

Pixel layer 들 간의 색상 공간 분포에 따른 공간적 분포를 이용한 영상 검색

안재현, 하성중, 이상화, 조남익
서울대학교 전기컴퓨터 공학부

jhahn@ispl.snu.ac.kr, oanchovy@ispl.snu.ac.kr, lsh529@snu.ac.kr, nicho@snu.ac.kr

Image Retrieval Using Color & Spatial Distribution between Pixel Layers

Jaehyun An, Seong Jong Ha, Sang Hwa Lee, Nam Ik Cho

School of Electronic Engineering and Computer Science, Seoul National University

요 약

본 논문에서는 컬러 영상의 검색을 위하여 영상을 색상 정보에 기반한 pixel layer (cluster)의 집합체로 모델링하고, 두 layer 간의 유사도를 각 layer 를 이루는 pixel 들의 색상 분포에 따른 공간적 분포를 이용하여 측정하는 기법을 제안한다. 먼저 pixel layering 단계에서는 HSV 색 공간에서 mean-shift clustering 알고리즘을 통해 초기 layer 들을 얻고, 비슷한 색상의 layer 들은 합쳐 영상의 soft segmentation 과 유사한 결과를 얻는다. 비교할 두 영상에서 pixel layering 을 한 후, 각 layer 를 이진화된 공간분포 지도로 형성하고 그 차이를 비교함으로써 유사도를 측정한다. 이 때, 사용하는 가중치로서 HSV 색 공간 분포의 비슷한 정도를 정의하는데, 이는 HSV 색 공간을 XYZ 의 3 차원 좌표로 설정하고, overlap 되는 pixel 수로 정의하였다. 본 논문에서 제안한 pixel layer 들 간의 색상 공간 분포에 따른 공간적 분포를 이용한 영상 검색 기법은 MPEG-7 에서 정의한 대표색상 기반의 영상 검색보다 우수한 성능을 보여주었다.

1. 서론

최근 개인용 디지털카메라와 스마트 단말기 보급의 증가로 사용자가 많은 양의 영상들을 저장하고 관리하게 되면서, 자동적인 영상 검색/분류 시스템에 대한 관심이 크게 증가하고 있다. 주요 단어의 분포상태를 중심으로 문서를 검색하는 BOW (Bag-of-Words) 방식은 자료검색에서 높은 정확도를 보여주고 있는 검증된 기술이지만, 영상을 검색하는 기술은 문서검색만큼의 신뢰도를 보여주지 못하고 있는 추세이다. 그러므로 최근에는 BOW 개념을 사용하여 특징추출 기술을 기반으로 영상에서 시각단어(visual word)를 추출하여 영상검색을 시도하는 연구가 활발히 진행되고 있다. 하지만 문서에서의 단어와 달리 시각단어는 영상의 다양성으로 인하여 명확히 규정되기 어렵고, 분류한 시각단어의 종류와 개수에 따라 검색되는 영상이 좌우되는 경향이 있기 때문에 좋은 결과를 보여주지 못하고 있다. 이렇듯 사용자가 특정한 대상을 검색하거나 비슷한 분위기의 영상을 검색하는 경우에 특징점을 이용한 기술보다는 색상을 이용한 기술이 더 효과적인 경우가 많다. 그러므로 본 논문에서는 색상정보를 기반으로 유사한 객체나 배경을 중심으로 영상을 검색하는 방법을 개발하고자 하였다.

기존의 색상 기반의 검색 기술을 살펴보면 다음과 같다. 우선, 다양한 색 공간의 히스토그램을 이용하는 방식들[1-3]은 통계적으로 유사한 색상 분포를 갖는 영상들을 검색하는데 효과적이지만, 히스토그램이라는 공간적 속성을 제거한 정보를 이용하므로 영상 속의 객체나 배경과 같은 내용 기반의 영상 검색에서는 효과적이지 못하다. 공간적 정보를 이용하기 위해서 [4]에서는 영상 내에서 같은 색상의 응집 여부를 이용하였으나,

공간 정보가 극히 적다는 단점이 존재한다. [5]에서는 색 정보를 반영하기 위해서 색상 히스토그램의 모멘트를 이용하고 공간 정보 반영을 위해서 웨이블릿 부이미지를 가우시안 분포로 근사한 뒤 계수를 사용하였다. 이 방법은 색 정보와 공간 정보를 분리해 다루었으므로 영상 간 유사도 정의 시 실험에 의해 각 특징량의 가중치를 결정해야 하는 단점이 있다. 또한 대표 색상을 이용해 유사 영상을 검색하는 MPEG-7 의 대표 색상 기술자[6]는 영상의 색상정보의 통계적 특성을 이용하는 방식으로 영상검색 기술을 개발하였으나, 두 영상이 같은 색상 분포를 가지고 있을 때 공간적 정보를 포함하고 있지 않으므로 매칭이 잘 안된다는 단점이 있다.

