

# Flexible Subblock을 이용한 영상 검색

고병철<sup>\*</sup>, 이상봉, 이해성, 변혜란  
연세대학교 컴퓨터과학과

## Image Retrieval Using flexible Subblocks

ByongChul Ko, SangBong Lee, Hae-Sung Lee, Hyeran Byun  
Dept. of Computer Science, Yonsei University

### 요 약

본 논문에서는 영상의 flexible subblock을 이용하여 영상내에 물체의 이동이나, 빛의 변화, 시각점(view-point)의 변화 등에 덜 민감한 영상 검색 방법을 제안한다. 특징 값으로는 Ohta 컬러 공간으로부터 1,2,3차 central 모멘트 값을 추출해 내고, 쌍직교 웨이블릿 변환을 통해 고주파 영역으로부터 수직-수평 방향 성분을 추출하여 인덱스화 시킴으로써 인덱싱을 위한 저장공간을 줄이고 계산 시간을 향상시킬 수 있었다. 아울러, 2개의 특징 값을 다단계(multi-step) K-NN 방법에 적용시킴으로써 사용자가 검색하고자 하는 가장 유사한 k 개의 영상만을 사용자에게 보여 주도록 설계하였다. 본 논문에서는 제안하는 알고리즘의 우수성을 증명하기 위해 RGB 색상 공간상에서 히스토그램 교차법(Histogram intersection)을 이용한 방법과, 본 논문의 알고리즘에 RGB, HSI 색상 공간을 그대로 적용하여 실험한 결과를 비교해 보았다. 추가적으로, 영상의 전역적인 유사성뿐만 아니라, 각 블록의 독립적인 특징 값을 이용하여 특정 블록에 대한 검색 환경도 제공하여 보다 의미있는 검색 환경을 제공하고 있다.

## 1. 서 론

최근 몇 년 동안 Visual data의 빠른 증가로 인해 영상의 내용기반 검색에 대한 많은 알고리즘들(QBIC, Netra, VisualSeek, PickToSeek)이 제안되어 오고 있다. 이러한 알고리즘들의 목적은 주어진 이미지나 사용자가 만든 이미지와 가장 유사한 이미지들을 빠른 시간 안에 검색하는 것이다.[1]

일반적으로 특징 값들은 데이터베이스에 인덱스(index)로 저장되게 된다. 따라서 주어진 이미지로부터 각 특징 벡터들 간의 거리를 인덱스로부터 계산하게 되고 이 거리 값이 미리 정의된 임계값 이하인 이미지를 결과 이미지로 보여주게 된다[2].

이때 영상 검색에서 가장 많이 사용되는 알고리즘이 바로 색상에 의한 검색 방법이다. 하지만 때때로 색상 성분은 잘못된 결과를 보여줄 수 있으므로 실제적인 사용에서는 질감(texture)이나 지역적 색상정보에 대한 indexing 방법의 결합이 필요하다. 본 논문의 구성은 다음과 같다. 2장에서는 기존의 색상 인덱싱 방법의 한계와 다른 대안 방법을 소개하고, 3장에서는 본 논문에서 제안하는 새로운 알고리즘을 제안한다. 마지막으로 4장에서 실험 결과 보여주고 5장에서 결론 및 향후 개선 방향을 제시한다.

## 2. 컬러 인덱싱

### 2.1 컬러 히스토그램의 한계

일반적으로 색상 히스토그램은 이미지의 회전, 이동 등에 덜 민감한 것으로 알려져 있다[3]. 하지만 색상 히스토그램을 이용한 인덱싱 방법에서는 시각적으로 비슷한 색상 히스토그램을 가지는 영상이라도 실제로는 검색되지 못하는 경우가 발생한다거나, 이와는 반대로 시각적으로는 전혀 다른 영상이더라도 색상 히스토그램에 대한  $L_1$  Euclidean distance를 계산할 때 동일한 영상으로 판단하는 경우가 발생한다[그림 1].

