

질감의존 색 특징을 이용한 내용기반 영상검색

정재웅*, 권태완, 박섭형
한림대학교 정보통신공학부
*와이더덴닷컴

A Texture-Dependent Color Feature for CBIR

Jae Woong Jeong*, Tae Wan Kwon, Seop Hyeong Park
Division of Information Engineering and Telecommunications, Hallym University
R&D Center, Widerthan.com*

Abstract

내용 기반 영상검색에서 다중 특징을 사용하여 영상을 검색하는 기존의 방법들은 영상에서 특징 간의 상관관계를 고려하지 않고 각 특징을 개별적으로 추출하여 검색에 사용한다. 따라서 특징 간의 최적의 가중치를 찾아야 하는 문제가 있다. 이 논문에서는 내용기반 영상검색을 위해 색과 질감 특징을 효과적으로 표현할 수 있는 새로운 특징 벡터인 CCE (channel color energy)를 제안한다. 실험을 통하여 제안하는 방법이 정규 가중거리 비교 방법에 비해 우수한 성능을 보이는 것을 확인하였다.

I. 서론

내용기반 영상검색이란 영상에 텍스트로 주석을 기입하는 방법이 아닌 영상 자체가 가지고 있는 저수준의 특징, 예를 들면 색 (color), 질감 (texture), 모양 (shape), 움직임 (motion)등을 수치화 하여 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. 내용기반 영상검색에 사용할 수 있는 특징으로는 색, 질감, 모양 등이 있으며, 대부분의 영상이 여러 가지 특징을 가지므로 영상 검색에 다중 특징을 사용하는 것이 바람직하다. 일반적인

영상에서 이런 특징들은 어느 정도 상관관계를 가지고 있다. 다중 특징을 사용하여 영상의 유사도를 측정하는 일반적인 방법으로는 개별적으로 추출된 각 특징들의 중요도에 따라 서로 다른 가중치를 부여하는 정규 가중거리 비교방법[1]이 있다. 그러나 다중 특징을 사용한 기존의 방법들은 영상에서 특징 간의 상관관계를 고려하지 않고 각 특징을 개별적으로 추출하여 검색에 사용하는 한계가 있다.

이에 따라 이 논문에서는 여러 가지 특징 중에서 색과 질감을 동시에 고려할 수 있는 새로운 특징벡터인 CCE 를 제안한다. 제안하는 특징벡터는 질감에 의존하는 색 정보를 표현하는 것으로 색 히스토그램에 기반하는 색 특징과 가보 필터에 기반하는 질감 특징을 사용한다.

이 논문의 구성은 다음과 같다. 장에서는 색과 질감 특징을 이용한 기존의 연구들을 살펴보고, 장에서는 CCE 를 사용한 내용기반 영상검색 방법에 대해 알아보겠다. 장에서는 실험 및 성능평가, 마지막으로 장에서는 결론을 맺는다.

II. 색과 질감특징을 이용한 내용기반 영상검색

2.1 색 특징을 이용한 내용기반 영상검색

색은 사람이 영상을 구할 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 사용할 때 가장 많이 사용하는 방법은 색 히스토그램에 기반 하는 것으로 평균 색 (average color), 색 히스토그램 (color histogram), 색 모멘트 (color moments), 누적 색 히스토그램(cumulative color histogram), 색 코릴로그램 (color correlogram), 국부 색 히스토그램 (local color histogram)[2]등의 방법이 있다.

2.2 질감 특징을 이용한 내용기반 영상검색

질감은 색과 더불어 사람이 영상을 인식할 때 중요하게 판단하는 요소 중 하나이다. 질감은 재질의 표면 차이에서 받는 느낌으로 거침, 부드러움, 규칙적임에 관한 척도라고 할 수 있다. 색이 영상의 픽셀 기반 특징이라 하면 질감은 공간적인 지역의 개념을 포함한다. 질감 특징에 자주 사용되는 방법으로는 이웃 픽셀 (pixel neighborhood)특성에 의한 방법, 히스토그램 기반 (histogram based)방법, 동시 발생 행렬 (co-occurrence matrix)에 의한 방법, 타무라 (Tamura) 질감 표현법[3], 가보 필터 (Gabor filter)[4]를 사용한 방법 등이 있다.

