

칼라 벡터각을 이용한 칼라 기반 영상 검색과 위치 추정

정회원 이 호 영*, 이 호 근*, 김 윤 태*, 남 재 열**, 하 영 호*

Color-Based Image Retrieval and Localization using Color Vector Angle

Ho-Young Lee*, Ho-Keun Lee*, Youn-Tae Kim*, Jae-Yeal Nam**,
and Yeong-Ho Ha* *Regular Members*

요 약

칼라가 물체 인식에 아주 효율적인 단서를 제공하지만 칼라 분포는 시청 조건과 카메라의 위치에 아주 큰 영향을 받는다. 생김새와 모양의 변화에 의한 칼라 분포 변화 문제를 해결하기 위해 본 논문에서는 밝기값의 변화에 영향을 받지 않고, 색상(hue) 성분에 민감한 칼라 벡터각(color vector angle)을 이용하여 칼라 에지를 추출한 후, 영상의 화소들을 평탄 화소와 에지 화소로 구분하여 칼라 특징값을 추출하였다. 에지 화소의 경우에는 에지 주위 칼라 쌍의 전체 분포를 *HLS* 색좌표계의 비균일 양자화를 통해 칼라 인접 히스토그램(color adjacency histogram)으로 표현하고, 평탄 화소의 경우에는 *HLS* 색좌표계의 비균일 양자화와 칼라 벡터각 균일 양자화를 통해 칼라 벡터각 히스토그램(color vector angle histogram)을 구성하여 공간적인 칼라분포를 표현하였다. 제안한 칼라 히스토그램을 이용하여 영상 검색에 적용하여 성능을 실험한 결과, 작은 bin의 수를 가지는 제안한 방법이 기존의 방법들보다 훨씬 효율적이고, 생김새와 모양의 변화에 아주 강건한 영상 검색이 가능하였고, 기존의 칼라 히스토그램 역투사 방법보다 훨씬 정확한 물체 위치 추정이 가능함을 확인할 수 있었다.

ABSTRACT

Even if color can provide the efficient visual cue for object recognition, color distribution is largely affected by viewing condition and camera position. To overcome the problem of color distribution changes caused by change of appearance and shape, we use the color vector angle, which is insensitive to variations in intensity, yet sensitive to differences in hue. First, color edge detection based on color vector angle is performed to classify the pixels of image into smooth and edge pixels. For edge pixels, the global distribution of pairs of colors around the edge is represented by color adjacency histogram in non-uniformly quantized *HLS* color space. For smooth pixels, color distribution is represented by color vector angle histogram, which is constructed by non-uniform quantization of *HLS* color space and uniform quantization of color vector angle. Experimental results on set comprising 5466 model images and 50 queries shows that the proposed image retrieval with small number of bins is more effective and robust to changes in appearance and shape than other methods. The result of object localization is also more accurate than conventional color histogram backprojection.

1. 서론

영상에서 칼라 분포는 물체 인식과 영상 검색에 유용하게 사용되어져 왔다. Swain의 칼라 인덱싱^[1]

을 기반으로 현재까지 칼라 분포는 많은 영상 검색 시스템의 중요 부분을 구성하고 있다. 영상에서 칼라 분포는 물체를 구별하는데 많은 도움을 줄뿐만 아니라, 물체의 회전과 이동에 많은 영향을 받지 않

* 경북대학교 전자전기공학부 칼라 및 영상처리 연구실(hylee@m80.knu.ac.kr), ** 계명대학교 공학부 컴퓨터공학과
논문번호 : K01078-0220, 접수일자 : 2001년 2월 20일

는 장점을 가진다. 대부분의 통계적인 방법들^[3,15,17]이 작은 크기의 특징 데이터(feature dataset)를 가진다 할 지라도, 신뢰성 있는 칼라 통계치를 얻기 위해서는 많은 연산 시간을 필요로 한다. 영상 검색을 용이하게 하기 위해서는 히스토그램이 반드시 영상에서 두드러진 칼라 분포 특징을 정확하게 표현할 수 있어야 하며, 히스토그램의 빈 크기(histogram bin-size)가 작으면 작을수록 더 빠른 속도로 영상의 검색을 수행할 수 있다.

칼라 히스토그램을 사용하는데 있어서 문제점은 히스토그램 구성 과정에서의 공간 정보 부족을 들 수 있다. 비록 동일한 물체를 촬영한 영상이라 할 지라도 서로 다른 관측 지점에서 촬영이 이루어질 경우에 두 영상은 서로 다른 칼라 히스토그램을 가지게 된다. 이러한 문제점은 데이터베이스의 크기가 아주 클 경우에 검색 결과에 아주 민감하게 반응하여, 대부분의 경우에 잘못된 검색결과를 가져오게 된다. 최근에 이러한 칼라 히스토그램의 문제점을 해결하고자 많은 방법들이 제안되어졌다.^[3,8,10] 제안된 대부분의 방법들은 영상을 몇 개의 부영역(sub-region)으로 나누거나, 영상을 비교할 때 공간적인 제한 조건을 주는 방법들이 대부분이었다. 그러나 영상을 부영역으로 분할하는 방법은 화소의 위치에 의존하므로 영상에서 아주 큰 생김새(appearance)나 모양(shape)의 변화가 있을 경우에는 검색의 결과가 아주 떨어지는 단점을 가지게 된다. 또 다른 접근 방법으로는 거리의 함수로 칼라의 공간적인 상관도 변화를 고려하는 correlogram 방법이 있다.^[3] 이 방법은 칼라들 사이의 국부적인 공간 상관도를 고려할 뿐만 아니라, 전체적인 칼라 분포를 잘 표현하므로 아주 우수한 검색 결과를 나타내지만 correlogram을 계산하는데 있어서 과도한 연산 시간을 필요로 하는 단점을 가진다.