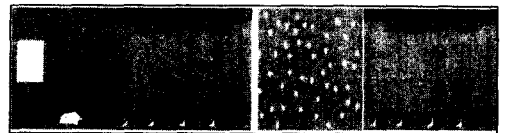


그림 1. 비슷한 히스토그램을 가지는 전혀 다른 영상

### 2.2 영상 검색을 위한 다른 접근 방법

색상 히스토그램 방법의 또 다른 문제점은 인덱싱을 위한 추가적인 저장공간의 필요성과 계산의 복잡성에 있다. 이에 대한 대안적인 방법으로, 히스토그램의 두드러진 특징 등을 이용하는 방법이 있다. 즉 영상의 컬러 분포가 확률 분포로 해석 될 수 있는 모멘트를 이용하는 것이다. 이 방법에서 만약 각 컬러 성분의 첫 번째 3개의 모멘트만을 사용한다면, 각 영상의 인덱스를 표현하기 위해 단지 9개의 부동 소수만이 필요하게 됨으로 인덱스의 수를 줄임으로써 저장공간을 절약할 수 있고, 검색 속도의 향상 및 강건성을 지니는 장점을 가지게 된다. 하지만 이 방법은 low-order의 모멘트가 빛의 변화에 민감한 단점을 가지고 있다[3].

**3. Flexible subblock을 이용한 영상검색 방법**

확률이론에 의해 p.d.f는 각각의 Central Moment(Moment)에 의해 특성화 될 수 있다. 따라서 만약 영상의 색상 분포를 확률분포(Probability distribution)로 해석할 수 있다면, 색상 분포 또한 각각의 모멘트에 의해 특성화 될 수 있다[2].

따라서 본 논문에서는 우선 영상의 지역적 특성을 이용하여 영상을 3X3의 9개 영역으로 분할 한 뒤에, 각각의 영역으로부터 1,2,3차 Central Moment를 Ohta[4] 색상공간(I1,I2,I3)으로부터 계산하여 인덱스를 구축한다. 여기서 1차 모멘트는 평균, 2차 모멘트는 분산, 3차 모멘트는 왜도(Skewness)를 의미한다.

$$\begin{aligned} I1 &= R+G+B \\ I2 &= R-B \\ I3 &= R-2G+B \end{aligned} \quad (1)$$

이렇게 해서 구해진 컬러 인덱스의 수는 한 영상에 대해 3(I1,I2,I3)X9(9개 영역)X3(1,2,3차 모멘트)=81개의 특징 값을 갖게 된다. 다음으로는 영상으로부터 방향성분을 추출한다.

이때 영상을 쌍직교 웨이블릿 프레임(Biorthogonal Wavelet Frame)변환을 통해서 X-Y의 경계선이 추출된 영상을 얻고 이 값들을 인덱스로 저장하게 된다. 이때도 마찬가지로 영상을 3X3의 9개 영역으로 나눈 뒤에 각 영역의 경계선 특징 값을 추출하게 된다. 이렇게 해서 구해진 방향성 인덱스의 수는 한 영상에 대해 9(9개 영역)X2(X-Y 방향 성분)+2(전체 영상의 방향성분)=20개의 특징 값으로 저장된다. 이렇게 추출된 특징 값을 이용하여 질의 영상과 데이터베이스 영상 사이에 9개 블록 중 2개 블록의 오차를 허용하여 동일한 배경을 가지는 영상이라도 빛의 변화나, 시각점(view-point)의 변화, 물체의 움직임등에 의해 잘못 판단 될 수 있는 한계를 극복하였다.

