) = \sum_{\{\mathbf{p} | i_q(\mathbf{p}) = c_i\}} |f_{lp}(\mathbf{p})|, \quad (11)$$

여기에서 F_{lp} 를 $n_1 n_2$ 크기의 저대역 통과 필터링 된 영상이라고 하면, $f_{lp}(\mathbf{p})$ 는 영상 F_{lp} 에서 좌표 \mathbf{p} 의 화소 값을 의미한다. $e_{lp}(c_i)$ 는 저대역 영상의 색 c_i 의 에너지를 의미한다.

모든 채널에 대한 CCE는 다음과 같이 정의된다.

$$\mathbf{e} = [e_{lp}, e_{00}, e_{01}, \dots, e_{(s-1)(o-1)}]. \quad (12)$$

CCE는 원본 영상의 휘도에 따라 해당 휘도의 값만큼 서로 다른 가중치를 부여받게 된다. 따라서 해당 휘도 값으로 정규화를 해야 한다.

원본 영상 I 의 휘도 영상을 I_g 라 하고, $i_g(\mathbf{p})$ 를 영상 I_g 에서 좌표 \mathbf{p} 의 화소 값, $i'_g(\mathbf{p}) = i_g(\mathbf{p}) + 1$ 이라 하자. 그리고 I'_g 를 필터링한 결과 영상을 F'_{so} 라 할 때, 특정 필터에 대한 색 에너지는 식 (18)과 같다.

$$\mathbf{e}'_{so} = [e'_{so}(c_1), e'_{so}(c_2), \dots, e'_{so}(c_m)], \quad (13)$$

$$e'_{so}(c_i) = \sum_{\{\mathbf{p} | i_g(\mathbf{p}) \in c_i\}} |\tilde{f}_{so}(\mathbf{p})|, \quad (14)$$

$$\tilde{f}_{so}(\mathbf{p}) = \frac{f'_{so}(\mathbf{p})}{i'_g(\mathbf{p})}. \quad (15)$$

여기에서 $f_{so}(\mathbf{p})$ 는 영상 F'_{so} 에서 좌표 \mathbf{p} 의 화소 값이다. 또한 e_{lp} 를 정규화한 e'_{lp} 도 이와 유사한 방법으로 계산한다.

정규화 된 CCE 특징은 다음과 같다.

$$\mathbf{e} = [e'_{lp}, e'_{00}, e'_{01}, \dots, e'_{(s-1)(o-1)}]. \quad (16)$$

두 영상의 CCE 특징 벡터들의 유사도로 가우시안 정규화된 L_1 놈을 사용하였다.

$$d(I, Q) = |e'_{dc} - e^Q_{dc}| + \sum_{i=0}^{S-1} \sum_{j=0}^{O-1} |e'_{ij} - e^Q_{ij}|, \quad (17)$$

$$e'_{ij} - e^Q_{ij} = \sum_{l=1}^m \frac{|e'_{ij}(c_l) - e^Q_{ij}(c_l)|}{\sigma(e'_{ij}(c_l))}. \quad (18)$$

$$e'_{lp} - e^Q_{lp} = \sum_{l=1}^m \frac{|e'_{lp}(c_l) - e^Q_{lp}(c_l)|}{\sigma(e'_{lp}(c_l))}. \quad (19)$$

여기에서 Q 는 질의영상, I 는 데이터베이스 내의 영상이다. m 은 양자화 된 색의 개수이며, S 는 전체 척도, O 는 전체 방위 개수이다. $e^Q_{ij}(c_l)$ 와 $e'_{ij}(c_l)$ 는 영상 Q 와 I 에 대하여 척도 i , 방위 j 인 채널에서 색 c_l 의 에너지를 나타낸다. 그리고 $\sigma(e'_{ij}(c_l))$ 와 $\sigma(e'_{lp}(c_l))$ 는 전체 데이터베이스에서 $e'_{ij}(c_l)$ 와 $e'_{lp}(c_l)$ 의 표준편차이다.

IV. 실험 및 성능 평가

제안하는 알고리듬의 성능을 평가하기 위해 색과 질감 특징의 정규 가중거리 비교 방법(RWC)과^[10]에서 제안한 SCGT (sequential chromatic Gabor transform)를 CCE와 비교하였다.