2.3 색과 질감특징을 이용한 내용기반 영상검색

다중 특징을 사용하여 영상 검색을 수행할 경우 일반적으로 사용하는 방법으로 정규 가중거리 비교 (regular weighted comparison)방법[1]이다.

$A = \{a_1, a_2, \dots, a_n\}$ 를 데이터베이스 내 영상의 특징, $Q = \{q_1, q_2, \dots, q_n\}$ 를 질의 영상의 특징이라 하자.

그리고 $d_i(a_i, q_i)$ 는 영상 A 와 영상 Q 의 i 번째 특징 간의 거리라고 정의하자. 이와 같이 정의할 때 두 영상 A 와 Q 의 정규 가중거리 비교 방법에 의한 거리는

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

와 같다. 여기에서 $\sum_{i=1}^n w_i = 1$ 이다.

다중 특징을 영상검색에 사용하려면 각각의 특징 벡터에 대한 적당한 가중치가 필요하게 되며 영상의 특징에 맞는 가중치의 선택은 매우 중요한 문제이다. 그러나 대부분의 경우 영상마다 뚜렷하게 구분이 되는 특징이 다르기 때문에 모든 영상에 적용할 수 있는 적당한 가중치를 선택하는 것은 매우 어려운 문제이다. 이에 따라 신경망 (neural net), 적합성 피드백 (relevance

feedback)등을 이용하여 문제를 해결하는 방법이 어려 분야에서 연구되고 있다. 그러나 이러한 방법들의 대부분은 개별적인 저 수준 특징 벡터 각각에 의존하여 가중치를 조절하게 되므로 특징간의 상관관계를 표현하기 어렵다는 단점이 있다.

III. 채널 색 에너지 (channel color energy)

CCE 는 색과 질감의 상관관계를 나타낼 수 있는 특징으로 가보 필터에 기반 하는 질감 특징과 색 히스토그램에 기반 하는 색 특징을 표현한다. CCE 의 추출 및 유사도 계산 과정은 다음과 같다.

3.1 필터 설계

가보 필터에 기반 하는 질감특징을 얻기 위해 2 차원 가보 필터 $g(x, y)$ 와 저대역 통과필터 $g_{lp}(x, y)$ 를 다음과 같이 설계한다.

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

$$g_{lp}(x, y) = \left(\frac{w}{2\pi\sigma_x\sigma_y} \right) \exp \left(-\frac{1}{2} \left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right] \right), \tag{2}$$

여기에서 $w = \frac{\sigma_{ud}}{\sigma_{ud}}$, $\sigma_{ud} = -\sigma_{ud} + \frac{U_l}{\sqrt{2 \ln 2}}$ 이며 $\sigma_{vd} = \sigma_{ud}$ 이다.

3.2 필터링

제안하는 특징은 색과 질감을 모두 사용한다. 따라서 원본 영상은 단색 영상이 아니고 색상을 가지는 영상이다. 이 영상으로부터 질감 특징을 추출하기 위해 원본 영상의 색 공간을 HSV (Hue Saturation Value) 색 공간으로 변환 후 V 채널에 대하여 필터링을 수행한다.

3.3 색 공간 선택 및 색 공간 양자화

이 논문에서 제안하는 색 특징은 색 히스토그램에 기반 한다. 따라서 색 공간을 그대로 적용할 경우 특징 벡터의 차원이 매우 커지게 된다. 따라서 특징 벡터의 차원을 줄이기 위해 색 공간을 양자화 한다. 실험에서는 각 채널이 2bits 로 양자화 된 RGB 색 공간을 사용하

였다.

3.4 원본 영상의 특정 색이 특정 채널에서 가지는 에너지 추출

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

$$\mathbf{e}_{so} = [e_{so}(c_1), e_{so}(c_2), \dots, e_{so}(c_m)], \quad (3)$$

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

여기에서 F_{so} 는 $n_1 \times n_2$ 크기의 가보 필터링 된 공간영역의 영상이며, $f_{so}(\mathbf{p})$ 는 영상 F_{so} 에서 좌표 \mathbf{p} 의 픽셀 값, $e_{so}(c_i)$ 는 스케일 s , 방위 o 에 해당하는 필터에 의해 필터링 된 영상의 색 c_i 에 대한 에너지를 의미한다. 그리고 저대역 통과 필터에 의한 색 에너지 \mathbf{e}_{so} 를 식 (3), (4)와 유사한 방법으로 계산한다.

