

자연영상 검색을 위한 색깔감 특징

*정재웅, 권태완, 박섭형

*와이더덴닷컴 기술연구소, 한림대학교 정보통신공학부

e-mail: *jwjeong@widerthan.com, twkwon@hallym.ac.kr, spark@hallym.ac.kr

A Color Texture Feature For Natural Image Retrieval

*Jae Woong Jeong, Tae Wan Kwon, Seop Hyung Park

*WiderThan.com Co., Ltd. Division of Information and Telecommunications

Engineering, Hallym University

Abstract

In the field of content-based image retrieval, various mathematical low-level features have been proposed to describe the perceptual content of images. Since most of the features are assumed to be independent of each other, one feature is extracted from images without any consideration of the other features. Recently proposed CCE and SCFT taking advantage of the correlation between color and texture have shown relatively good performance. In this paper, the performance of CCE, SCFT, and the traditional regular weighted comparison method are evaluated. Simulation results with natural images have shown that CCE outperforms the other methods.

I. 서론

컴퓨터와 인터넷의 발전으로 멀티미디어 데이터의 양이 급속도로 증가하고 있다. 대량의 멀티미디어 데이터를 효과적으로 활용하기 위한 방안으로 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 동텍스트기반 영상검색의 단점을 극복할 수 있는 방법이다.

내용기반 영상검색에 사용할 수 있는 특징으로는 색, 질감, 모양 등이 있으며, 특히 색과 질감 특징이 가장 널리 사용된다. 대부분의 영상, 특히 일반적인 자연영상의 경우 색과 질감 특징을 모두 가지므로 영상 검색에 색과 질감 특징을 모두 사용하는 것이 바람직하다. 이들 특징은 수치화 되어 데이터베이스에 저장되며 나중에 이 수치화된 특징 벡터들을

사용하여 영상 검색을 수행하게 된다.

색과 질감 특징을 수치화 할 때 색 히스토그램과 가보 필터를 많이 사용한다. 두 가지 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.

그러나 다중 특징을 이용한 기존의 방법들은 각 특징을 개별적으로 추출하여 검색에 사용하기 때문에 영상에서 특징간의 상관관계를 고려하지 못하는 한계가 있다. 예를 들면 특징 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. 그러나 일반적인 영상의 경우 영역을 분할하는 것이 매우 어렵기 때문에 이 방법을 사용하는 것이 쉽지 않은 않다.

이런 이유로 색과 질감 사이의 관계를 고려하는 특징 벡터를 사용해야 한다. 이런 특징 벡터로는 SCGT(sequential chromatic Fourier transform)[3]와 CCE(channel color energy)[4]가 있다. 이 논문에서는 색과 질감 사이의 관계를 고려하는 두 가지 특징 벡터 중에서 CCE를 간략히 설명하고 실험을 통해 SCGT, RWC와의 성능을 비교한다.

이 논문의 구성은 다음과 같다. II장에서는 색과 질감 특징의 상관관계를 이용하는 특징벡터인 CCE에 대해 간략히 살펴보고, III장에서는 실험 및 검토, 마지막으로 IV장에서는 결론을 맺는다.

II. 색 질감 특징

영상 분석이나 영상 분류 분야에서 색과 질감은 대단히 중요한 특징으로 인식되어 왔지만 많은 경우 이 둘이 완전히 독립적인 경우로 간주 되었다. 따라서 질감 특징에 관한 기존의 연구는 컬러 영

상이 아닌 흑백 영상에서 추출하는 것이 일반적이었다. 그러나 최근에 영상 분류에 색깔감(color texture) 특징을 이용하면 흑백 질감 특징만을 사용하는 것보다 우수한 성능을 얻을 수 있다는 연구 결과가 발표되는 등[3], 색과 질감 사이의 상관성을 활용하기 위한 연구가 진행되고 있다. [3]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 색깔감 특징으로 사용하였다.

반면 CCE는 흑백 영상에만 가보 필터링을 적용하여 채널 별 영상을 구한 후에, 원 색영상에서 동일한 색을 가진 화소들을 조사하여 그 화소들과 같은 위치에 있는 채널 영상의 화소들의 에너지를 색깔감 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 24개의 가보 필터와 하나의 저대역 통과 필터를 사용하여 25 채널 신호로 분리하여 $n_1 n_2$ 크기의 가보 필터링 된 공간영역의 영상 F_{so} 와 $n_1 n_2$ 크기의 저대역 통과 필터링 된 영상 F_{lp} 를 얻는다.

