

Saliency Map을 이용한 대표 색상 기반의 영상 검색

안재현 이상화 조남익

서울대학교 전기·정보공학부

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

Dominant Color Based Image Retrieval using Saliency Map

An, Jae-Hyun Lee, Sang-Hwa Cho, Nam-Ik

Department of Electrical and Computer Engineering, Seoul National University

요약

본 논문에서는 객체 위주의 컬러 영상 검색을 위하여 영상의 saliency map을 이용해 객체 중심의 영상을 생성하고, 객체와 그 주변 영역에서의 대표 색상이 가지는 통계적 특성과 공간적 분포 정보를 이용하는 방법을 제안한다. 먼저, 영상의 saliency map을 이진화하여 영상을 객체/배경으로 분할하고 객체를 중심으로 객체/배경의 비율이 일정한 일정 크기의 영상을 생성한다. 생성된 영상에서 대표 색상을 추출하고, 각 색상이 영상에서 어떻게 분포하는가를 나타내는 이진 공간분포 지도를 형성한다. 그 후 영상 간의 대표 색상마다 이진 공간분포의 차이를 비교함으로써, 색상의 통계적 특성과 공간적 분포가 동시에 반영된 특징으로 영상을 검색한다. 본 논문에서 제안한 saliency map을 이용한 대표 색상 기반의 영상 검색 기법은 기존의 대표 색상 기반의 영상 검색보다 우수한 성능을 보여준다.

1. 서론

최근 카메라나 휴대용 단말기의 보급으로 많은 양의 영상들을 저장하고 관리하는 일이 보편화되면서, 자동적인 영상 검색/분류 시스템의 필요성이 증가하고 있다. 영상 검색은 태그와 같은 텍스트를 이용하는 방법과 영상 자체의 내용을 기반으로 하는 방법이 있다. 내용 기반의 방법으로는 영상에서 특징점 기반으로 시각단어(visual word)를 추출하여 영상 검색을 시도하는 연구가 많지만, 시각단어는 영상의 다양성으로 인하여 명확히 규정되기 어렵고 분류한 시각단어의 종류와 개수에 따라 검색되는 영상의 결과가 좌우되는 경향이 있다. 이렇듯 사용자가 특정한 대상을 검색하거나 비슷한 분위기의 영상을 검색하는 경우에 특징점을 이용한 기술보다는 색상을 이용한 기술이 효과적인 경우가 많다. 특히 영상에서 가장 두드러지는 색상을 나타내는 대표 색상을 기반으로 하여 영상을 검색하는 방법에 대한 연구는 최근까지도 활발히 진행되고 있다.

MPEG-7의 대표 색상 기술자[1,2] 방식은 영상의 색상 정보의 통계적 특성을 이용하는 방식으로 영상 검색 기술을 개발하였으나, 두 영상 간의 유사도를 계산하는 방법이 사람이 인식하는 유사성과 차이가 있다는 한계점이 있다. [3]에서는 새로운 유사도 계산법을 제안하였지만 영상의 객체가 아닌 배경에서의 대표색상이 유사할 경우에도 유사한 영상으로 매칭되는 단점이 있다. 이를 해결하기 위하여 [4]에서는 배경보다는 객체를 중심으로 검색하기 위하여 각 대표 색상에 가중치를 두어 유사도를 계산하는 방법을 제안하였다. 마찬가지로 본 논문에서는 객체 위주의 영상 검색에 초점을 맞추어, 영상의 saliency map을 이용하여 영상의 객체 영역을 찾고 객체와 그 주변 영역에서의 대표 색상 정보를 기반으로 영상을 검색하는 방법을 제안한다.

2. 제안하는 알고리즘

본 논문에서는 쿼리 영상과 유사한 영상을 검색하기 위하여 두 영상을 saliency map을 이용하여 객체를 중심으로 하는 일정 크기의 영상으로 생성하는 방법, 생성된 쿼리 영상의 대표 색상을 클러스터링(clustering)하여 공간분포 지도를 생성하는 방법, 그리고 생성된 공간분포 지도를 이용하여 두 영상의 유사도를 측정하는 방법을 제안한다.

2.1 객체 중심의 영상 생성

먼저, 객체 중심으로 영상을 검색하기 위해서 영상에서 객체의 위치를 찾고, 객체와 그 주변 영역으로만 이루어진 영상을 생성한다. 영상을 객체/배경으로 분할하는 알고리즘은 여러 가지가 존재하지만, 본 논문에서는 객체/배경에 대한 seed가 필요 없고 계산량도 간단한 saliency map[5]을 이용한다. 영상의 saliency map을 이진화함으로써 객체/배경 영역을 얻고 객체와 배경의 비율이 일정하도록 객체를 중심으로 영상을 자른다. 이 영상은 일정한 크기를 갖고 그 안의 객체의 크기가 일정하므로 본 논문의 2.3에서 제안하는 이진 공간분포 지도를 이용하는 유사도 계산 방법이 타당성을 가진다는 것을 보여준다. 그림 1(a)와 (b)는 쿼리 영상과 saliency map을 보여주고, 그림 1(c)는 제안하는 방법을 통하여 얻은 객체 중심의 영상을 보여준다.