이와 같이 기존 기술의 문제점을 개선하기 위하여 본 논문에서는 영상의 색상 정보를 이용하여 영상을 pixel layer 들의 집합체로 모델링하고, 각 layer 의 색상 분포에 따른 공간적 분포를 고려함으로써 쿼리 영상과 유사한 영상을 검색하는 기법을 제안한다. 본 논문의 구성은 다음과 같다. 2 절에서는 제안하는 알고리즘으로 색상 정보를 이용한 pixel layering 방법과 색상 분포에 따른 공간적 분포를 이용하여 유사도를 측정하는 방법을 설명한다. 3 절에서는 제안한 방법과 MPEG-7 standard 방법을 이용한 비교 실험 결과를 제시하며, 4 절에서는 본 논문에 대한 결론을 정리하도록 하겠다.

2. 제안하는 알고리즘

본 논문에서는 컬러 영상의 검색을 위하여 영상을 색상 정보를 이용하여 pixel layer 들의 집합체로 모델링하는 pixel layering 방법과 비교할 두 영상의 유사도를 측정하기 위하여 layer 들 간의 색상 분포에 따른 공간적 분포를 이용하는

방법을 제안한다.

2.1 색상 정보를 이용한 pixel layering

컬러 영상의 검색을 위하여 먼저 영상을 색상 정보를 이용해 pixel layer 들의 집합체로 모델링한다. 즉, 영상 I 는 다음과 같이 여러 개의 layer 들의 집합체로 표현된다.

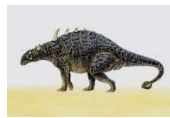
$$I = \bigcup_{i=1}^N L_i$$

여기서 L_i 는 N 개의 추출된 layer 들 중에 i 번째 layer 를 의미한다. 초기 layering 은 HSV 색 공간에서 mean-shift clustering 알고리즘을 통해 얻는다. 이 결과는 색상 공간에서 픽셀들을 대략적으로 clustering 한 것이므로, 이를 좀 더 refinement 하는 과정이 필요하다. 이 과정에서는 각 cluster 의 중심 값의 차이가 적으면 layer 를 결합하여 영상의 soft segmentation 과 유사한 결과를 얻도록 한다. 실제 색상은 어두워질수록 채도 값에 따른 색상 변화가 크지 않기 때문에, 영상에서 색상 정보는 채도 값과 명도 값이 매우 작지 않을 때에만 유효하므로 refinement 하는 과정은 다음 식과 같다.

$$L_i^{new} = L_j \cup L_k \text{ if } \|c_j - c_k\|_2 \leq \epsilon_c$$

where

$$c_j = \begin{cases} (h_j, s_j, v_j) & \text{if } s_j > T_s, v_j > T_v \\ v_j & \text{otherwise} \end{cases}$$



(a)



(b)



(c)

그림 1. Pixel layering 단계의 예. (a) 쿼리 영상, (b) 초기 layering 결과, (c) refinement 과정 후 얻어진 layering 결과

여기서 L_i^{new} 는 L_j 와 L_k 를 결합한 layer 이고, c_j 와 c_k 는 각 layer 에 해당하는 cluster 의 중심 값으로 채도 값 (s_j)과 명도 값 (v_j)이 매우 작지 않을 때에는 색상 (h_j), 채도, 명도로 3 차원 벡터를 나타내고, 그렇지 않으면 명도 값만 나타내는 값이다. 그림 1 은 이 과정을 보여준다.

그림 1.(a) 는 쿼리 영상이고, 그림 1.(b) 와 (c)는 색상 정보를 이용한 초기 pixel layering 의 결과와 refinement 과정 후 얻어진 layering 의 결과를 보여준다.

2.2 색상 분포에 따른 공간적 분포를 이용한 유사도 측정

2.2.1 색상 분포에 따른 가중치 설정

쿼리 영상과 DB 영상의 유사도를 측정하기 위하여, 두 영상을 pixel layering 한 후에 유사한 색상을 가지는 layer 간의 색상 분포가 얼마나 비슷한 지에 대한 가중치를 정의한다. 그림 2 는 HSV 색 공간 원기둥 모형과 이를 XYZ 의 3 차원 좌표로 설정한 것을 보여준다.