I 를 모든 색 채널을 가지는 $n_1 n_2$ 크기의 원본 색 영상이며, $p = (x, y)$ 는 원본 영상에서 좌표 $i(p)$ 는 좌표 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.

$$e_{so} = [e_{so}(c_1), e_{so}(c_2), \dots, e_{so}(c_m)], \quad (1)$$

$$e_{so}(c_i) = \sum_{\{p | i(p) = c_i\}} |f_{so}(p)|, \quad (2)$$

여기에서 i_q 는 화소 i 를 양자화한 값을 의미하며, $f_{so}(p)$ 는 영상 F_{so} 에서 좌표 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 에너지 값을 의미한다.

그리고 저대역 통과 영상의 색 에너지는 다음과 같이 정의된다.

$$e_{lp} = [e_{lp}(c_1), e_{lp}(c_2), \dots, e_{lp}(c_m)], \quad (3)$$

$$e_{lp}(c_i) = \sum_{\{p | i(p) = c_i\}} |f_{lp}(p)|, \quad (4)$$

여기에서 $f_{lp}(p)$ 는 영상 F_{lp} 에서 좌표 p 의 화소 값을 의미하며, $e_{lp}(c_i)$ 는 저대역 영상의 색 c_i 의 에너지를 의미한다.

이제 모든 채널에 대한 CCE는 다음과 같이 정의된다.

$$e = [e_{lp}, e_{00}, e_{01}, \dots, e_{(s-1)(o-1)}]. \quad (5)$$

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, $i_g(p)$ 를 영상 I_g 에

서 좌표 p 의 화소 값, $i'_g(p) = i_g(p) + 1$ 이라 하자. 그리고 I'_g 를 필터링한 결과 영상을 F'_{so} 라 할 때, 특징 필터에 대한 색 에너지는 식 (2)과 같다.

$$e'_{so} = [e'_{so}(c_1), e'_{so}(c_2), \dots, e'_{so}(c_m)], \quad (6)$$

$$e'_{so}(c_i) = \sum_{\{p | i(p) = c_i\}} |\tilde{f}_{so}(p)|, \quad (7)$$

$$\tilde{f}_{so}(p) = \frac{f'_{so}(p)}{i'_g(p)}. \quad (8)$$

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

정규화 된 CCE 특징은 다음과 같다.

$$e = [e'_{lp}, e'_{00}, e'_{01}, \dots, e'_{(s-1)(o-1)}]. \quad (9)$$

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

$$d(I, Q) = |e'_{dc} - e'_{dc}| + \sum_{i=0}^{s-1} \sum_{j=0}^{o-1} |e'_{ij} - e'_{ij}|, \quad (10)$$

$$e'_{ij} - e'_{ij} = \sum_{l=1}^m \frac{|e'_{ij}(c_l) - e'_{ij}(c_l)|}{\sigma(e_{ij}(c_l))}. \quad (11)$$

$$e'_{lp} - e'_{lp} = \sum_{l=1}^m \frac{|e'_{lp}(c_l) - e'_{lp}(c_l)|}{\sigma(e_{lp}(c_l))}. \quad (12)$$

여기에서 Q 는 질의영상, I 는 데이터베이스 내의 영상이다. m 은 양자화 된 색의 개수이며, S 는 전체 척도, O 는 전체 방위 개수이다. $e'_{ij}(c_l)$ 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.

III. 실험 및 성능 평가

SCFT와 CCE의 성능을 평가하기 위해 색과 질감 특징의 정규 가중거리 비교 방법(RWC)과 SCGT, CCE를 비교하였다.

RWC에 사용되는 색 특징과 질감 특징으로는 색 히스토그램 분포와 가보 필터를 이용한 방법을 각각 사용하였다. 색 히스토그램은 R, G, B 채널이 각각 4개의 bin으로 균일하게 양자화된 RGB 색 공간을 사용하였으며, 가보 필터의 척도와 방위는 각각 4와 6을 사용하였다. RWC를 위한 두 특징의 가중치는 반복적인 실험을 통해서 가장 좋은 성능을 얻을 수 있도록 정했다.

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 두렷하게 구분할 수 있는 18장을 선택하였다. 표 1은 실험에 사용한 평가 영상 18장을 정리한 것이다.

평가 척도로는 다음과 같이 정의되는 정확률(precision)과 재현률(recall)을 사용하였다[5].

