

대표 색상의 공간적 분포를 이용한 영상 검색

안재현 이상화 조남익

서울대학교 전기·컴퓨터 공학부

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

Image Retrieval Using Spatial Distribution of Dominant Colors

An, Jaehyun Lee, Sang Hwa Cho, Nam Ik

School of Electronic Engineering and Computer Science, Seoul National University

요약

본 논문에서는 컬러영상의 검색을 위하여 대표색상의 통계적 특성과 그 공간적 분포를 이용하는 기법을 제안한다. 대표색상은 HSV 색상공간에서 빈도수가 가장 높은 1-8개의 색상영역으로 선택되는데, 대표색상의 값과 각 대표색상이 영상에서 어떻게 분포하느냐에 따라서 검색되는 영상의 종류와 객체의 형태가 크게 달라진다. 본 논문에서는 영상을 일정한 개수의 작은 영역으로 나누고 각 영역에서 각 대표색상이 충분히 존재하는지를 평가하여, 각 대표색상이 영상평면에서 분포한 상태를 이진화된 공간분포 지도로 표시한다. 이로부터 영상간의 대표색상마다 이진공간분포의 차이를 비교함으로써, 색상의 통계적 특성과 공간적 분포가 동시에 반영된 특징으로 영상을 검색한다. 본 논문에서 제안한 대표색상의 공간적 분포 모델을 이용한 영상검색 기법은 MPEG-7에서 정의한 대표색상 기반의 영상 검색보다 우수한 성능을 보여주었다. 이진공간분포 지도의 생성 방법 및 거리측정 파라미터에 대한 최적화가 이루어지면 영상검색 성능이 더욱 개선될 것으로 예상되며, 웹기반의 영상검색 시스템에 활용될 수 있을 것으로 기대된다.

1. 서론

주요 단어의 분포상태를 중심으로 문서를 검색하는 BOW (Bag-of-Words) 방식은 인터넷이나 자료검색에서 높은 정확도를 보여주고 있는 검증된 기술이지만, 영상을 검색하는 기술은 아직 문서검색 만큼의 신뢰도를 보여주지 못하고 있다. 오래전부터 MPEG-7에서는 영상의 색상정보의 통계적 특성을 이용하는 방식으로 영상검색 기술을 개발하였다 [1]. 다양한 색공간의 히스토그램을 이용하는 기존의 방식들은 통계적으로 유사한 색상분포를 갖는 영상들을 검색하는데 효과적이지만, 히스토그램이라는 공간적 속성을 제거한 정보를 사용하기 때문에, 영상속의 객체나 배경과 같은 내용 기반의 영상검색에서는 효과적이지 못한 것으로 보고되고 있다. 최근에는 문서검색에서 사용되는 BOW 개념을 SIFT [8], SURF [9] 같은 표현자 (descriptor) 기반의 특징추출 기술에 적용하여 영상검색을 시도하는 BOF (Bag-of-Feature) 기술에 대한 연구가 활발하게 진행되고 있다 [2, 7]. BOF 기술에서는 영상으로부터 추출된 SURF 또는 SIFT 벡터를 일정한 개수로 양자화 또는 분류하여 시각단어 (visual words)를 우선 추출하고, 시각단어를 문서에서의 단어처럼 간주하여 영상을 문서검색과 같은 방법으로 영상을 찾는다. 그러나 BOF 기반의 영상검색도 아직까지는 좋은 결과를 보여주지 못하고 있는데, 이러한 이유로는 문서에서의 단어와는 달리 시각단어는 영상의 다양성으로 인하여 명확히 규정되기 어려우며, 분류한 시각단어의 종류와 개수에 따라서 검색되는 영상이나 객체가 크게 좌우되는 경향이 있기 때문이다. 또한 SIFT나

SURF 특징벡터가 충분히 검출되지 않거나 객체나 배경의 자세변환이 심한 경우에도 BOF를 이용한 방법으로는 검색이 불가능하다.

본 논문에서는 영상검색에 있어서 MPEG-7에서 정의한 대표 색상의 통계적 특성을 이용하면서, 대표 색상의 공간적인 분포를 함께 고려하여 객체의 형태와 배경을 반영하는 영상검색 기법을 다룬다. 본 논문에서는 사용자가 특정한 대상을 검색하거나 비슷한 분위기의 영상을 검색하는 경우에 초점을 맞추고자 하는데, 이 때 특징점을 이용한 BOF 기반의 기술보다는 색상을 이용한 영상검색이 더 효과적인 경우가 많다. 본 논문에서는 색상정보를 기반으로 유사한 객체나 배경을 중심으로 영상을 검색하는 방법을 중점적으로 개발하고자 하였다.