RWC를 위한 색 특징과 질감 특징으로 색 히스토그램 분포와 가보 필터를 이용한 방법을 각각 사용하였다. 색 히스토그램은 R, G, B 채널이 각각 4개의 범으로 균일하게 양자화 된 RGB 색 공간을 사용하였으며, 가보 필터의 척도와 방위는 각각 4와 6을 사용하였다. RWC를 위한 두 특징의 가중치는 반복적인 실험을 통해서 구하였다.

실험 영상으로는 15,000여장의 코렐(Corel) 영상과 5,400여장의 MPEG 실험 영상, 2,500여장의 VisTex 영상을 사용하였다. 질의 영상으로는 해당 질의 영상에 적합한 결과 영상을 뚜렷하게 구분할 수 있는 18장을 선택하였다. 표 1은 실험에 사용한 평가 영상 18장을 정리한 것이다.

평가 척도로는 정확률(precision)과 재현률(recall)을 사용하였으며 다음과 같이 정의된다^[11].

표 1. 실험에 사용한 영상.

Table 1. Test images used in the experiments.

corel106009	corel107021	corel108036
corel114005	corel75084	corel40093
MPEG00041	MPEG00148	MPEG02933
MPEG04997	Bark.0000	Brick.0004
Fabric.0017	Fabric.0018	Food.0001
Food.0008	Flower.0002	Leaves.0011

$$\text{정확률} = \frac{\text{검색된 결과 영상 중 적합한 영상의 개수}}{\text{검색된 영상의 전체 개수}}$$

$$\text{재현률} = \frac{\text{검색된 결과 영상 중 적합한 영상의 개수}}{\text{전체 영상 중 적합한 영상의 개수}}$$

표 2, 3, 4는 각각 검색 영상의 수를 10, 30, 50 장으로

했을 때 RWC, SCGT, CCE의 정확률과 재현률을 비교한 것이다. 각 입력 영상에 대해서 상대적으로 성능이 우수한 방법의 결과에 영암을 표시하였다. 또한 RWC에 사용되는 가중치는 각 입력 영상에 대해서 정확률과 재현률의 값을 최대로 하는 값을 실험에 의해서 찾은 값이다. 이 가중치를 찾을 때 0과 1사이의 0.1 간격의 수를 사용하여 얻은 결과 중에서 최적의 값을 선정한 것이다. 표에서 영상에 따라서 최적 가중치가 달라지는 것을 볼 수 있으며, MPEG, Corel과 같은 자

표 2. 10 장의 영상을 검색했을 때 RWC, SCGT, CCE의 정확률과 재현률 비교.

Table 2. Comparison of RWC, SGCT, and CCE in terms of precision and recall when 10 images are retrieved.

질의 영상	정확률			재현률			RWC 가중 치
	RWC	SCGT	CCE	RWC	SCGT	CCE	
corel106009	0.2	0.1	0.1	0.03	0.02	0.02	0.9
corel107021	0.9	0.3	0.9	0.09	0.03	0.09	0.7
corel108036	0.8	0.4	0.6	0.08	0.04	0.06	0.9
corel114005	0.8	0.3	1	0.08	0.03	0.1	0.9
corel40093	0.6	0.8	0.5	0.06	0.08	0.05	0.7
corel75084	0.8	0.4	0.7	0.09	0.05	0.09	0.9
mpeg00041	1	0.4	1	0.67	0.27	0.67	0.7
mpeg02933	0.4	0.1	0.5	0.5	0.67	0.83	0.8
mpeg04997	0.8	0.3	0.8	0.89	0.33	0.89	0.7
mpeg00148	1	0.4	1	0.63	0.25	0.63	0.7
Bark.0000	1	0.6	1	0.16	0.09	0.16	0.8
Brick.0004	1	0.9	1	0.16	0.14	0.16	0.7
Fabric.0017	1	1	1	0.16	0.16	0.16	0.8
Fabric.0018	1	1	1	0.16	0.16	0.16	0.8
Food.0001	1	0.4	1	0.16	0.06	0.16	0.8
Food.0008	0.7	0.1	0.6	0.11	0.02	0.09	0.7
Flower.0002	1	1	1	0.16	0.16	0.16	0.8
Leaves.0011	1	0.9	1	0.16	0.14	0.16	0.8
최대값 개수	15	4	13	15	4	13	

표 3. 30 장의 영상을 검색했을 때 RWC, SCGT, CCE의 정확률과 재현률 비교.