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

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

표 2, 3, 4는 검색 영상의 수를 각각 10, 30, 50 장으로 했을 때 RWC, SCGT, CCE의 정확률과 재현률을 비교한 것이다. 각 입력 영상에 대해서 상대적으로 성능이 우수한 방법의 결과를 짙은 색으로 표시하였다. 또한 RWC에 사용되는 가중치는 각 입력 영상에 대해서 정확률과 재현률을 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 자연 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 더 중요한 역할을 하는 것을 알 수 있다.

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, 세로축은 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.

IV. 결론

이 논문에서는 색과 질감 특징 사이의 관계를 고려하지 않는 기존의 방법(RWC)과 두 특징의 관계를 고려할 수 있는 두 가지 특징 벡터인 SCFT와 CCE의 성능을 비교 하였다. SCFT는 가보 필터를 색 영상에 적용하기 위하여 푸리에 변환을 확장하여 사용하며, 각 채널 신호의 크기 에너지와 위상 에너지를 색질감 특징으로 사용한다. 반면 CCE는 원본 영상의 흑백 영상을 가보 필터에 의해 다중 채널로 분리한 후 각 채널 별 색의 에너지로 정의된다.

RWC를 사용하는 경우에는 검색 성능이 최고가 되는 색과 질감의 가중치가 영상에 따라서 다르다는 것을 실험으로 확인하였고, 이




렇게 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.

참고 문헌

- [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] 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.
- [4] 정재용, 권태완, 박성형, 질감의존 색 특징을 이용한 내용기반 영상검색, 대한전자공학회 하계종합학술대회 제26권 제1호, July, 2003.
- [5] H.Miller, Wo. Miller, 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.

표 1. 실험에 사용한 영상

		
corel106009	corel107021	corel108036

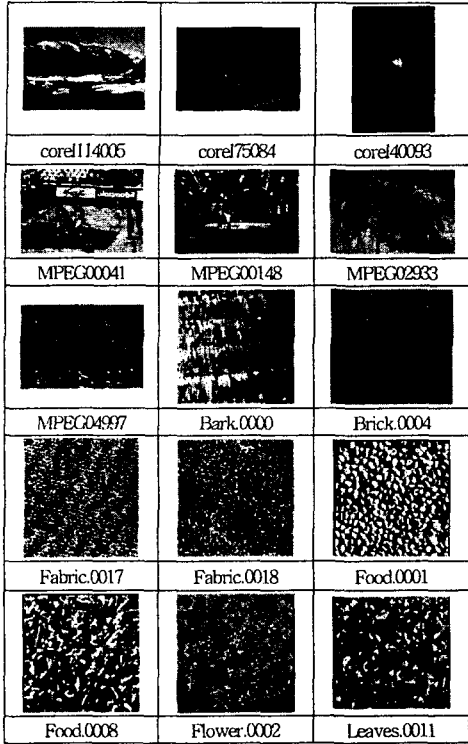


표 2. 10 장의 영상을 검색했을 때 RWC, SCGT, CCE의 정확률과 재현률 비교

질의 영상	정확률			재현률			RWC 가중치
	RWC	SCGT	CCE	RWC	SCGT	CCE	
corel106009	0.2	0.1	0.1	0.05	0.02	0.02	0.9
corel107021	0.3	0.3	0.3	0.07	0.03	0.03	0.7
corel108036	0.3	0.4	0.6	0.04	0.04	0.06	0.9
corel114005	0.8	0.3	0.3	0.08	0.03	0.03	0.9
corel40093	0.6	0.5	0.5	0.06	0.05	0.05	0.7
corel75084	0.3	0.4	0.7	0.04	0.05	0.09	0.9
mpeg00041	0.4	0.4	0.3	0.05	0.27	0.03	0.7
mpeg02933	0.4	0.1	0.5	0.5	0.67	0.03	0.8
mpeg04997	0.3	0.3	0.1	0.33	0.03	0.03	0.7
mpeg00148	0.4	0.4	0.1	0.25	0.03	0.03	0.7
Bark.0000	0.6	0.6	0.6	0.09	0.03	0.03	0.8
Brick.0004	0.9	0.9	0.9	0.14	0.03	0.03	0.7
Fabric.0017	0.9	0.9	0.9	0.03	0.03	0.03	0.8
Fabric.0018	0.9	0.9	0.9	0.03	0.03	0.03	0.8
Food.0001	0.4	0.4	0.4	0.06	0.03	0.03	0.8
Food.0008	0.7	0.1	0.6	0.02	0.09	0.03	0.7
Flower.0002	0.9	0.9	0.9	0.14	0.03	0.03	0.8
Leaves.0011	0.9	0.9	0.9	0.14	0.03	0.03	0.8
최대값 개수	15	4	13	15	4	13	