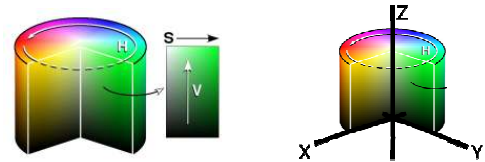


그림 2. HSV 색 공간 원기둥 모형과 XYZ 3 차원 좌표.

그림 2 에서 보이는 바와 같이, HSV 색 공간을 XYZ 의 3 차원 좌표로 표현할 수 있다. 그러므로 layer 에 해당되는 pixel 들의 색상 분포를 3 차원으로 표현할 수 있고 이 분포를 X-Y, Y-Z, 그리고 Z-X 의 2 차원으로 투영시킨다. 비교할 두 layer 의 색상 분포가 유사하면 overlap 되는 pixel 수가 많게 되므로, 전체 영상의 pixel 수 대비 overlap 되는 pixel 수를 가중치로 설정한다. 즉, 쿼리 영상의 k 번째 layer 와 DB 영상의 i 번째 layer 간의 가중치 w_{ki} 는 X-Y, Y-Z, Z-X 평면으로 투영시켰을 때의 가중치들의 평균값으로 나타낼 수 있다.

2.2.2 공간적 분포를 이용한 두 영상의 유사도 측정

컬러 영상에서는 그 색상이 어떻게 분포하느냐에 따라서 검색되는 영상의 종류와 객체의 형태가 크게 달라진다. 그러므로 본 논문에서는 각 layer 영상을 일정한 개수의 작은 영역으로 나누고 색상이 영상평면에서 분포한 상태를 이진화된 공간분포 지도로 표시한다. 즉, 쿼리 영상의 k 번째 layer 영상 L_k^Q 의 공간적 분포를 나타내는 이진 공간분포 지도 M_k^Q 는 영상을 $n \times n$ 블록으로 나누고, 각 블록에서 layer 에 속하는 pixel 수가 얼마나 되느냐에 의해 결정된다. 블록 p 에서의 이진 공간분포 지도를 식으로 나타내면 다음과 같다.

$$M_k^Q(p) = \begin{cases} 1, & \text{if } \frac{N(A)}{N(B)} > \gamma \\ 0, & \text{otherwise} \end{cases}$$

여기서 $N(A)$ 는 각 블록에 속하는 모든 pixel 의 개수이고, $N(B)$ 는 각 블록에서 각 layer 에 속하는 pixel 의 개수이다. 마찬가지로 DB 영상의 i 번째 layer 영상에 대한

이진 공간분포 지도 M_i^{DB} 를 생성하면, 두 layer 간의 유사도 $dist(L_k^Q, L_i^{DB})$ 는 다음과 같이 계산된다.

$$dist(L_k^Q, L_i^{DB}) = \sum_{p=1}^{n \times n} H_{ki}(p)$$

where

$$H_{ki}(p) = \begin{cases} 1, & \text{if } M_k^Q = 1 \text{ and } M_i^{DB} = 1 \\ 0, & \text{otherwise} \end{cases}$$

결과적으로 색상 분포에 따른 공간적 분포를 이용하여 쿼리 영상과 DB 영상의 유사도 $S_{Q,DB}$ 는 다음과 같이 구할 수 있다.

$$S_{Q,DB} = \sum_{k=1}^n \sum_{i=1}^m \omega_{ki} \cdot dist(L_k^Q, L_i^{DB})$$

여기서 n 은 쿼리 영상의 총 layer 개수이고, m 은 쿼리 영상의 k 번째 layer 와 유사한 색상을 가진 DB 영상의 layer 개수이다. 그림 3 은 쿼리 영상의 k 번째 layer 와 DB 영상의 i 번째 layer 간의 유사도 측정의 예를 보여준다.

3. 실험 결과

본 논문에서 제안한 방법을 검증하기 위해 Core1K Database 영상을 이용하였다. 이는 10 개의 class 들로 이루어진 1000 장의 영상으로, 그림 4,5 는 MPEG-7 대표 기술자를 사용한 결과와 제안하는 방법을 사용한 검색결과로써, 쿼리 영상과 가장 유사하다고 생각되는 상위 20 장의 영상을 뽑은 것을 보여준다. 그림에서 보는 바와 같이, 제안하는 영상 검색 방법은 영상에서의 색상 분포를 가중치로 하여 공간적 분포와 결합하여 고려하기 때문에 기존의 방법보다 우수한 검색 결과를 보여준다.