제안한 방법에서는 칼라 벡터각을 이용하여 영상에서 생김새나 모양의 변화가 발생할 경우에 나타나는 검색 효율의 저하 문제를 해결하였다. 본 논문에서는 먼저 색상과 채도의 변화에 민감한 칼라 벡터각^[4]을 이용하여 칼라 에지를 검출하였다. 생김새나 모양의 변화가 발생한다 할지라도 칼라 에지를 구성하는 칼라 쌍(color pair)은 변하지 않게 되므로, 에지 화소에 대해서는 에지를 이루는 칼라 쌍에 대해 칼라 인접 히스토그램(color adjacency histogram)을 구성하고, 평탄 화소에 대해서는 칼라 벡터각을 균일 양자화하여 칼라 벡터각 히스토그램(color vector angle histogram)을 구성하였다. 제안

한 칼라 인접 히스토그램과 칼라 벡터각 히스토그램은 HLS 칼라 좌표계의 비균일 양자화를 통해 구성되므로 유사한 칼라가 히스토그램 구성 시에 서로 다른 빈으로 할당되는 것을 방지할 수 있다.

Swain은 영상에 존재하는 물체의 위치를 찾아내기 위한 방법으로 칼라 히스토그램 역투사(color histogram backprojection) 방법을 제안하였는데, 이 방법은 영상의 해상도와 히스토그램의 해상도의 영향에 민감하지 않을 뿐 아니라, 아주 적은 연산량을 가지고서도 효율적으로 물체의 위치를 추정할 수 있는 방법이다. 그러나 히스토그램 역투사 방법은 영상내의 동일한 칼라를 가지는 모든 화소가 동일한 비 히스토그램(ratio histogram)를 가지게 되므로, 공간적인 칼라 분포 정보를 가지지 못하는 기존의 칼라 히스토그램을 이용할 경우에 위치 추정의 정확도가 아주 떨어지는 단점을 가진다. 그래서 본 논문에서는 제안한 칼라 인접 히스토그램과 칼라 벡터각 히스토그램을 이용하여 공간 정보 부족의 문제점을 해결하였다.

본 논문의 구성은 먼저 II장에서 제안한 영상 검색 방법의 전체적인 흐름을 살펴보고, III장에서 칼라 벡터각을 이용한 칼라 에지 추출과 칼라 인접 히스토그램 및 칼라 벡터각 히스토그램의 구성 방법에 대해 살펴본다. IV, V장에서 히스토그램 역투사 방법을 이용한 목표 물체 위치 추정과 제안한 칼라 특징을 이용한 유사도 측정에 대해 각각 살펴본 후, 마지막으로 다양한 영상들로 구성된 영상 데이터 베이스를 이용한 실험 결과를 살펴본다.

II. 제안한 영상 검색 시스템의 구성

그림 1은 제안한 칼라 히스토그램의 구성을 나타낸다. 먼저 입력 화소에 대해 입력 화소를 중심 화소로 하는 3×3 마스크를 적용하여 중심 화소와 8개의 이웃 화소가 이루는 칼라 벡터각들 가운데 최대값을 구한다. 구해진 최대 벡터각이 문턱값보다 작은 경우에는 색차(color difference)가 거의 없는 평탄 화소로 간주하고, 그렇지 않은 경우에는 색차가 큰 에지 화소로 분류한다. 에지 화소에 대해서는 칼라 인접 히스토그램을 통해 표현하고, 평탄 화소에 대해서는 칼라 벡터각 히스토그램을 구성한다. 입력 영상 내의 모든 화소에 대해 히스토그램을 구성하여 입력 영상에 대한 최종적인 공간적인 칼라 분포를 구하게 된다.

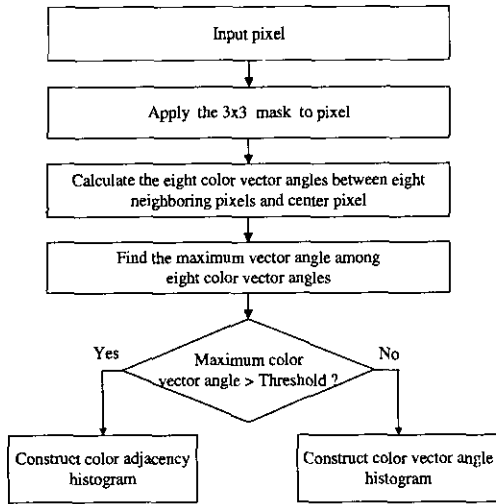


그림 1. 제안한 칼라 히스토그램 구성의 블록도

III. 제안한 칼라 특징 정보 추출

3.1 벡터각을 이용한 칼라 에지 추출

RGB 칼라 영상의 경우, RGB 각 채널의 값들이 서로 높은 상관 관계를 가지므로, RGB 공간에서의 단순한 유클리디언 거리(Euclidean distance)로 색차를 구한다는 것은 불가능하다.^[4] 유클리디언 거리는 밝기값(intensity)의 변화에 매우 민감하므로, 동일한 칼라를 가지는 영역의 두 화소가 광원 조사 방향의 변화가 존재할 경우에 아주 큰 유클리디언 거리값을 가지게 된다. 일반적인 영상 획득과정에서 각 화소 위치에서의 값은 물체에 조사되는 광원(E)과 물체 표면 반사 계수(ρ)의 곱으로 표현된다.

