

칼라 인접성과 기울기를 이용한 내용 기반 영상 검색

김홍염, 이호영, 김희수, 하영호
경북대학교 전자전기공학부
전화 : (053) 940-8835 / 팩스 : (053) 957-1194

Content-based Image Retrieval Using Color Adjacency and Gradient

Hong-Yan Jin, Ho-Young Lee, Hee-Soo Kim and Yeong-Ho Ha
School of Electronics and Electrical Engineering Kyungpook National University
E-mail : yha@ee.knu.ac.kr

Abstract

This paper proposes a color-based image retrieval method using color adjacency and gradient. In proposed method, both the adjacency of different colors and gradient of a color in homogeneous region are considered as features of an image. The gradient, defined as the maximum distance along the direction with largest change of color, is computed for each pixel to determine whether the center color is similar or different to the neighboring colors. Therefore the problems caused by uniform quantization, which is popularly used at most existing retrieval, can be avoided. And furthermore, the storage of the feature is reduced by the proposed binary representation.

I. 서론

칼라 특징(color feature)은 내용 기반 영상 검색에 있어서 중요한 도구이다. 질감(texture)이나 형태(shape) 등 다른 특징(feature)들에 비하여 칼라 정보는 시점(viewpoint)나 영상의 회전 등 외부조건에 영향을 적게 받는다. 칼라 히스토그램(color histogram)은 영상에서 물체의 회전이나 변형 등에 강건한 특징으로 칼라 기반 영상 검색에서 가장 보편적으로 사용되고 있다. 칼라 히스토그램을 이용한 검색 방법[1]은 간단하고 검색속도가 빠른 장점이 있지만 영상에서의 전체적인 칼라 분포만을 나타내기 때문에 물체에 대한 위

치 정보가 고려되지 않는 단점을 가진다. 예를 들어 서로 다른 영상이지만 비슷한 칼라 분포를 가진 경우에 비슷한 영상으로 잘못 판단하게 되고, 같은 물체가 확대되거나 축소되는 경우에도 히스토그램에 큰 변화를 가져오게 되어 서로 다른 영상으로 구별되게 된다. 그리고 히스토그램 구성 과정에서의 균일 양자화로 인해 칼라정보가 왜곡되어 잘못된 검색 결과를 가져오고, 히스토그램 빈(bin)의 크기가 커질 경우에 특징정보의 저장량이 증가하는 등 많은 단점을 가진다. 최근에 위치정보를 고려하는 방법[2]들이 많이 제안되고 있지만 대부분 특징정보의 증가를 수반하며, 특징정보의 추출과 비교과정이 복잡해지게 된다. 최근 멀티미디어 저장 기술의 발달로 영상 데이터베이스의 크기가 증가함에 따라, 영상 검색 과정에서 적은 특징 정보량과 빠른 특징 정보의 추출 및 비교의 중요성은 점차 증대되고 있다.

본 논문에서는 영상에서 나타나는 칼라들의 분포를 표현함에 있어서 경계부분에서의 칼라의 인접성과 평탄한 영역에서의 기울기(gradient)를 고려하여, 비슷한 칼라가 서로 다른 히스토그램 빈에 양자화 됨으로 인하여 생기는 오차를 감소시키는 방법을 제안한다.

II. 제안한 칼라 인접성과 기울기를 이용한 특징 정보 추출

기존의 칼라 인접성을 이용한 graph intersection 방법[3]에서는 양자화된 칼라공간에서 서로 다른 빈에 속하는 칼라들의 인접한 확률을 이용하지만, 균일 양자화에 의하여 비슷한 색들이 다른 색으로 잘못 구분

될 수 있다. 그래서 본 논문에서는 인접화소들의 실제 색차를 계산하여 일정한 크기 이상의 색차를 가지는 칼라들만을 시각적으로 다른 색으로 분류하도록 하였다. 아래 블록도는 제안한 특징 추출 과정을 나타낸다.