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

질의 영상	정확률			재현률			RWC 가중치
	RWC	SCGT	CCE	RWC	SCGT	CCE	
corel106009	0.1	0.17	0.1	0.06	0.05	0.05	0.9
corel107021	0.23	0.23	0.4	0.14	0.07	0.12	0.7
corel108036	0.2	0.2	0.4	0.11	0.06	0.12	0.9
corel114005	0.3	0.17	0.39	0.14	0.05	0.03	0.9
corel40093	0.63	0.77	0.4	0.19	0.12	0.07	0.7

corel75084	0.43	0.27	0.67	0.15	0.16	0.23	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	0.33	0.7
mpeg04997	0.27	0.13	0.13	0.89	0.44	0.44	0.7
mpeg00148	0.4	0.2	0.17	0.75	0.38	0.75	0.7
Bark.0000	0.6	0.3	0.3	0.41	0.14	0.47	0.8
Brick.0004	0.3	0.73	0.83	0.34	0.39	0.39	0.7
Fabric.0017	0.9	0.9	0.9	0.47	0.47	0.47	0.8
Fabric.0018	0.9	0.93	0.93	0.44	0.44	0.44	0.8
Food.0001	0.2	0.2	0.2	0.09	0.09	0.09	0.8
Food.0008	0.07	0.27	0.27	0.03	0.13	0.13	0.7
Flower.0002	0.6	0.6	0.6	0.28	0.28	0.28	0.8
Leaves.0011	0.53	0.53	0.53	0.25	0.25	0.25	0.8
최대값 개수	10	3	11	10	3	11	

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

질의 영상	정확률			재현률			RWC 가중치
	RWC	SCGT	CCE	RWC	SCGT	CCE	
corel106009	0.12	0.1	0.2	0.08	0.03	0.03	0.9
corel107021	0.3	0.2	0.32	0.19	0.1	0.16	0.7
corel108036	0.38	0.16	0.32	0.18	0.08	0.16	0.9
corel114005	0.32	0.18	0.18	0.16	0.09	0.39	0.9
corel40093	0.66	0.6	0.36	0.33	0.34	0.18	0.7
corel75084	0.34	0.22	0.3	0.21	0.14	0.32	0.8
mpeg00041	0.28	0.24	0.3	0.3	0.8	0.33	0.7
mpeg02933	0.08	0.04	0.1	0.67	0.33	0.33	0.7
mpeg04997	0.16	0.1	0.1	0.89	0.56	0.56	0.7
mpeg00148	0.24	0.16	0.1	0.75	0.5	0.5	0.7
Bark.0000	0.56	0.24	0.24	0.44	0.19	0.57	0.8
Brick.0004	0.66	0.66	0.7	0.52	0.55	0.55	0.7
Fabric.0017	0.9	0.9	0.9	0.47	0.47	0.47	0.8
Fabric.0018	0.7	0.82	0.82	0.55	0.64	0.64	0.8
Food.0001	0.14	0.14	0.14	0.11	0.11	0.11	0.8
Food.0008	0.1	0.16	0.16	0.08	0.13	0.13	0.7
Flower.0002	0.46	0.46	0.46	0.36	0.36	0.36	0.8
Leaves.0011	0.88	0.4	0.4	0.69	0.31	0.31	0.8
최대값 개수	10	2	12	10	2	12	

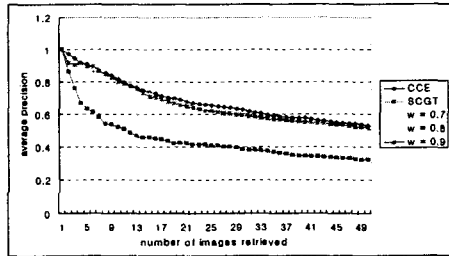


그림 1. RWC, SCGT, CCE의 평균 정확률 비교

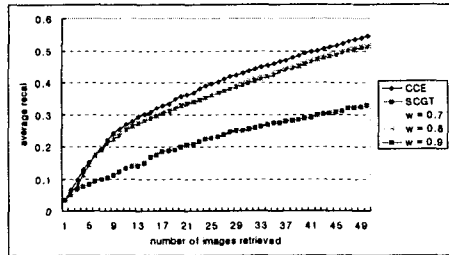


그림 2. RWC, SCGT, CCE의 평균 재현률 비교