$$L(x, y) = E(x, y)\rho(x, y) \quad (1)$$

그래서 RGB 색공간에서 두 화소가 동일한 표면 반사 계수를 가지는 영역의 화소들이고, 동일한 칼라를 가지지만 그 밝기값이 다른 유색 광원이 조사된다면, 각 위치에서의 값은

$$V_1 = aV_2, \quad a > 0 \quad (2)$$

와 같은 관계를 가진다. 비록 두 화소가 동일한 표면 반사 계수를 가지는 영역의 화소이지만 두 화소 사이의 유클리디언 거리는 광원의 밝기값 차이로 인하여 상당한 차이를 가지게 된다.

그래서 제안한 방법에서는 RGB 색공간에서 다른 색공간으로의 변환 없이 광원의 밝기값에 영향을 받지 않고, 두 화소사이의 색상(hue)과 채도(satura-

tion)를 잘 구분해낼 수 있는 칼라 벡터각^[4]을 이용하여 색차를 구하였다. 두 칼라 화소사이의 벡터각을 구하는 과정은 아래와 같다.

$$\sin(\theta)_{v_1, v_2} = \left(1 - \frac{(V_1^T V_2)^2}{V_1^T V_1 V_2^T V_2}\right)^{1/2} \quad (3)$$

여기서 V_1, V_2 는 RGB 공간상에서의 두 화소이다. 정확한 벡터각을 구하기 위해서는

$$\theta_{v_1, v_2} = \sin^{-1}\left(1 - \frac{(V_1^T V_2)^2}{V_1^T V_1 V_2^T V_2}\right)^{1/2} \quad (4)$$

를 구해야 되지만 RGB 좌표계에서 최대 벡터각의 크기가 90° 이고, sine 함수의 경우에 $0^\circ \sim 90^\circ$ 사이에서는 증가함수이므로, 본 논문에서는 식(3)를 이용하여 벡터각 기반의 색차를 구하였다. 그림 2는 4가지의 칼라 샘플에 대한 벡터각을 보여준다. 칼라 A, B와 C, D는 각각 동일한 색상을 가지지만 밝기값이 차이가 있는 화소들을 각각 나타내는데, 동일 색상을 가지더라도 밝기값의 차이에 따라 상당한 유클리디언 거리 차이를 가지고, 서로 다른 색상임에도 불구하고 동일한 유클리디언 거리 차이를 가지는 것을 확인할 수 있다. 벡터각의 경우에는 색상이나 채도의 차이가 나는 경우에는 벡터각의 차이가 발생하나, 동일 색상이지만 밝기값이 차이가 나는 경우에는 두 화소가 동일 직선 상에 놓이게 되므로 벡터각의 차이가 발생하지 않게 된다.

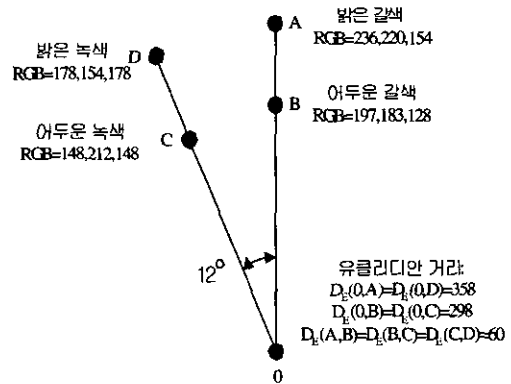


그림 2. 4가지 칼라 샘플에 대한 벡터각의 평면 기하학

제안한 방법에서는 영상이 입력되면 3×3 마스크 내의 중심화소와 이웃한 8개의 화소들과의 벡터각을 구한 후, 총 8개의 벡터각 가운데서 최대 벡터각이 문턱값보다 높은 경우에 중심화소를 에지 화소

로 간주하였다. 중심 화소가 에지 화소로 분류될 경우에는 최대 벡터각을 구성하는 중심 화소와 이웃 화소를 가지고서 칼라 인접 히스토그램을 구성하게 된다.