Table 3. Comparison of RWC, SGCT, and CCE in terms of precision and recall when 30 images are retrieved.

질의 영상	정확률			재현률			RWC 가중치
	RWC	SCGT	CCE	RWC	SCGT	CCE	
corel106009	0.1	0.17	0.1	0.06	0.08	0.05	0.9
corel107021	0.47	0.23	0.4	0.14	0.07	0.12	0.7
corel108036	0.47	0.2	0.4	0.14	0.06	0.12	0.9
corel114005	0.3	0.17	0.93	0.14	0.05	0.28	0.9
corel40093	0.63	0.77	0.4	0.19	0.23	0.12	0.7
corel75084	0.43	0.27	0.67	0.15	0.16	0.25	0.9
mpeg00041	0.4	0.37	0.47	0.8	0.73	0.93	0.7
mpeg02933	0.1	0.07	0.17	0.67	0.33	0.83	0.7
mpeg04997	0.27	0.13	0.3	0.89	0.44	1	0.7
mpeg00148	0.4	0.2	0.4	0.75	0.38	0.75	0.7
Bark.0000	0.6	0.3	1	0.41	0.14	0.47	0.8
Brick.0004	0.9	0.73	0.83	0.47	0.34	0.39	0.7
Fabric.0017	1	1	1	0.47	0.47	0.47	0.8
Fabric.0018	1	0.93	0.93	0.47	0.44	0.44	0.8
Food.0001	1	0.2	1	0.47	0.09	0.47	0.8
Food.0008	0.3	0.07	0.27	0.14	0.03	0.13	0.7
Flower.0002	1	0.6	1	0.47	0.28	0.47	0.8
Leaves.0011	1	0.53	1	0.47	0.25	0.47	0.8
최대값 개수	10	3	11	10	3	11	

연 영상들에서는 최적 가중치가 0.7~0.9 사이에 분포하며, 텍스처 영상인 VisTex 영상에서는 최적 가중치가 0.7~0.8 사이에 분포하는 것을 알 수 있다. 즉, 자연 영상과 텍스처 영상을 비교하면 질감 특성의 비중이 텍스처 영상에서 조금 큰 것을 알 수 있다. 그렇지만 두 경우 모두 질감 특성보다는 색 특성이 훨씬 중요한 역할을 하는 것을 알 수 있다.

표 4. 50 장의 영상을 검색했을 때 RWC, SCGT, CCE의 정확률과 재현률 비교.

Table 4. Comparison of RWC, SGCT, and CCE in terms of precision and recall when 50 images are retrieved.

질의 영상	정확률			재현률			RWC 가중치
	RWC	SCGT	CCE	RWC	SCGT	CCE	
corel106009	0.12	0.1	0.12	0.1	0.08	0.1	0.9
corel107021	0.38	0.2	0.32	0.19	0.1	0.16	0.7
corel108036	0.38	0.16	0.32	0.19	0.08	0.16	0.9
corel114005	0.32	0.18	0.76	0.16	0.09	0.38	0.9
corel40093	0.66	0.68	0.36	0.33	0.34	0.18	0.7
corel75084	0.34	0.22	0.52	0.21	0.14	0.32	0.8
mpeg00041	0.28	0.24	0.28	0.93	0.8	0.93	0.7
mpeg02933	0.08	0.04	0.1	0.67	0.33	0.83	0.7
mpeg04997	0.16	0.1	0.18	0.89	0.56	1	0.7
mpeg00148	0.24	0.16	0.26	0.75	0.5	0.81	0.7
Bark.0000	0.56	0.24	0.72	0.44	0.19	0.56	0.8
Brick.0004	0.96	0.66	0.7	0.75	0.52	0.55	0.7
Fabric.0017	1	1	1	0.78	0.78	0.78	0.8
Fabric.0018	1	0.7	0.82	0.78	0.55	0.64	0.8
Food.0001	1	0.14	1	0.78	0.11	0.78	0.8
Food.0008	0.2	0.1	0.16	0.16	0.08	0.13	0.7
Flower.0002	1	0.46	1	0.78	0.36	0.78	0.8
Leaves.0011	0.88	0.4	0.94	0.69	0.31	0.73	0.8
최대값 개수	10	2	12	10	2	12	