2.2 대표 색상의 추출과 그에 따른 이진 공간 분포 지도의 생성

대표 색상을 구하기 위한 가장 보편적인 방법은 색상 히스토그램에서 가장 많은 빈도수를 가진 색상을 대표 색상으로 정의하는 것이다

[1]. 일반적으로 사용되는 RGB 색 공간은 색상들 간의 의존성이 높고 이들 간의 근접성이 색상의 유사성을 나타내지 못한다는 단점을 가지므로, 본 논문에서는 HSV 색 공간을 이용하여 쿼리 영상 Q 의 HSV 색상 히스토그램[6]에서 대표 색상 C_i 을 추출한다. 하지만 추출된 대표 색상은 히스토그램의 bin(bin)의 개수를 어떻게 양자화 하느냐에 따라 달라질 수가 있으며, 추출된 색상의 bin이 인접한 경우 대표 색상이 유사해지는 결과를 낳을 수가 있다. 그러므로 추출된 대표 색상의 개선(refinement)을 위하여 히스토그램에서 빈도수가 0인 영점bin(zero-bin)을 검출하고, 두 영점bin 사이의 대표 색상들이 하나의 클러스터를 생성하도록 하여 최종적으로 4개에서 8개 사이의 대표 색상 클러스터 D_k 를 추출한다.

또한, 컬러 영상에서는 그 색상이 어떻게 분포하느냐에 따라서 검색되는 영상의 종류와 객체의 형태가 크게 달라진다. 그러므로 본 논문에서는 대표 색상의 공간적 분포를 나타내는 이진 공간분포 지도를 생성한다. 즉, 쿼리 영상을 $n \times n$ 블록으로 나누고, 각 블록 bl 에서 추출된 대표 색상 D_k 에 속하는 픽셀 수가 어느 문턱치 값 이상이면 1, 그렇지 않으면 0인 이진 공간분포 지도 M_k^Q 를 생성한다. 마찬가지로 주어진 데이터베이스 영상에 대하여 쿼리 영상의 대표 색상에 대한 공간적 분포를 나타내는 공간분포 지도 M_k^{DB} 를 생성한다. 그림 1은 공간분포 지도를 생성함에 있어서 본 논문에서의 대표 색상의 클러스터링 방법이 효과적이라는 것을 보여준다.

2.3 유사도 계산

쿼리 영상 Q 와 데이터베이스 영상 DB 간의 유사도는 대표 색상이 유사할수록, 그리고 그 대표 색상의 공간적 분포가 유사할수록 높은 값을 가질 것이라고 여겨진다. 또한, 대표 색상은 객체 중심으로 생성된 영상에서 추출되므로 D_k 는 객체 또는 객체 주변의 배경에 속하게 된다. 그러므로 객체와 객체 주변의 배경에 속하는 대표 색상의 부분 집합 D_F 와 D_B 을 구하여, 다음 식 (1)과 같이 DB 에서 객체와 배경에 대한 대표 색상이 존재하느냐에 따라서 두 영상이 비교할 대상이 되는지 판단한다.

$$\left(\frac{\sum_{k \in F} d(M_k^Q, M_k^{DB})}{N_F} \geq \tau_F \right) \cap \left(\frac{\sum_{k \in B} d(M_k^Q, M_k^{DB})}{N_B} \geq \tau_B \right) \quad (1)$$

where

$$d(M_k^Q, M_k^{DB}) = \sum_{bl} H(M_k^Q(bl), M_k^{DB}(bl))$$

여기서 $H(M_k^Q(bl), M_k^{DB}(bl))$ 는 쿼리 영상과 데이터베이스 영상의 동일한 위치의 블록 bl 에서의 이진 AND 연산이며, N_F 와 N_B 는 쿼리 영상의 객체/배경에 속하는 픽셀의 수이다. 두 영상이 비교할 대상으로 판단되면, 쿼리 영상과 데이터베이스 영상 간의 유사도 $Score$ 는 다음 식 (2)와 같이 계산된다.

$$Score(Q, DB) = \sum_k \omega_k \times d(M_k^Q, M_k^{DB}) \quad (2)$$

$$\omega_k = \exp\left(-\frac{(m_k^Q - m_k^{DB})^2}{\sigma^2}\right)$$

여기서 m_k^Q (또는 m_k^{DB})는 3차원 벡터로 쿼리 영상 Q (또는 DB)에서 k 번째 대표 색상 클러스터 D_k 에 대응하는 HSV 색 공간의 평균 픽셀 값이다.