3.2 HLS 칼라 좌표계의 비균일 양자화

많은 연구들이 히스토그램의 크기를 줄이고, 색상 성분을 효과적으로 나타내기 위해 CIELUV와 CIELAB와 같은 균일 좌표계를 이용한다. 그러나 이러한 균일 좌표계들은 RGB 좌표계로부터 직접적으로 구할 수 없을 뿐만 아니라, 좌표계 변환 시에 많은 연산량을 필요로 한다. 그래서 본 논문에서는 히스토그램을 구성함에 있어서 인간 시각이 칼라를 인식하는 과정과 유사하고, RGB 값으로부터 이주적인 연산으로 변화가 가능한 HLS 좌표계를 이용하였다. 그림 3은 빨간색 색상 성분에 대한 채도와 밝기값의 분포를 나타내는데, 채도값에 따라서 인지되는 칼라가 서로 달라지는 것을 알 수 있다. 그래서 본 논문에서는 각 칼라 성분을 채도값에 따라 무채색 영역(achromatic area), 낮은 채도 영역(low chromatic area), 높은 채도 영역(high chromatic area)으로 구성된 세 개의 부영역(subregion)으로 나누어서 각 부영역에 따라 서로 다른 성분을 이용하여 양자화를 하였다. 무채색 영역(그림 3의 영역 1)의 경우에는 화소가 색상 성분을 가지더라도 인간 시각의 인식에는 밝기값(L)이 중요하게 되므로, 밝기값을 이용하여 양자화 하였으며, 낮은 채도 영역(그림 3의 영역 2)의 경우에는 인간 시각이 어느 정도의 색상 성분을 인지할 뿐만 아니라 밝기값에 따라 인간 시각에 인지되는 칼라가 달라지므로 색상 성분(H), 밝기값(L)를 이용하여 양자화 하였다. 높은 채도 영역(그림 3의 영역 3)의 경우에는 밝기값과 채도값에 관계없이 색상 성분에 따라 인간 시각에 칼라 인식되므로 색상 성분(H)을 양자화 하였다.



그림 3. 채도와 밝기값에 따른 칼라 분포의 예

제안한 비균일 양자화는 총 32개의 대표 칼라(representative color)를 구성하게 되는데, 각 부영역에서 HLS 좌표계의 각 성분에 대한 양자화 레벨의 수는 표 1과 같다. 각 부영역의 양자화에서 밝기값은 균일 양자화를 이용하였고, 색상 성분에 대해서는 칼라 명명법(color naming)^[18]의 기준이 되는 7개의 칼라로 비균일 양자화 하였다. 식 (5)는 색상 성분의 비균일 양자화를 나타낸다.

$$H = \begin{cases} 0 & \text{if } 330 < Hue \leq 22 \\ 1 & \text{if } 22 < Hue \leq 45 \\ 2 & \text{if } 45 < Hue \leq 70 \\ 3 & \text{if } 70 < Hue \leq 155 \\ 4 & \text{if } 155 < Hue \leq 186 \\ 5 & \text{if } 186 < Hue \leq 278 \\ 6 & \text{if } 278 < Hue \leq 330 \end{cases} \quad (5)$$

표 1. 세 부영역에서 각 성분에 대한 양자화 레벨의 수

영역 \ 칼라 수	무채색 영역			낮은 채도 영역			높은 채도 영역		
	H	L	S	H	L	S	H	L	S
32	1	4	1	7	3	1	7	1	1

3.3 칼라 인접 히스토그램의 구성

3x3 마스크의 중심 화소가 에지 화소로 분류되면, 중심 화소와 최대 벡터각을 이루게 되는 3x3 마스크내의 이웃 화소를 이용하여 칼라 인접 히스토그램을 구성하게 된다. 그림 4는 칼라 인접 히스토그램의 구성 예를 나타낸다. 그림 4(a)에서 굵은 실선으로 표시된 화소는 에지 화소로 판명된 화소를 나타내고, 화살표는 에지 중심 화소에 대해 3x3 마스크 내에서 최대 벡터각을 구성하는 중심 화소와 이웃 화소를 연결한다. 그림 4(b)에서 x축과 y축은 각각 중심 화소와 최대 벡터각을 이루는 이웃 화소에 대해 제안한 비균일 양자화 방법으로 양자화된 대표 칼라(C_i)를 나타낸다. 그림 4(a)의 경우 6개의 칼라 에지 화소가 존재하는데, 6개의 에지 화소 모두 C₂ 대표 칼라를 가진다. 6개의 에지 화소 가운데 3개의 에지 화소는 이웃 화소가 C₁ 대표 칼라를 가지고, 나머지 3개의 에지 화소는 C₃ 대표 칼라를 가진다. 그래서 그림 4(b)에서처럼 x축 C₂에 대해 y축 C₁, C₃에 각각 3개의 화소가 누적되게 된다.

그러나 칼라 인접 히스토그램을 제안한 32개의 대표 칼라를 가지는 양자화 방법을 이용하여 구성할 경우에 칼라 인접 히스토그램의 빈 크기는 32x

32로 아주 커지게 된다. 일반적으로 영상 검색 속도와 히스토그램을 저장하는데 필요한 저장 공간은 히스토그램의 빈 크기에 의존하므로, 제안한 칼라 인접 히스토그램을 이용하기 위해서는 효율적으로 빈 수를 줄이는 방법이 필요하다.