세 방법의 검색 성능을 비교해 보면, 자연 영상의 경우 CCE의 성능이 우수하며, 텍스처 영상의 경우는 RWC와 CCE가 비슷한 성능을 나타내는 것을 확인할 수 있다. 그러나 RWC는 색과 질감 특징을 선형 결합할 때 모든 입력 영상에 대해서 반복 실험을 통해서 찾은 최적 가중치를 사용한 결과라는 것에 유의해야 할 것이다. 반면에 CCE는 색, 질감 사이의 최적 가중치를 찾기

위한 별도의 노력이 필요 없다는 장점이 있다.

그림 1과 2는 RWC와 SCGT, CCE의 종합적인 성능을 비교한 그래프이다. 이 그림에서 가로축은 검색 영상 수를 의미하고, 세로 축은 실험에서 사용한 모든 입력 영상에 대해서 정확률과 재현률을 구하고 그 평균값을 계산하여 나타낸 것이다. 이 때 RWC에서는 가중치를 0.7, 0.8, 0.9로 고정하여 그 결과를 제시하였다. 그래프에 의하면 SCGT 보다는 RWC가, RWC 보다는 CCE의 성능이 우수한 것을 알 수 있다.

V. 결 론

이 논문에서는 색과 질감을 동시에 고려할 수 있는 새

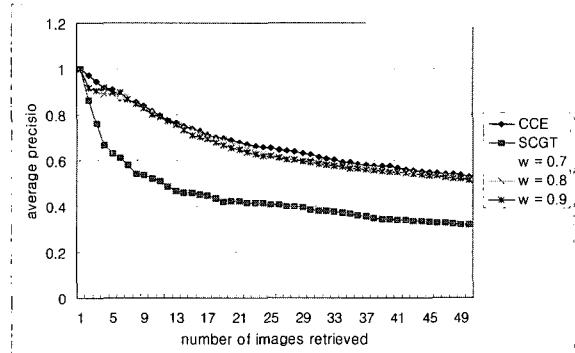


그림 1. RWC, SCGT, CCE의 평균 정확률 비교.
Fig. 1. Comparison of RWC, SGCT, and CCE in terms of precision.

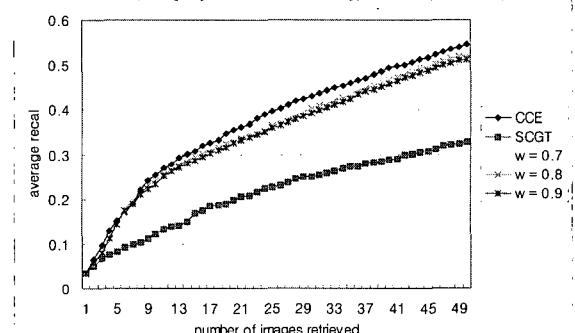


그림 2. RWC, SCGT, CCE의 평균 재현률 비교.
Fig. 2. Comparison of RWC, SGCT, and CCE in terms of recall.

로운 특징벡터인 CCE를 제안하였다. CCE는 원본 영상의 흑백 영상을 가보 필터에 의해 다중 채널로 분리한 후 각 채널 별 색의 에너지로 정의된다.

제안한 특징 벡터의 성능을 평가하기 위해 색과 질

감 사이에 가중치를 적용한 RWC 그리고 색과 질감을 같이 고려한 SCGT와의 비교 실험을 수행하였다. RWC를 사용하는 경우에는 검색 성능이 최고가 되는 색과 질감의 가중치가 영상에 따라서 다르다는 것을 실험으로 확인하였다. 이렇게 실험적으로 결정된 최적 가중치를 사용한 RWC의 결과와 SCGT, CCE의 성능을 비교할 때 CCE가 가장 우수하다는 것을 확인할 수 있었다.