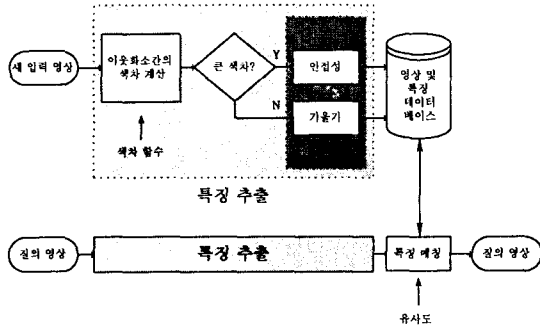


그림 1. 알고리즘 블록도

영상의 각 화소들에 대해 3×3 마스크를 적용하여 주변화소들과의 색차를 구하고 그 중 가장 큰 값을 일정한 문턱치와 비교한다. 문턱치보다 큰 경우에는 그 색차를 가지는 두 칼라가 경계부분에 위치하고 있으므로, 그림2에서처럼 인접성을 나타내는 2차원 배열에 표시한다. 문턱치보다 작은 경우에는 인접화소들이 비슷한 칼라를 가진 평탄한 영역이므로 그 색차를 중심 화소의 기울기[4]로 하여 그림3과 같이 기울기 분포에 표시한다.

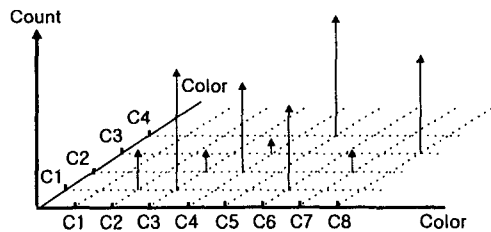


그림 2. 인접성 정보

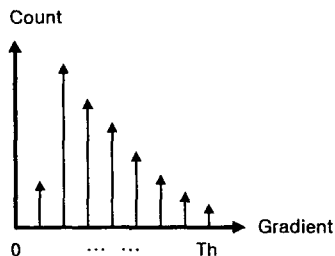


그림 3. 기울기 분포

그림 3에서 기울기의 최대치는 색차를 분류하는 문턱치 Th 로 하였다.

화소들간의 색차를 구하고 문턱치를 결정함에 있어서 본 논문에서는 인간의 시각특성에 더욱 적합한 vector-angular distance 방법을 사용하였다. RGB 공간에서 보편적으로 사용되는 유클리디언 공식을 이용할 경우에 칼라간의 색차는 단순히 두 칼라의 공간적 거리로 나타나므로 실제적인 색차를 표현하지 못한다. 그림4는 RG 평면($B=0$)에서의 칼라 분포를 나타내는데, $P_1 P_2$ 와 $P_3 P_4$ 는 같은 유클리디언 거리를 가지지만 P_1 과 P_2 위치의 칼라는 거의 비슷하고 P_3 과 P_4 위치에서는 서로 다른 칼라로 나타난다는 것을 알 수 있다.

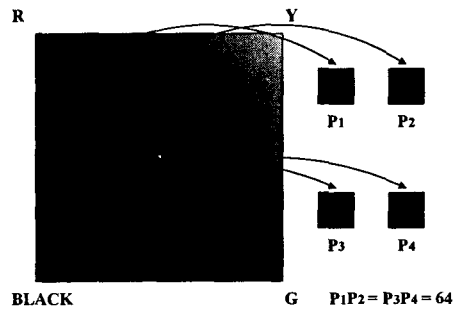


그림 4. 유클리디언 거리 및 시각적 색차

본 논문에서 사용한 vector-angular distance[5]는 두 칼라 벡터가 이루고 있는 각도와 실제 색차의 관계를 고려한 것으로 RGB 공간상의 두 점 $P_1 P_2$ 사이의 정규화된 색차를 다음과 같이 구한다.

$$\delta(P_1, P_2) = 1 - \left[1 - \frac{2}{\pi} \cos^{-1} \left(\frac{P_1 \cdot P_2}{|P_1| |P_2|} \right) \right] \cdot \left[1 - \frac{|P_1 \cdot P_2|}{\sqrt{3 \cdot 255^2}} \right] \quad (1)$$

수식으로부터 벡터간의 각도와 두 점사이의 거리가 동시에 색차에 영향을 주고, 각도가 클수록, 거리가 멀수록 색차는 커지게 된다는 것을 알 수 있다.

영상의 전체 화소들에 대해 양자화하는 대신 vector-angular distance를 구하여 영상의 에지 특성을 구하고, 경계부분의 화소와 평탄한 영역의 화소를 더욱 정확히 구분함으로써 균일 양자화에 의해 칼라 정보가 왜곡되는 것을 피할 수 있다.