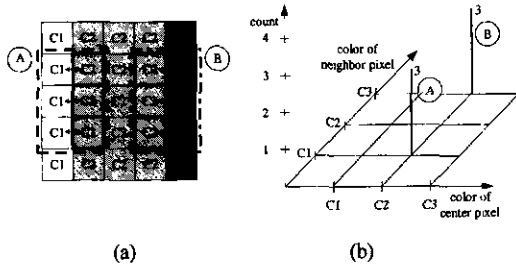


그림 4. 칼라 인접 히스토그램의 구성
(a) 입력 영상 (b) 칼라 인접 히스토그램

3.4 칼라 인접 히스토그램의 이진화 표현

칼라 인접 히스토그램은 칼라 에지 주위에 인접한 칼라들에 대한 히스토그램이므로 대부분의 영상의 경우에 대해 많은 비어있는 빈들을 가지게 된다. 그래서 칼라 인접 히스토그램의 각 빈 값들의 직접적인 자세한 비교보다는 주요한 칼라 에지를 구성하는 칼라 성분을 비교하는 것이 효율적이라 할 수 있다. 이렇게 주요한 칼라 에지를 이루는 칼라들만을 비교할 경우에 특징값의 저장 공간을 효율적으로 줄일 수 있을 뿐만 아니라, 칼라 영상의 스캔 과정 중에 생길 수 있는 영상의 잡음에 아주 강한 특성을 가지게 된다. 그림 5는 칼라 인접 히스토그램의 문턱치화 과정을 나타내는데, 각 히스토그램 빈의 값이 문턱값 보다 큰 경우에는 해당 빈이 주요한 에지를 구성하는 칼라 분포라고 생각하고, 그렇지 않은 경우에는 주요하지 않은 칼라 에지 성분으로 간주한다. 그림 5에서 사용된 문턱값은 15이다.

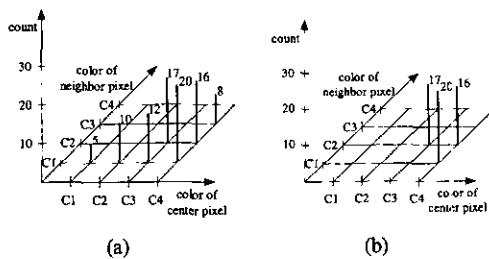


그림 5. 칼라 인접 히스토그램의 문턱치화
(a) 칼라 인접 히스토그램
(b) 주요 칼라 인접 히스토그램

주요한 칼라 에지 성분에 대해서는 이진수 '1'을 할당하고, 그렇지 않은 경우에는 이진수 '0'을 할당하면 그림 5(b)의 문턱치화된 칼라 인접 히스토그램은 아래와 같이 이진 행렬로 표현할 수 있다.

$$C = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (6)$$

여기서 행렬의 각 행은 이웃 화소의 대표 칼라를 나타내고, 각 열은 중심화소의 대표 칼라를 각각 나타낸다. 식 (6)의 이진 행렬은 3개의 칼라 쌍 (C_1, C_4), (C_2, C_3), (C_2, C_4)가 서로 인접하고 있음을 나타낸다. 이상과 같이 칼라 인접 히스토그램을 이진화하여 나타내어도 히스토그램 빈 수의 변화는 없게 되는데, 이진 행렬로 표현된 칼라 인접 히스토그램의 빈 수를 줄이기 위해 본 논문에서는 그림 6에서처럼 이진수로 표현된 이진 행렬의 각 행을 십진수로 변환하였다. 십진수로 변환하게 되면 제안한 비균일 양자화 방법의 양자화 레벨의 총 개수와 동일한 개수만큼의 십진수가 발생되므로, 32×32 개의 칼라 인접 히스토그램의 빈 수는 32개로 줄어들게 된다. 아래의 식은 식 (6)에 대한 십진수 변환 결과를 나타낸다.

$$C = [1 \ 3 \ 0 \ 0]^T \quad (7)$$

Binary stream	Decimal number
$\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$	1
$\begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix}$	3
$\begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$	0
$\begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$	0

그림 6. 십진수 변환

3.5 칼라 벡터각 히스토그램의 구성

중심 화소가 평탄 화소로 분류되면 중심 화소와 최대 벡터각을 이루게 되는 3×3 마스크내의 이웃 화소를 이용하여 칼라 벡터각 히스토그램을 구성하게 된다. 일반적으로 영상에는 에지 화소보다 평탄 화소가 더 많이 존재하게 되므로, 제안한 칼라 벡터각 히스토그램의 구성 시에 평탄 화소에 대한 대표 칼라의 수를 늘리면 검색 효율을 높일 수 있게 된다. 이를 위해 제안한 비균일 양자화 방법의 양자화 레벨 수를 증가시킬 경우, 시각적으로 유사한 칼라 성분을 가지는 칼라 에지들이 서로 다른 대표 칼라로 표현되어 오히려 칼라 인접 히스토그램의 성능