그러나 CCE는 다른 특징벡터에 비해 특징 벡터의 차원과 계산량이 많은 단점을 가지고 있다. 따라서 향후 연구 과제로 CCE의 특징 벡터 차원을 줄이고 계산량을 줄이기 위한 방법이 연구되어야 할 것이다.

참 고 문 현

- [1] Young Rui, Thoas S. Hang and Shih-Fu Chang, "Image Retrieval : Current technique, promising directions, and open issues," Journal of Visual Communication and Image Representation, Vol. 10, pp. 39-62, 1999.
- [2] P. Piamsa-NGA, N. A. Alexandridis, S. Srakaew, G. Blankenship, G. Papakonstantinou, P. Tsanakas and S. Tzafestas, "Multi-Feature Content Based Image Retrieval," Proc. of International Conference on Computer Graphics and Imaging, 1998.
- [3] Jing Huang, Color-Spatial Image Indexing and Applications, Phd Thesis, Cornell University, August 1998.
- [4] M. J. Swain and D. H. Ballard, "Color Indexing," International Journal of Computer Vision, vol. 7, no. 1, pp.11-32, 1991
- [5] Jing Huang, S. Ravi Kumar, Mandar Mitra, Wei-Jing Zhu and Ramin Zabih, "Image Indexing Using Color Correlograms," Proc. of IEEE Computer Vision and Pattern Recognition Conference. San Juan, Puerto Rico, June 1997.
- [6] R. Haralick, "Statistical and Structural Approaches to Texture," Proc. of the IEEE, vol. 67, no. 5, pp.786-804, 1979.
- [7] H. Tamura, S. Mori, and T. Yamawaki, "Texture features corresponding to visual perception," IEEE

- Trans. Sys. Man, and Cybernetics, vol. 8, no. 6, pp.460-473, 1978.
- [8] B. S. Manjunath and W. Y. Ma, "Texture Features for Browsing and Retrieval of Image Data," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 8, Aug, 1996.
- [9] Yossi Rubner, Perceptual Metrics for Image Database Navigation, Phd Thesis, Stanford University, May 1999.
- [10] Christoph Palm and Thomas M. Lehmann, "Classification of color textures by Gabor filtering," Machine Graphics & Vision, vol. 22, no. 2/3, pp.195-219, 2002.
- [11] H.Mller, Wo. Mller, D. McG. Squire, S. M. Maillet and T. Pun, "Performance Evaluation in Content-Based Image Retrieval: Overview and Proposals, Pattern Recognition Letters, 22, 5, pp. 593-601, 2001.

저자 소개

정재웅(정회원)

2001년 2월 한림대학교 정보통신공학부 학사. 2003년 2월 한림대학교 대학원 전자공학과 석사. 2003년 2월~현재 (주)와이더랜드컴 기술연구소에서 VoiceXML IDE 개발. <주요 관심분야: 내용기반 영상검색, VoiceXML.>



권태완(학생회원)

2002년 한림대학교 정보통신공학부 학사. 2002년 2월 ~ 현재 한림대학교 대학원 전자공학과 석사과정 재학. 주요 관심분야: 내용기반 영상검색, embedded linux, XML

박섭형(정회원)



1984년 서울대학교 제어계측공학과 학사. 1986년 서울대학교 대학원 제어 계측공학과 석사. 1990년 서울대학교 대학원 제어계측공학과 박사. 1990년 9월 ~ 1992년 9월 한국생산기술연구원 HDTV개발사업단에서 상공부 G7 과제인 “유럽형 HDTV 수상기 디코더 개발”에 참여. 1992년 10월 ~ 1998년 2월 KT 연구개발본부 통신시스템개발센터와 통신망 연구소에서 “HDTV 전송기술”과 “ATM 망에서 VoD, 영상회의 서비스 시스템, 서비스관리 시스템” 개발. 1993년 10월 ~ 1994년 9월 일본 NTT Human Interface Laboratories에서 “압축동영상의 후처리 알고리듬” 개발. 1998년 3월 ~ 현재 한림대학교 정보통신공학부 교수. <주요 관심분야: 영상 및 음성신호 처리, 내용기반 영상 검색, VoiceXML 응용, 이동통신용 음악 전처리>