**3.1 유사성 측정**

본 논문에서는 모멘트와 방향성분의 두 가지 특징 값을 이용하여 multi-step k-nearest neighbor 검색 방법을 사용하였다. 첫 번째 단계에서 사용자는 검색하고자 하는 영상의 수를 k로 입력하고, 수식 (2)에 의해 측정된 각 블록의 차이 값이 초기 임계값인  $d_{max}$  이하인 블록의 수가 7개 이상이라면 1차 후보 동일 영상으로 판단하고, 이 후보 영상들을 이용하여 다

시 2차 특징 값인 방향성 성분의 임계값( $d_{max}$ -수식(3))을 결정한 뒤 이 값에 따라 우선 순위(수식 (4))에 의해 k 만큼의 유사 영상을 사용자에게 보여주게 된다.

$$d_{mom}(Q, D) = \sum_{i=1}^r (W_1|E_q^i - E_d^i| + W_2|\sigma_q^i - \sigma_d^i| + W_3|S_q^i - S_d^i|)$$

$d_{mom}$ : 질의 영상과 데이터베이스 영상사이의 대응되는 한 블록의 차이값  
 $r$ : 색상 수

$E-F, \sigma-\zeta, s-t$ : 두 영상간의 1,2,3차 central moments

$$D = \left| \frac{V_q}{H_q} - \frac{V_t}{H_t} \right| \quad (3)$$

$V_q, H_q$ : 질의 영상의 수직, 수평 성분

$V_t, H_t$ : 데이터베이스 영상의 수직, 수평 성분

$$Score = w_1 \cdot G_{mom} + w_2 \cdot D \quad (4)$$

$G_{mom}$ : 전역 모멘트 차이값,  $D$ : 방향성분 차이값

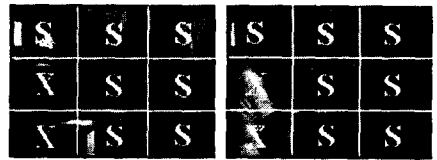


그림 2. flexible subblock에 의한 1차 후보 유사 영상(S:유사 블록, X:비 유사블록)

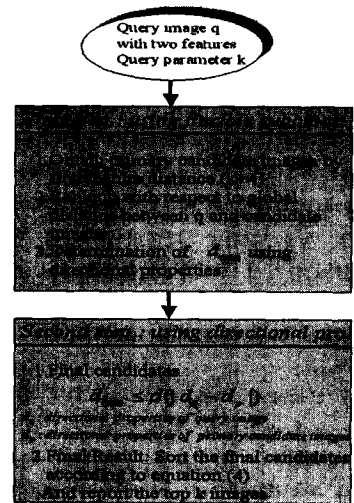


그림 3. 다단계 K-NN 알고리즘

**4. 실험 및 결과**

본 논문에서는 웹으로부터 수집한 1,500개의 영상과, 정확성 테스트를 위해 코넬 대학(<http://cs.cornell.edu/home/rdz/>)

ccv.html)으로부터 얻은 90개의 영상을 사용하였다. 본 논문에서 제안하는 알고리즘의 우수성을 증명하기 위해 RGB 색상 공간상에서 히스토그램 교차법(Histogram intersection)을 이용한 방법과, 본 논문의 알고리즘에 RGB, HSI 색상 공간을 그대로 적용하여 실험한 결과를 비교해 보았다. 이 때 사용된 각 색상 모멘트에 대한 가중치는 표 1과 같다.

	H	S	I
1차 모멘트	3	2	1
2차 모멘트	3	1	1
3차 모멘트	1	1	1

표 1. 각 컬러 공간에 대한 가중치

최종 실험 결과는 그림 4와 같고, 이때 정확율은 recall에 의해 계산하였다. 검색 결과는 그림 5,6과 같다.

본 논문에서는 전역적인 질의 이외에 아울러 영상의 일부분을 블록 단위로 질의 할 수 있는 환경을 제공한다.

4.1 블록에 의한 질의

사용자는 영상 전체에 대한 질의 이외에 사용자가 관심을 갖는 영상 일부분에 대해 검색 할 수 있다. 이때 사용자는 블록의 동일 공간만을 검색하는 absolute 검색과 자신의 블록을 포함하여 나머지 블록들을 모두 검사하는 relative 검색 중 하나를 선택 할 수 있다. 블록에 의한 질의 결과는 그림 7,8과 같다.