을 떨어뜨리게 된다. 그래서 본 논문에서는 양자화 레벨 개수를 증가시키지 않고, 평탄 화소에 대해서는 영상들 간의 변별력을 높이기 위해 중심 화소와 3×3 마스크내의 이웃 화소와의 최대 칼라 벡터각을 균일 양자화 하여 이용하였다. 3×3 마스크 내의 평탄 화소인 중심 화소에 대해서는 제안한 비균일 양자화 방법으로 생성된 32개의 대표 칼라로 표현하고, 평탄 화소의 변별력 증가를 위해 그림 7과 같이 중심 화소와 이웃 화소 사이의 최대 칼라 벡터각을 균일 양자화 하여 히스토그램을 구성하였다. 중심 화소가 평탄 화소이면, 3×3 마스크 내에서 이웃 화소와 중심 화소가 이루는 칼라 벡터각의 최대 값이 칼라 에지 검출에 사용되는 문턱값 보다 적게 된다. 그래서 그림 7에서처럼 칼라 에지 검출에 사용되는 문턱값의 1/2을 양자화 스텝으로 하여 최대 칼라 벡터각을 균일 양자화 하였다. 최대 칼라 벡터각이 균일 양자화 스텝보다 적은 경우에는 그림 7의 영역 A로 양자화되고, 그렇지 않은 경우에는 영역 B로 양자화 된다. 그림 8은 최종적인 칼라 벡터각 히스토그램의 구성을 나타낸다. 평탄 중심 화소가 32개의 대표 칼라로 표현되고, 각 중심 화소에 대해 최대 칼라 벡터각이 2개의 bin으로 표현되므로 총 64(32×2)개의 bin으로 칼라 벡터각 히스토그램은 구성되게 된다.

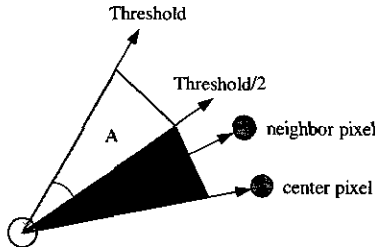


그림 7. 최대 칼라 벡터각의 균일 양자화

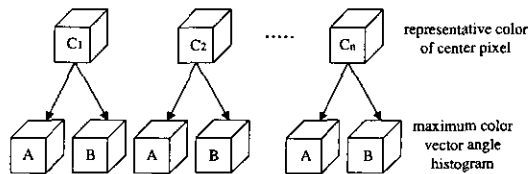


그림 8. 칼라 벡터각 히스토그램의 구성

VI. 목표 물체 위치 추정

Swain은 임의의 영상 내에서 찾고자하는 물체의

위치를 추정하기 위해 칼라 히스토그램 역투사 방법을 제안하였다.^{[1],[3]} 이 방법은 일정한 크기의 형태로 표현된 목표 물체 표면의 RGB 칼라 히스토그램과 임의의 영상 전체의 칼라 히스토그램을 비교하는 방법으로서 실시간 처리가 가능하면서 영상의 해상도와 히스토그램의 해상도에 민감하지 않은 방법이다. 칼라 히스토그램 역투사 방법은 임의의 영상 전체 화소가 일정한 크기의 형태로 표현된 목표 물체 내에 있을 가능성(likelihood)을 구하여 가장 높은 가능성 값을 영상 내에 존재하는 창 형태의 부영상(subimage) 위치를 찾아내는 방법이다. 이러한 목표 물체 위치를 수식적으로 표현하기에는 목표 영상이 다양한 크기를 가지므로, 일반적으로 일정한 크기의 창(window) 형태로 목표 물체를 표현하고 창의 중심 위치를 목표 물체의 위치로 가정한다.^[3] 찾고자하는 목표 물체 영상(target object image)을 Q , 목표 영상을 찾고자하는 대상 영상을 I 라 할 때 ($Q \subseteq I$), 먼저 아래와 같이 칼라 c 에 대한 히스토그램 비(ratio histogram)를 구한다.

$$\pi_{c,h}(I/Q) = \min \left\{ \frac{H_c(Q)}{H_c(I)}, 1 \right\} \quad (8)$$

영상 내에 존재하는 모든 부영역에 대해 아래와 같이 각 부영역에 대한 가능성을 구한다.

$$\prod_p = \sum_{q \in I_1} \pi_{p,q,h}(I/Q) \quad (9)$$

최종적인 목표 물체의 위치는 아래 식에 의해 결정된다.

$$\max \prod_p(I/Q) \quad (10)$$

목표 물체 영상 Q 에서 칼라 c 가 모든 다른 칼라 성분 보다 우세할 경우, 즉 $\pi_{c,h}(I/Q)$ 가 최대값일 경우에 칼라 히스토그램 역투사 방법은 부영역 I_1 가 많은 칼라 c 화소를 가지고 있으면, 목표 물체가 존재하지 않더라도 가능성 값이 증가하는 문제점을 가지고 있다.^[3] 그림 9에서처럼 Q 영상이 80개의 검정 화소와 20개의 흰색 화소를 가지고 있고, I 영상이 200개의 검정 화소와 200개의 흰색 화소를 가지고 있을 경우에는 $\pi_{b,h} = 0.4, \pi_{w,h} = 0.1$ 이 되어서 잘못된 목표 물체를 추출하게 된다.

그러나 제안한 칼라 특징 정보를 이용할 경우에는 칼라 인접 히스토그램을 통하여 부영상에서의 에지 성분에 대한 칼라를 비교하므로 앞서 언급한

칼라 히스토그램 역투사 방법의 문제점은 존재하지 않는다.