III. 특징 정보 표시와 비교

2차원 배열로 나타내는 인접성 정보는 bin(bin)수가 커짐에 따라 방대한 양의 저장량을 필요로 하게 된다. 본 논문에서는 기존의 방법에서처럼 히스토그램의 실제 값들을 저장하고 질의 영상과 데이터베이스 영상의 대응하는 히스토그램 bin들을 모두 비교하는 것이 아니라, 일부 주요 성분들만을 선택하여 영상의 에지 특성을 이진 배열로 표시하여 대응하는 이진 부호를 비교한다.

그림2에서 전체 에지 정보의 일정한 비례를 차지하는 주요 성분들을 선택하여 부호 '1'로, 나머지를 '0'으로 표시하여 전체 영상에서의 인접성 정보를 다음과 같이 이진 배열로 나타낸다.

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \dots & \dots & & & & & & \end{bmatrix} \quad (2)$$

즉

$$C = \begin{bmatrix} 36 \\ 16 \\ 1 \\ 8 \\ \vdots \end{bmatrix} \quad (3)$$

이렇게 함으로써 특징 정보를 훨씬 적은 양으로 저장하고 빠른 속도로 비교할 수 있다. 이진 배열간의 비교는 graph intersection 방법[3]에서의 유사도(similarity measure) 수식을 적용하여 배열에 들어있는 부호 '1' 즉 주요 에지의 개수를 고려한다.

$$S(I, M) = \frac{NC(I, M)}{N(I) + N(M) - NC(I, M)} \quad (4)$$

식에서 $N(I)$ 는 이진 배열 I 에 나타나는 부호 '1'의 개수를 표시하고 $NC(I, M)$ 은 두 배열 I, M 에서 동일한 위치에 "1"이 있는 경우 즉 동일한 칼라로 구성된 에지의 개수를 표시한다. 예를 들어 영상 I, M 에서 i 번째 칼라의 이진 배열이 각각 아래와 같을 때

$$B_c(I) = (1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1 \ 1 \ 0) \\ B_c(M) = (0 \ 1 \ 0 \ 1 \ 1 \ 1 \ 0 \ 1 \ 1)$$

그 유사도는 다음과 같이 구해진다.

$$NC_i(I, M) = 2, \quad N_i(I) = 4, \quad N_i(M) = 5 \\ S_i(I, M) = \frac{2}{5+4-2} \approx 0.28$$

기울기 히스토그램의 비교에서는 히스토그램 인터섹션(histogram intersection) 방법에서처럼 아래와 같은 유사도를 이용하였다.

$$S(I, M) = \frac{\sum_{j=1}^8 \min(I_j, M_j)}{\sum_{j=1}^8 M_j} \quad (5)$$

질의 영상 I 와 모델 영상 M 의 기울기 히스토그램에서 각 대응 bin I_j 와 M_j ($j=1, 2, \dots, 8$)들의 공통부분을 계산하고 모델 영상으로 정규화한 것이다.

인접성 성분(I^a)과 기울기 성분(I^g)에 대한 전체 유사도는 다음과 같다.

$$S(I, M) = \frac{NC(I^a, M^a)}{N(I^a) + N(M^a) - NC(I^a, M^a)} + \frac{\sum_{j=1}^8 \min(I_j^g, M_j^g)}{\sum_{j=1}^8 M_j^g} \quad (6)$$

IV. 실험 결과

영상 검색에서 결과의 상대평가를 위하여 보편적으로 사용하는 검색 정확도(retrieval accuracy)는 recall과 precision이 있다.

$$Recall = \frac{\text{relevant retrieved}}{\text{all relevant}} \quad (7)$$

$$Precision = \frac{\text{relevant retrieved}}{\text{all retrieved}} \quad (8)$$

Recall은 검색된 관련 영상(relevant image)의 개수가 전체 관련 영상에서 차지하는 비례를 나타내고 precision은 검색된 관련 영상의 개수가 모든 검색된 영상에서의 비례이다.

본 논문에서는 검색된 영상의 rank까지 고려한 정확도 측정 방법인 ANMRR[6]을 사용하였다. 각 질의 영상 q 에 대하여 관련 영상의 개수가 $NG(q)$ 일 때 검색 영상의 개수는 아래와 같다.

$$K(q) = \text{Min}\{4 \cdot NG(q), 2 \cdot GTM\} \quad (9)$$

$$GTM = \text{Max}_q \{NG(q)\} \quad (10)$$

첫 $K(q)$ 개의 결과 영상 중 정확히 검색된 관련 영상들의 rank를 기록하고 검색되지 못한 관련 영상들에게는 $K(q)+1$ 값을 주어 평균치를 구하고, 다시 $[0,1]$ 로 정규화하여 Q 개의 질의에 대해 평균을 구한다.