4. 결론

본 논문에서는 컬러 영상의 검색을 위하여 영상의 색상 정보를 이용하여 영상을 pixel layer 들의 집합체로 모델링하고, 각 layer 들 간의 색상 분포에 따른 공간적 분포를 이용하였다.

초기 pixel layering 은 mean-shift clustering 알고리즘을 이용하였고, refinement 과정을 거쳐 soft segmentation 와 유사하게 layer 들을 생성하였다. 그런 다음 layer 들 간의 색상 분포와 공간적 분포를 이용하여 유사도를 측정하였다. 각 layer 에서 색상이 공간적으로 분포한 상태를 이진화된 공간분포 지도로 모델링 하였고, 이를 색상 분포에 따라 가중치를 두어 비교함으로써 영상을 검색하였다. 본 논문에서 제안한 영상검색 기법은 기존의 MPEG-7 standard 영상 검색 기법보다 더욱 정확한 결과를 보여주었다.

감사의 글

이 논문은 2012 년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. 2012-0000913).

5. 참고문헌

[1] M. Swain and D. Ballard, " Color indexing," IJCV, vol. 7, no. 1, pp. 11-32, 1991.

[2] M. Stricker and A. Dimai, " Color indexing with weak spatial constraints," In Storage and Retrieval for Image and Video Databases (SPIE), pp. 26-40, 1996.

[3] R. Brnuelli and O. Mich, " Histograms analysis for image retrieval." PR, vol. 34, pp. 1625-1637, 2001.

[4] G. Pass, R. Zabih, and J. Miller, " Comparing images using color coherent vectors," In Proc. of the ACM Multimedia Conference, pp. 65-73, 1996.

[5] M. K. Mandal, T. Aboulnasr, and S. Panchanathan, " Image indexing using moments and wavelets," IEEE Transactions on Consumer Electronics, vol. 42, no. 3, pp. 557-565, Aug. 1996.

[6] ISO/IEC 15938-3/FDIS Information Technology - Multimedia Content Description Interface - Part 3 Visual, Jul. 2001, ISO/IEC/JTC1/SC29/WG11 Doc. N4358.

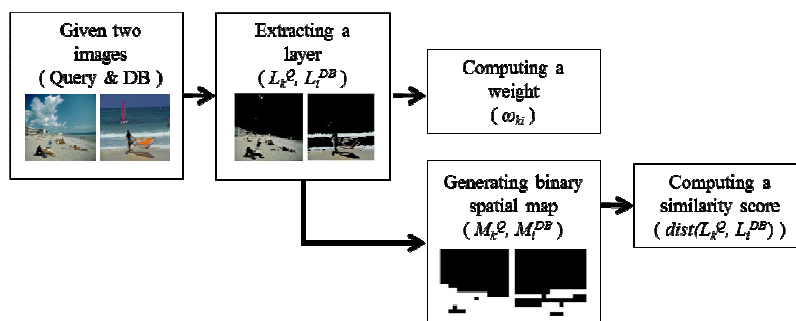
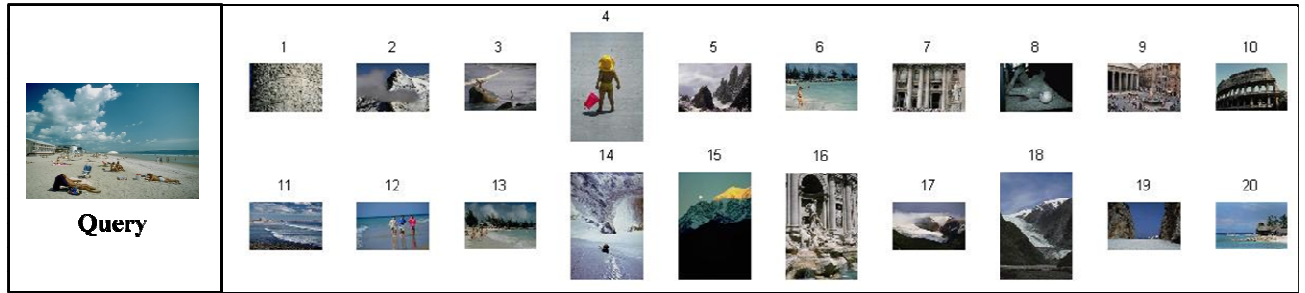
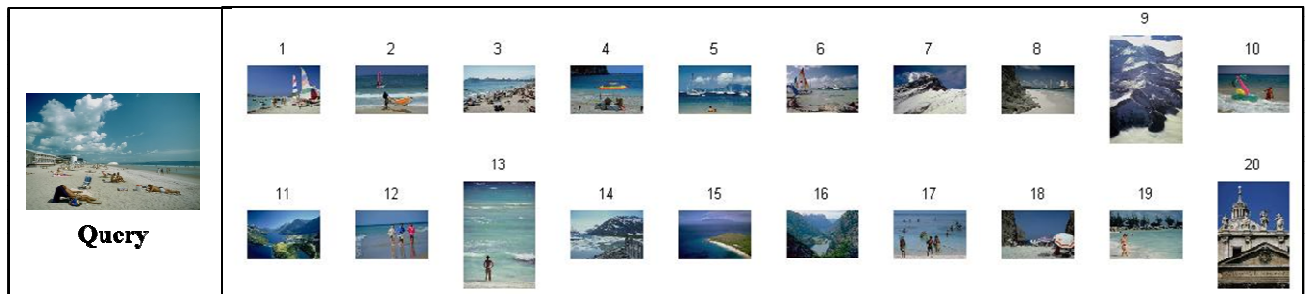