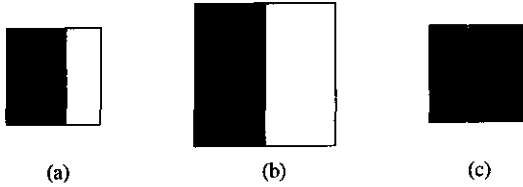


그림 9. 히스토그램 역투사의 잘못된 추출 예 (a) 목표 물체 질의 영상 Q (b) 영상 I (c) 추출된 목표 물체 영상

V. 유사도 측정

5.1 칼라 인접 히스토그램의 유사도 측정

칼라 인접 히스토그램의 각 히스토그램 빈은 칼라 에지를 구성하는 칼라 쌍(color pair)에 대한 정보를 이진열(binary array)의 형태로 표현하므로 기존의 히스토그램 인터섹션(histogram intersection)의 방법으로 유사도를 측정하기는 어렵다. 그래서 본 논문에서는 임의의 영상 I 에 대한 k 번째 대표 칼라에 대한 이진열을 $B_{C_k}(I) = (b_1^k \ b_2^k \ \dots \ b_N^k)$ 라 할 때, 질의 영상 I 와 데이터 베이스 영상 J 간의 대표 칼라 C_k 에 대한 인터섹션을 아래와 같이 구하였다.

$$S_k(I, J) = \frac{NC_k^*(I, J)}{N_k^*(I) + N_k^*(J) - NC_k^*(I, J)} \quad (11)$$

여기서 $N_k^*(I)$ 와 $N_k^*(J)$ 는 각각 I 영상과 J 영상의 k 번째 대표 칼라에 대한 이진열에서 1의 개수를 나타내며, $NC_k^*(I, J)$ 는 두 영상 I, J 의 k 번째 대표 칼라에 대한 이진열에서 동일한 위치에 존재하는 1의 개수를 나타낸다. 에지 주위의 칼라의 분포가 동일할수록 유사도값이 증가함을 식(11)에서 확인할 수 있다. 예를 들어 $B_{C_k}(I) = [0 \ 0 \ 1 \ 0 \ 1]$, $B_{C_k}(J) = [0 \ 0 \ 1 \ 1 \ 0]$ 라 하면, $NC_k^*(I, J) = 1$, $N_k^*(I) = 2$, 그리고 $N_k^*(J) = 2$ 가 되므로 인터섹션의 결과값은

$$S_k(I, J) = \frac{1}{2+2-1} \approx 0.33 \quad (12)$$

가 된다. N 개의 대표 칼라로 양자화 하였을 경우에 대한 전체 인터섹션 결과값은 아래와 같다.

$$S_{pair}(I, J) = \frac{1}{N} \sum_{k=1}^N S_k(I, J) \quad (13)$$

5.2 칼라 벡터각 히스토그램의 유사도 측정

칼라 벡터각 히스토그램의 유사도 측정에는 기존의 히스토그램 인터섹션(histogram intersection) 방법을 이용하였다.^[1] $H_{R(k)}$ 와 $H_{I(k)}$ 를 각각 데이터 베이스 영상과 질의 영상의 칼라 벡터각 히스토그램이라 할 때 질의 영상과 데이터 베이스 영상의 거리 측정은 아래와 같다.

$$S_{vec}(I, J) = 1 - \frac{1}{2T} \sum_{k=1}^T |H_{R(k)} - H_{I(k)}| \quad (14)$$

여기서 T 는 전체 히스토그램 빈 값(histogram bin value)의 합으로써 본 논문에서는 다양한 크기를 가지는 모든 영상에 대해서 검색 결과가 각 영상의 크기에 영향을 받지 않게 하기 위해서 각 영상에서 모든 히스토그램 빈 값의 합으로 각 히스토그램 빈을 정규화 하여 영상의 크기에 관계없이 각 히스토그램 빈값의 합이 T 가 되도록 하였으며, 실험에서는 T 값을 255로 하였다. 아래는 각 히스토그램 빈 값의 정규화 과정에 대한 의사 코드(pseudo code)를 나타낸다.

```
for k=0; k<= total_number_of_histogram_bin;k++
    total += H(k);
```

```
for k=0; k<= total_number_of_histogram_bin;k++
    H(k) = T × ⌊ H(k)/total ⌋
```

5.3 전체 유사도 측정

칼라 인접 히스토그램과 칼라 벡터각 히스토그램이 결합된 증대 히스토그램(augmented histogram)의 전체 유사도 측정 방법은 아래와 같다.

$$S(I, J) = \alpha S_{pair}(I, J) + \beta S_{vec}(I, J) \quad (15)$$

여기서 α 와 β 는 가중치이며, 본 논문에서는 0.5를 사용하였다.

VI. 실험 및 결과

제안한 색 정보 추출 알고리즘의 성능 평가를 위한 실험을 위해 본 논문에서는 자연, 인물, 동식물 등 다양한 종류의 5466 개의 영상으로 데이터 베이스를 구성하였다. 먼저 질의 영상을 결정하고, 각각의 질의 영상에 대해 관련성(relevant)이 있는 영상들의 집합인 ground-truth 영상을 결정하여야 하는데,

본 실험에서는 총 50개의 질의 영상을 결정하고, 각 질의 영상에 대한 ground-truth 영상들을 결정하였다.

성능 평가를 위해 사용한 정확도 측정 방법으로는 ANMRR, precision, recall을 이용하였다.^[9,12,19] 성능 평가값인 ANMRR를 구하기 위한 과정은 아래와 같으며, 평균 rank값을 나타내는 ANMRR은 그 값이 낮을 수록 검색 결과가 우수함을 나타내고, precision과 recall은 그 값이 높을 수록 검색 결과가 우수함을 나타낸다.