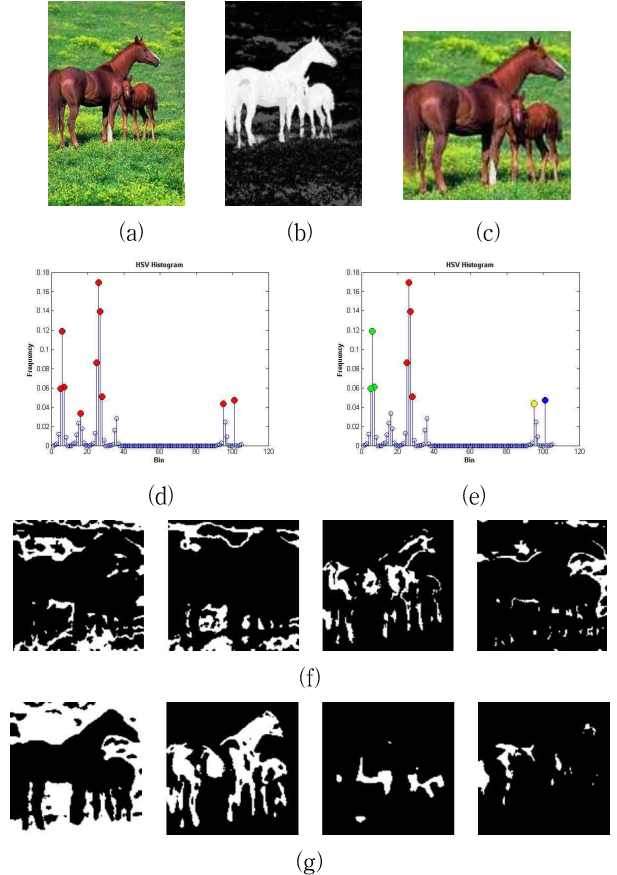


그림 1. 대표 색상의 클러스터링과 이진 공간분포 지도의 예. (a) 쿼리 영상, (b) 쿼리 영상의 saliency map, (c) 객체 중심의 영상, (d) HSV 색상 히스토그램에서의 대표 색상 $C_i (i = 10)$, (e) 클러스터링된 대표 색상 $D_k (k = 4)$, (f) $C_i (i = 1, \dots, 4)$ 에 대한 공간분포 지도, (g) D_k 에 대한 공간분포 지도

3. 실험결과

본 논문에서 제안한 방법을 검증하기 위해서 Core1K 데이터베이스 영상을 이용하였다. 이는 10개의 class 들로 이루어진 1000장의 영상이다. 그림 2, 3은 기존의 방법[4-6]과 제안하는 방법을 사용한 검색 결과로서, 쿼리 영상(왼쪽 윗부분의 영상)과 가장 유사하다고 생각되는 상위 20장의 영상을 뽑은 것을 보여준다. 그림에서 보는 바와 같이, 제안하는 영상 검색 방법은 객체 위주로 하여 대표 색상의 통계적 특성과 공간적 분포를 결합하여 고려하기 때문에 기존의 방법보다 우수한 검색 결과를 보여줄 수 있다.

4. 결론

본 논문에서는 객체 위주의 컬러 영상 검색을 위하여 영상의 saliency map을 이용하여 영상의 크기를 객체 중심으로 정규화하고, 정규화된 영상 간 대표 색상의 통계적 특성과 공간적 분포를 이용하여 유사도를 계산하는 방법을 제안하였다. 대표 색상은 HSV 히스토그램에서 빈도수가 높은 색상 영역들을 클러스터링하여 얻으며, 각 대표 색상이 영상에서 분포하는 상태를 이진 공간 지도로 표시하였다. 그런 다음 추출된 대표 색상이 객체와 객체 주변의 배경에 속하는지 판단하고, 이진 공간 분포 지도를 비교함으로써, 영상을 검색하였다. 본 논문에서 제안한 saliency map을 이용한 대표 색상 기반의 영상 검색 기법은 기존의 대표 색상 기반의 방법들보다 더욱 정확한 결과를 보여주었다.

감사의 글

본 연구는 미래창조과학부 및 정보통신산업진흥원의 대학 IT 연구센터지원사업의 연구결과로 수행되었음 (NIPA-2013-H0301-13-4005).