기존에 색상 기반의 검색 기술을 살펴보면 다음과 같다. 우선, [3]에서는 쿼리 영상과 데이터베이스 영상과의 색상 히스토그램 간의 거리를 계산함으로써 두 영상을 비교하였다. 하지만 히스토그램 자체는 공간적 정보를 포함하고 있지 않으므로 좋은 결과를 보여주지 못한다. 공간적 정보를 이용하기 위해서 [4]에서는 영상 내에서 같은 색상의 응집 여부를 이용한 Color Coherence Vectors를 제안하여 기존의 색상 히스토그램 방법을 개선하였으나, 공간 정보가 극히 적다는 단점이 존재한다. [5]에서는 색 정보를 반영하기 위해서 색상 히스토그램에 대한 모멘트를 이용하였고, 공간 정보 반영을 위해서 웨이블릿 부이미지를 가우시안 분포로 근사한 뒤 그 계수를 사용하였다. 이 방법은 특징량 추출 시 색 정보와 공간 정보를 분리해 다루었기 때문에 영상 간 유사도 정의 시 실험에 의해 각 특징량의 가중치를 결정해야 하는 단점이 있다. 또한 대표 색상을 이용해 유사 영상을 검색하는 MPEG-7의 대

표 색상 기술자는 대표적인 색상 기술자로서, 영상의 색상 분포를 효율적으로 간단하게 표현한 것이다. 하지만 두 영상이 같은 색상 분포를 가지고 있을 때, 공간적 정보를 포함하고 있지 않으므로 매칭이 잘 안 된다는 단점이 있다.

이와 같이 기존 기술의 문제점을 개선하기 위하여 본 논문에서는 대표 색상의 공간적 분포를 표현하고 그 차이를 쉽게 구할 수 있는 공간분포 지도를 제안한다. 본 논문의 구성은 다음과 같다. 2절에서는 대표 색상의 통계적 특성(히스토그램)을 모델링하는 방법을 소개하고, 3절에서는 대표색상의 공간분포를 모델링하는 이진공간분포 지도 생성 방법을 설명한다. 4절에서는 제안한 방법과 기존 MPEG-7 방법을 이용한 비교 실험 결과를 제시하며, 5절에서 본 논문에 대한 향후 개선 사항과 결론을 정리하도록 하겠다.

2. 제안하는 알고리즘

본 논문에서는 쿼리 영상과 유사한 영상을 검색하기 위하여 쿼리 영상에 대한 히스토그램을 HSV 색 공간에서 설계하는 방법과, 대표 색상의 공간적 분포를 이용하여 두 영상의 유사도를 측정하는 방법을 제안한다.

2.1 HSV 색상 히스토그램

히스토그램은 영상의 명암 값의 분포를 보여주기 위해 사용되며, 두 히스토그램의 유사도를 계산함으로써 영상을 비교할 수 있다. 색상 히스토그램이란 색상 공간 (color space) 을 일정 구간으로 나누고, 각 구간에 분포하는 픽셀들의 개수를 측정하는 것으로 다음 식과 같이 표현된다.

$$h_{A,B,C}(a,b,c) = N \cdot \Pr(A = a, B = b, C = c)$$

여기서 A, B, C 는 색상 공간에 해당하는 3 개의 색상 채널을 의미하고, N 은 영상 전체의 픽셀 수를 의미한다.

일반적으로 사용되는 RGB 색 공간은 색상들 간의 의존성이 높고 이들 간의 근접성이 색상의 유사성을 나타내지 못한다는 단점을 가진다. 따라서 이러한 단점을 보완하기 위해 본 논문에서는 HSV 색 공간을 이용한다. 이는 색상(Hue), 채도(Saturation) 그리고 명도(Value) 로 구성되어 있는 색 공간으로, 3 개의 구성 요소의 정의에 따라 HSV 색 공간은 그림 1과 같이 원뿔 모형으로 나타난다.