6.1 히스토그램 빈 표현

보다 객관적인 성능 평가를 위하여 각 영상의 특징 벡터를 저장하는데 필요한 저장 공간을 비교하였다. 제안한 방법의 경우에는 칼라 인접 히스토그램의 경우 32개의 각 대표 칼라에 대해 32 비트가 필요하고, 칼라 벡터각 히스토그램의 경우에 64개의 빈 수를 가지고, 최대 히스토그램 빈값이 255이므로 총 1536(32×32+64×8) 비트가 각 영상마다 필요하다. 기존의 125 빈과 252 빈 칼라 히스토그램의 경우도 최대 히스토그램의 빈값을 255로 하여 각 빈을 8비트로 표현하였다. 그래서 125 빈과 252 빈도 각각 1000 비트와 2016비트가 필요하게 된다.

6.2 기존 방법들과의 성능 비교

표 2는 기존의 방법들과 제안한 방법의 전체 50개의 질의 영상 집합에 대한 검색 결과를 나타낸다. 표에서도 알 수 있듯이 제안한 방법이 비교적 적은 히스토그램 빈 수와 저장 용량에서도 아주 우수한 검색 효율을 가짐을 확인할 수 있다. Graph intersection 방법^[8]의 경우 영상마다 저장 용량이 가변되므로 정확하게 저장 용량을 비교할 수 없고, 전체 빈 수도 다른 대표 칼라와의 상관 관계를 고려하므로 실제적으로는 64개를 넘어서게 된다.

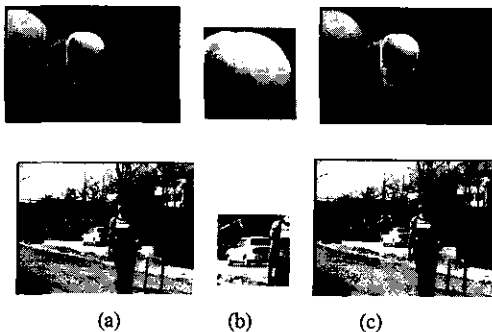


그림 11. 목표 물체 위치 추정 (a) 칼라 히스토그램 결과 (b) 목표 물체 질의 영상 (c) 제안한 방법

그림 11은 칼라 히스토그램 역투사 방법의 성능 비교를 나타낸다. 기존의 칼라 히스토그램 기반의 Swain 방법은 정확하게 목표 물체 위치를 찾아내지 못하지만, 제안한 칼라 특징 정보를 이용한 방법의 경우에는 정확하게 목표 물체 위치를 찾아내는 것을 확인할 수 있다.

표 2. 기존의 영상 검색 방법들과의 결과 비교

	ANMRR	Average Recall	Average Precision
Color Histogram 64 bins (512 bits)	0.22	0.20	0.80
Color Histogram 125 bins(1000 bits)	0.19	0.22	0.85
Color Histogram 252 bins (2016 bits)	0.15	0.22	0.87
Joint Graph Intersection 125 bins	0.14	0.22	0.89
Proposed 96 bins (1536 bits)	0.11	0.24	0.91

VII. 결론

본 논문에서는 칼라 벡터각을 이용한 영상 검색 방법과 물체 위치 추정 방법을 제안하였다. 기존의 칼라 히스토그램 방법들이 가지는 공간 정보의 부족 문제를 보완하기 위해 제안한 칼라 인접 히스토그램은 기존의 과도한 연산량을 필요로 하는 방법들과 달리 큰 연산량의 증가 없이 카메라의 이동 및 줌 등에 의해 발생하는 생김새 및 모양의 변화에 큰 영향을 받지 않고 안정된 검색 결과를 나타내는 것을 확인할 수 있었다. 그리고 입력 영상에서 예지가 존재하지 않은 평탄 영역들에 대해서도 검색 효율을 높이기 위해 중심 화소와 이웃 화소와의 색차를 나타내는 칼라 벡터각을 이용하므로 보다 효율적인 검색 결과를 얻을 수 있었다. 또 칼라 인접 히스토그램의 표현에 있어서 각 대표 칼라마다 칼라 인접의 위치 정보를 이진수로 표현하여 특징 벡터 크기의 아주 큰 증가 없이 효율적으로 히스토그램을 표현할 수 있는 방법을 제안하였다. 그리고 히스토그램 역투사 방법을 통한 물체 위치 추정에 있어서도 제안한 칼라 특징 벡터를 이용할 경우에 칼라 예지 정보를 이용하므로 기존의 칼라 히스토그램을 이용한 방법보다 영상 내에 미세하게 존재하는 칼라 정보의 영향을 아주 많이 줄일 수 있었다.

2001년 2월: 경북대학교 대학원 전자공학과 졸업
(공학석사)

2001년 3월~현재: 경북대학교 대학원 전자공학과
박사과정

<주관심 분야> 칼라 영상 처리, 칼라 하프토닝

남 재 열(Jae-Yeal Nam) 정회원
제 22 권 제 4 호 참조

하 영 호(Yeong-Ho Ha) 정회원
제 25 권 제 8B 호 참조