그림 3. 쿼리 영상의 k 번째 layer 와 DB 영상의 i 번째 layer 간의 유사도 측정의 예.

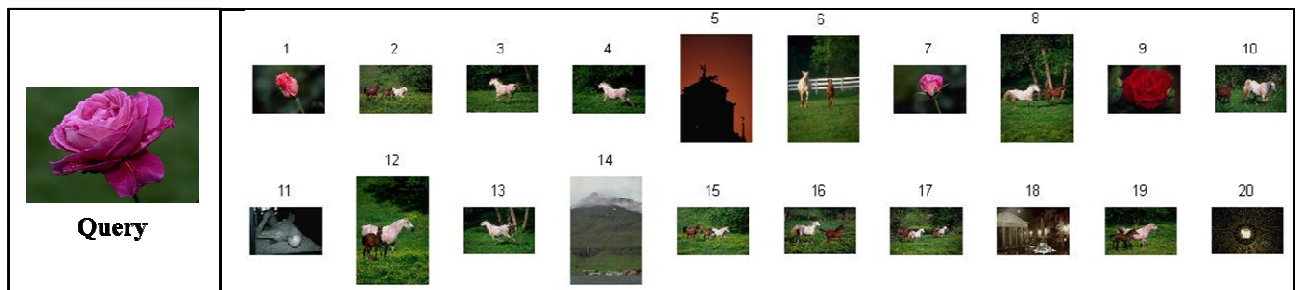


(a)

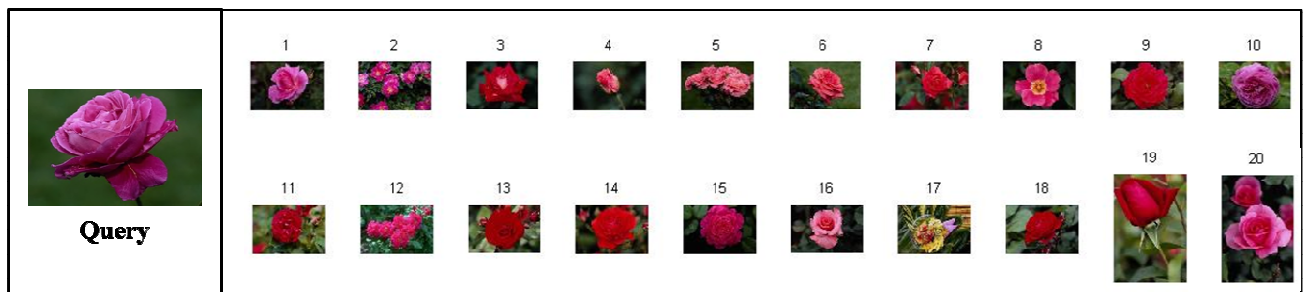


(b)

그림 4. 실험 결과. (a) 대표 색상 기술자를 사용한 결과, (b) 제안하는 방법을 사용한 결과.



(a)



(b)

그림 5. 실험 결과. (a) 대표 색상 기술자를 사용한 결과, (b) 제안하는 방법을 사용한 결과.