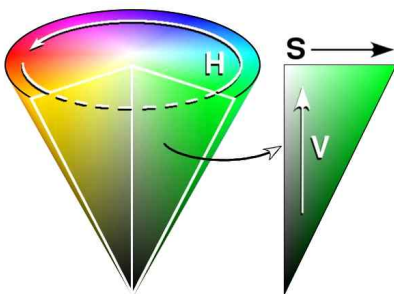


그림 1. HSV 색 공간 원뿔 모형

색상 값 H 는 가장 과장이 긴 빨강을 0도로 하였을 때 상대적인 배치 각도를 의미하고, 0~360도의 범위를 갖는다. 채도 값 S 는 특정한 색상의 가장 진한 상태를 100%로 하였을 때 진하기의 정도를 나타내며, 명도 값 V는 흰색을 100%, 검정을 0%로 하였을 때 밝기의 정도를 나타낸다. 원뿔 모형에서 보이는 바와 같이, 명도 0%는 검정을 의미하므로 단 하나의 점으로 표현되며 원뿔의 꼭지점에 해당한다. 또한 실제 색상은 어두워질수록 채도 값에 따른 색상 변화가 크지 않기 때문에 높은 명도에 비해 채도값이 나타내는 폭이 줄어들는다. 이러한 사실을 이용하면 영상에서 색상 정보는 채도 값과 명도 값이 매우 작지 않을 때에만 유효하다는 것을 알 수 있다. 그러므로 $N_h N_s$ 개수의 빈을 가지는 HS 색상 히스토그램은 채도 값과 명도 값이 특정 문턱치 값 이상인 값 ($T_s = 0.1, T_v = 0.2$) 에 대하여만 생성한다. 나머지 색상 정보가 없는 픽셀은 명도 값으로만 표현이 가능하므로, N_v 개수의 빈을 가지는 V 색상 히스토그램을 생성한다. 결과적으로 생성된 HSV 색상 히스토그램은 $N = N_h N_s + N_v$ 개수의 빈으로 이루어진다 [6]. 이러한 히스토그램 구조를 바탕으로 빈도수가 높은 1-8개의 대표색상을 추출한다.

2.2 대표 색상의 공간적 분포를 이용한 유사도 측정

대표 색상은 주어진 영상에서 가장 두드러지는 색상이 무엇인가 하는 정보이다. 이를 구하기 위한 가장 보편적인 방법은 색상 히스토그램에서 가장 많은 빈도수를 가진 색상을 대표 색상으로 정의한다.

쿼리 영상 Q 의 HSV 색상 히스토그램에서 최소 1 개에서 8 개 사이의 대표 색상 D 를 추출한 후, 영상의 픽셀 값이 각 대표 색상에 속하면 1, 그렇지 않으면 0 인 이진 맵 B^Q 을 생성한다. 쿼리 영상의 대표 색상의 공간적 분포를 나타내는 이진 공간분포 지도 M^Q 는 B^Q 를 $n \times n$ 블록으로 나누고, 각 블록에서 대표 색상에 속하는 픽셀 수가 얼마나 되는가에 의해 결정된다. B^Q 와 M^Q 는 N_i ($i \in D$) 개의 채널로 구성되어 있다. 각 채널에서의 공간분포 지도 M_i^Q 의 픽셀 위치 p 에서의 값은 다음과 같이 정의된다.

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

여기서 $N(A)$ 는 픽셀 p 가 속하는 블록에서의 대표 색상 i 값에 속하는 픽셀의 개수이고, $N(B)$ 는 각 블록에 속하는 모든 픽셀의 개수를 의미한다. 마찬가지로 주어진 데이터베이스의 영상 DB 에 대하여, 쿼리 영상의 대표 색상에 대한 공간적 분포를 나타내는 공간분포 지도 M^{DB} 를 생성한다. 그림 2는 쿼리 영상의 HSV 히스토그램으로부터 추출한 2 개의 대표 색상의 공간분포 지도를 구하는 과정을 보여준다.

결과적으로 쿼리 영상 Q 와 데이터베이스의 영상 DB 간의 유사도 S 는 다음과 같이 계산된다.

$$S_{Q,DB} = \sum_{i \in D} Hit_i(p),$$

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

여기서 $Hit_i(p)$ 는 쿼리 영상과 데이터베이스 영상에서 픽셀 위치 p 에서의 값이 같은 대표 색상에 속하면 1 이고, 그렇지 않으면 0 인 이진 값을 갖는다.