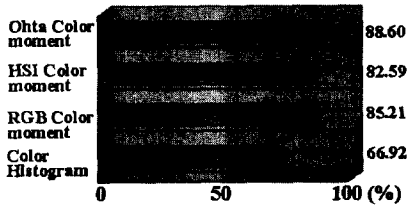


그림 4. 영상 검색 정확율(%)

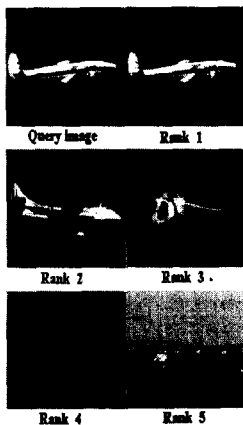


그림 5. 검색결과(k=5)

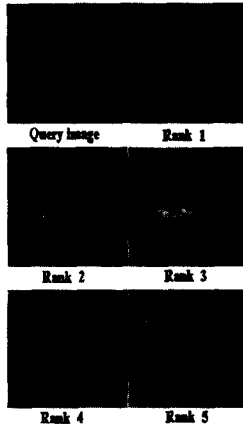


그림 6. 검색결과(k=5)

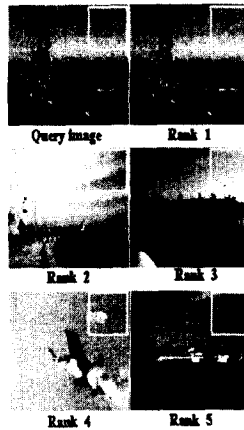


그림 7. Absolute 블록에 대한 검색결과(k=5)

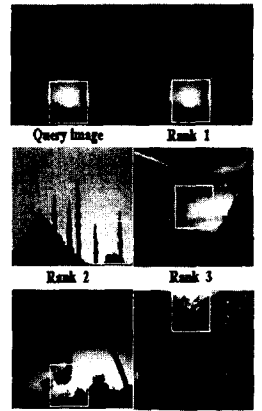


그림 8. Relative 블록에 대한 검색결과(k=5)

5. 결론 및 향후 개선 방향

본 논문에서는 영상의 flexible subblock을 이용하여 영상내에 물체의 이동이나, 빛의 변화, 시각점(view-point)의 변화 등에 덜 민감한 영상 검색 방법을 제안하였다. 또한 이때 Ohta 색상 공간으로부터 1,2,3차 central 모멘트 값을 추출해 내고, 쌍직교 웨이블릿 변환을 통해 고주파 영역으로부터 수직-수평 방향 성분을 추출하여 인덱스화 시킴으로써 인덱스를 위한 저장공간을 줄이고 계산 시간을 향상시킬 수 있었다. 아울러, 2개의 특징 값을 다단계(multi-step) K-NN 방법에 적용 시킴으로서 사용자가 검색하고자 하는 가장 유사한 k개의 영상만을 사용자에게 보여 주도록 하고 있다. 추가적으로, 영상의 전역적인 유사성 뿐만 아니라, 각 블록의 독립적인 특징 값을 이용하여 특정 블록에 대한 검색이 가능하도록 설계하였다.

하지만 보다 의미 있는 내용기반 검색 환경을 제공하기 위해서는 색상 및 방향성 정보 이외에 모양 등에 대한 추가적인 정보를 고려해야 할 것이다.

6. 참고문헌

- [1] Shih-Fu Chang, William Chen, Hari Sundaram, "Semantic Visual Templates: Linking Visual Features to Semantics", *International conference on Image Processing (ICIP '98)*, Chicago, Illinois, October 4-7, 1998.
- [2] Markus Stricker, Markus Orengo, "Similarity of Color Images", *Storage Retrieval Image Video Database III*, 2420, February 1995, 381-392.
- [3] F.Idris and S. Panchanathan, "Review of Image and Video Indexing Techniques", *Journal of Visual Communication and Image Representation*, Vol. 8, No. 2, June, pp. 146-166,1997
- [4] Y-I Ohta, T.Kanade, and T. Sakai, "Color information for region segmentation", *Comp. Grap. And Img. Proc.*,13:222-241, 1980