이제 모든 채널에 대한 색 에너지 특징 벡터(CCE)는 다음과 같이 정의된다.

$$\mathbf{e} = [\mathbf{e}_{lp}, \mathbf{e}_{00}, \mathbf{e}_{01}, \dots, \mathbf{e}_{0(O-1)}, \mathbf{e}_{11}, \dots, \mathbf{e}_{(S-1)(O-1)}]. \quad (5)$$

3.5 정규화

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

3.6 유사도 계산

두 영상의 CCE 특징 벡터를 위한 유사도 (Similarity Measure)를 측정하기 위해 가우시안 정규화 (Gaussian Normalization) 된 L_1 놈을 사용하였다.

$$d(I, Q) = |\mathbf{e}'_p - \mathbf{e}^o_p| + \sum_{i=0}^{S-1} \sum_{j=0}^{O-1} |e'_i - e^o_i|, \quad (6)$$

$$\mathbf{e}'_i - \mathbf{e}^o_i = \sum_{l=1}^m \frac{|e'_i(c_l) - e^o_i(c_l)|}{\sigma(e_{ij}(c_l))}, \quad (7)$$

$$\mathbf{e}'_p - \mathbf{e}^o_p = \sum_{l=1}^m \frac{|e'_p(c_l) - e^o_p(c_l)|}{\sigma(e_{lp}(c_l))}, \quad (8)$$

여기에서 Q 는 질의영상, I 는 데이터베이스 내의 영상이다. m 은 양자화 된 색의 개수이며, s 는 전체 스케일, o 는 전체 방위 개수이다. $e^o_{ij}(c_i)$ 와 $e'_i(c_i)$ 는 영상 Q 와 I 에 대하여 스케일 i , 방위 j 인 채널에서 색 c_i 의 에너지를 나타낸다. 그리고 $\sigma(e_{ij}(c_i))$ 와 $\sigma(e_{lp}(c_l))$ 는 전체 데이터베이스에서 $e_{ij}(c_i)$ 의 $e_{ij}(c_i)$ 표준편차 이다.

IV. 실험 및 성능평가

제안하는 알고리즘의 성능을 평가하기 위해 색과 질감 특징의 정규 가중거리 비교 방법 (RWC)을 제안한 알고리즘과 비교하였다. RWC 를 위한 색 특징으로는 색 히스토그램을 사용하였으며, 질감 특징으로는 가보 필터를 이용한 방법이 사용되었다. 색 히스토그램은 R, G, B 채널이 각각 4 개의 bin으로 균일하게 양자화 된 RGB 색 공간을 사용하였으며, 스케일 4, 방위 6 개인 가보 필터를 사용하였다. RWC 를 위한 두 특징의 가중치는 반복적인 실험에 의해 선택하였다. 실험 환경은 AMD Athlon 1.4GHz, 1G RAM, Windows 2000 에서 Visual C++ 6.0 에 의해 제작 되었다. 실험 영상으로는 15,000 여장의 코렐(Corel) 영상과 5,400 여장의 MPEG 실험 영상, 1,800 여장의 직물 디자인 영상을 사용하였다. 질의 영상으로는 해당 질의 영상에 적합한 결과 영상을 뚜렷하게 구분할 수 있는 10 장을 선택하였다. 평가 척도로는 정확률 (precision)과 재현률 (recall)을 사용하였다[5]

표 1 은 10 개 질의 영상에 대한 실험 결과를 정리한 것으로, 각 질의 영상에 대하여 10 개, 30 개, 50 개의 영상을 검색한 경우 RWC 와 CCE 의 성능을 정리한 것이다. 표 2 에서 음영이 있는 부분은 상대적으로 성능이 우수한 경우를 나타내며, 정확률과 재현률은 순서대로 10 개, 30 개, 50 개에 대한 것을 의미한다. 표 2 에 의해 영상들마다 질감과 색에 대한 가중치가 다르게 나오는 것을 확인 할 수 있다. 또한 검색 개수에 따라 시스템의 성능이 어느 정도 영향을 받는 것을 알 수 있다.