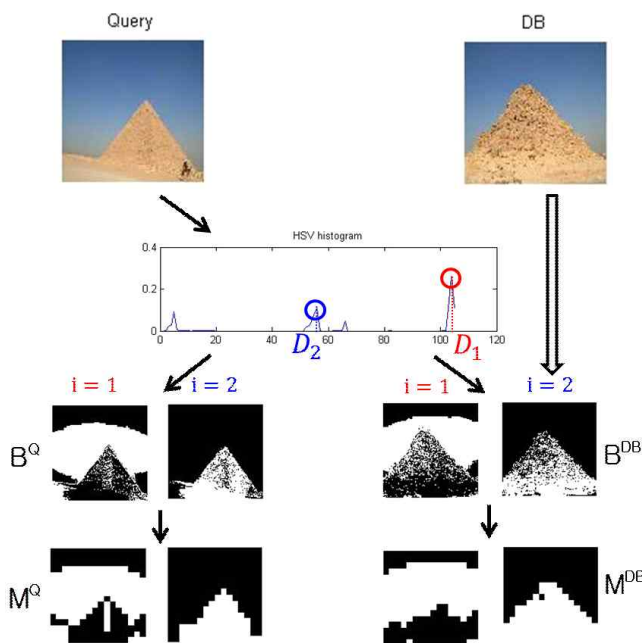


그림 2. 쿼리 영상의 대표 색상에 대한 공간분포 지도 M^Q 와 M^{DB} 를 구하는 과정. D_i ($i = 1, 2$) 는 i 번째 대표 색상을 의미한다.

3. 실험 결과

본 논문에서 제안한 방법을 검증하기 위해 68 장의 쿼리 영상과 318 장의 데이터베이스 영상을 이용하였다. HSV 색 공간에서의 색상 히스토그램의 빈의 개수는 $N_h = 10$, $N_s = 10$ 그리고 $N_v = 5$ 로 $10 \times 10 + 5 = 105$ 개로 양자화 하였고, 대표 색상 D 의 개수는 8 개로 사용하였다. 각 대표 색상의 공간적 분포를 위하여 영상을 16×16 로 분할하고 각 블록에서 각 대표 색상의 값을 가지는 픽셀이 차지하는 비율이 30% ($\gamma = 0.3$) 이상일 때 1의 값을 가지는 이진 맵을 형성하였다. 그림 3, 4 는 MPEG-7 의 대표 색상 기술자를 사용한 결과와 제안하는 방법을 사용한 검색결과로서, 쿼리 영상과 가장 유사하다고 생각되는 상위 10 장의 영상을 뽑은 것을 보여준다. 그림에서 보는 바와 같이, 제안하는 영상 검색 방법은 영상에서의 객체나 배경 형태를 고려하여 대표 색상의 분포를 반영하기 때문에, 기존의 방법보다 우수한 검색 결과를 보여준다. 또한 공간분포지도를 생성하거나

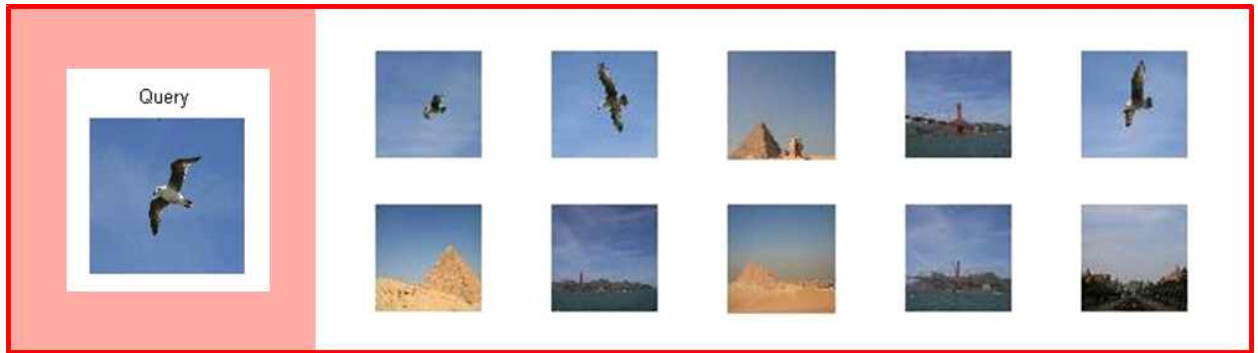
비교하는 과정이 정수연산으로 단순하기 때문에, 모바일 기기에서와 같은 제한된 연산기에서도 큰 제약없이 사용이 가능하여, BOF 영상 검색 기술을 대체할 수 있을 것으로 기대된다.