$$AVR(q) = \sum_{j=1}^{NG(q)} \frac{Rank(j)}{NG(q)} \quad (11)$$

$$ANMRR = \frac{1}{Q} \sum_{q=1}^Q \frac{AVR(q) - 0.5 - \frac{NG(q)}{2}}{K(q) + 0.5 - \frac{NG(q)}{2}} \quad (12)$$

관련영상들의 rank가 높을수록 ANMRR는 작아진다. 본 논문에서는 서로 다른 크기를 가진 5000여 개의 자연영상을 포함한 데이터베이스에서 50개의 질의 영상에 대하여 제안한 방법을 적용한 결과 기존의 방법들보다 좋은 정확도를 얻었다.

표 1. 검색 정확도 비교

	Recall	Precision	ANMRR
Histogram intersection	0.82	0.21	0.22
Graph intersection	0.89	0.22	0.17
Proposed	0.92	0.24	0.11

표1은 칼라 분포를 64개의 칼라로 표시할 때 제안한 방법과 히스토그램 인터섹션 및 graph intersection 방법을 비교한 것이다. 제안한 방법은 평균 30여개의 특징 성분만으로 세 가지 정확도에서 모두 가장 좋은 결과를 가져왔다.

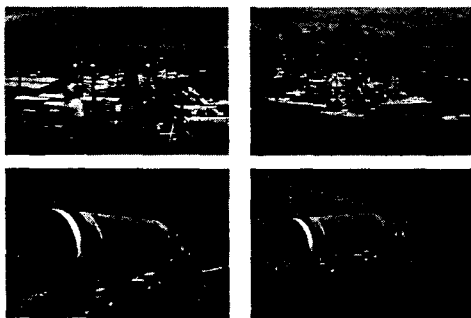


그림5. 질의 영상과 관련 영상

그림5의 왼쪽 부분은 두 개의 질의 영상이고 오른쪽 부분은 각 질의와 관련되는 영상으로 기존의 히스토그

램 인터섹션 방법과 graph intersection 방법에서 검색되지 못한 영상들이다. 하지만 제안한 방법을 사용하여 top rank로 정확히 검색되었다. 제안한 방법이 영상의 내용에 근거하여 유사성을 판단하는 인간의 시각적 특성에 더욱 적합하고 여러 가지 외부조건 특히 물체의 부재(absence)나 영상의 확대, 축소(zoom in/out)등에 강건하다는 것을 알 수 있다.

V. 결론

이 논문에서는 내용 기반 영상 검색에서 칼라의 인접성을 이용하여 영상의 에지 특성을 반영하고 동시에 평탄한 영역의 칼라 기울기를 고려하여 비슷한 칼라가 서로 다른 양자화 축에 들어감으로 생기는 오차를 감소시켰다. 영상의 특징을 표현함에 있어서 기존의 방법들처럼 히스토그램 빈들의 실제 값을 모두 저장하고 비교한 것이 아니라 주요한 성분만을 선택하여 이진수로 나타내어 비교하였다. 이렇게 함으로써 영상 특징 정보의 양을 줄이고 검색 속도를 증가시키는 동시에 여러 가지 외부조건에 강건한 검색 결과를 얻었다.

참고문헌

- [1] M.J.Swain and D.H.Ballard, "Color indexing," *Int.J.Computer Vision*, 7(1), 1991, pp. 11-32.
- [2] J.Huang, S.R.Kumar, M.Mitra, W.J.Zhu, and R.Zabih, "Image indexing using color correlograms," *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, June 1997, pp. 762-768.
- [3] I.K.Park, I.D.Yun, and S.U.Lee, "Color image retrieval using hybrid graph representation," *Image and Vision Computing*, 17, 1999, pp. 465-474.
- [4] G.Pass, and R.Zabih, "Comparing images using joint histograms," *Multimedia Systems* 7, 1999, pp. 234-20.
- [5] D.Androutsos, K.N. Plataniotis, and A.N. Venetsanopoulos, "A vector angular distance measure for indexing and retrieval of color," *Storage & Retrieval for Image and Video Databases VII*, SPIE-3656, San Jose, USA, 1999, pp. 604-613.
- [6] ISO/IEC JTC1/SC29/WG1/ "Core experiment on MPEG-7 color and texture descriptors," *Doc. N2819, MPEG Vancouver Meeting*, July, 1999.