참고 문헌

- [1] ISO/IEC 15938-3/FDIS Information Technology - Multimedia Content Description Interface - Part 3 Visual, Jul. 2001, ISO/IEC/JTC1/SC29/WG11 Doc. N4358.
- [2] A. Yamada, M. Pickering, S. Jeannin, L.C. Jens, "MPEG-7 Visual Part of Experimentation Model Version 9.0-Part 3 Dominant Color," ISO/IEC JTC1/SC29/WG11/N3914, Pisa, 2001.
- [3] N.-C. Yang, W.-H. Chang, C.-M. Kuo, T.-H. Li, "A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval," *J. Visual Commun. Image Rep.*, 19(20), pp. 92-105, 2008.
- [4] A. Talib, M. Mahmuddin, H. Husni, and L.-E. George, "A weighted dominant color descriptor for content-based image retrieval," *J. Visual Commun. Image Rep.*, 24(3), pp. 345-360, 2013.
- [5] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global Contrast based Salient Region Detection," *Proc. of CVPR*, pp. 409-416, 2011.
- [6] P. Perez, C. Hue, J. Vermaak, and M. Gangnet, "Color-Based Probabilistic Tracking," *Proc. of ECCV*, pp. 661-675, 2002.



그림 2. 실험 결과. (a) MPEG-7 DCD(Dominant Color Descriptor) [2], (b) LBA(Linear Block Algorithm) [3], (c) WDCD(Weighted DCD) [4], (d) 제안하는 방법

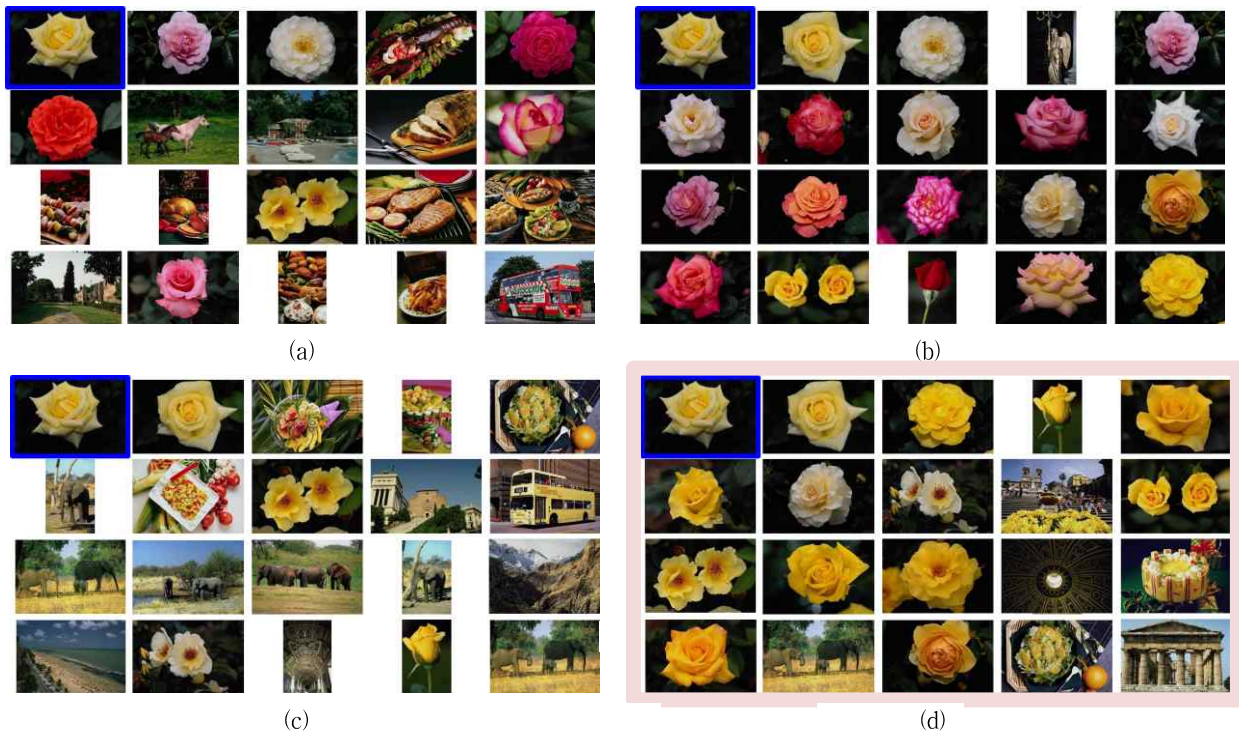


그림 3. 실험 결과. (a) MPEG-7 DCD(Dominant Color Descriptor) [2], (b) LBA(Linear Block Algorithm) [3], (c) WDCD(Weighted DCD) [4], (d) 제안하는 방법