4. 결론

본 논문에서는 대표색상의 통계적 특성과 그 공간적 분포를 이용하여 영상을 검색하는 기법을 제안하였다. 대표색상은 HSV 색상공간에서 빈도수가 높은 1-8개의 색상영역으로 선택되며, 대표색상이 영상에서 분포하는 상태를 이진공간지도로 표시하였다. 본 논문에서는 영상을 일정한 개수의 작은 영역으로 나누고, 각 영역에서 각 대표색상이 충분히 존재하는지를 평가하여 각 대표색상이 공간적으로 분포한 상태를 이진화된 공간분포 지도로 모델링하였다. 그런 다음 대표색상의 이진공간분포 지도를 비교함으로써, 영상을 검색하였다. 본 논문에서 제안한 대표색상의 공간적 분포 모델을 이용한 영상검색 기법은 MPEG-7에서 정의한 영상 검색보다 더욱 정확한 결과를 보여주었다. 향후 이진공간분포 지도의 생성 방법 및 파라미터에 대한 최적화를 통하여 검색 성능을 개선할 수 있을 것으로 기대되며, 웹기반의 컬러영상 시스템에 적용할 수 있을 것으로 판단된다.

5. 참고문헌

- [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] Josef Sivic, and Andrew Zisserman, "Video Google: A Text Retrieval Approach to Object Matching in Videos," IEEE International Conference on Computer Vision. Oct. 2003.
- [3] M. Swain, and D. Ballard, "Color Indexing." *International Journal of Computer Vision*, vol. 7, No. 1, pp. 11-32, 1991.
- [4] G. Pass and R. Zabih, "Histogram refinement for content-based image retrieval," *Proc. of IEEE Workshop on Applications of Computer Vision*, pp. 96-102, 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] P. Perez, C. Hue, J. Vermaak, and M. Gangnet, "Color-Based Probabilistic Tracking," *Proc. of ECCV*, pp. 661-675, 2002.
- [7] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bag-of-features: spatial pyramid matching for recognizing natural scene categories," *Proc. of CVPR*, pp. 2169-2176, 2006.
- [8] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, pp. 91-110, 2004.
- [9] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, "SURF: Speeded Up Robust Features," *Computer Vision and Image Understanding*, pp. 346-359, 2008.

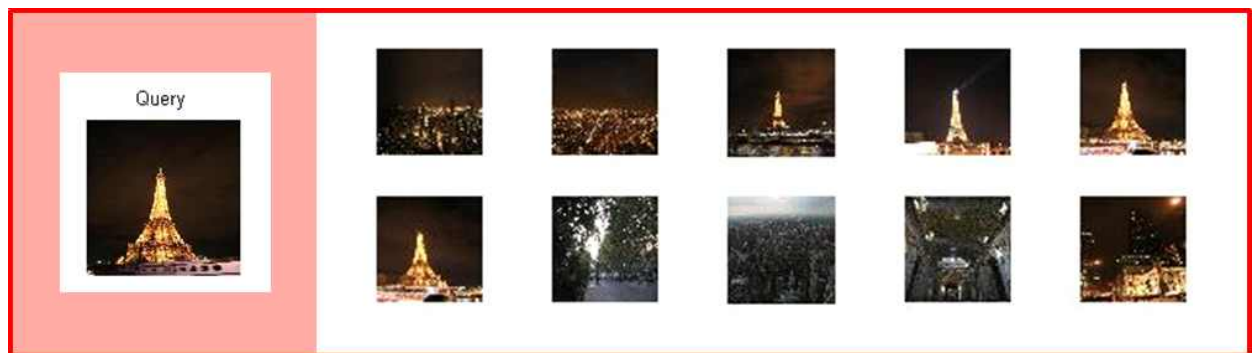


(a)



(b)

그림 3. 실험 결과. (a) 대표 색상 기술자를 사용한 결과, (b) 제안하는 방법을 사용한 결과.



(a)



(b)

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