

論文 99-36S-9-9

# 컬러 히스토그램과 컬러 텍스처를 이용한 내용기반 영상 검색 기법

(Content-based Image Retrieving Using Color Histogram and Color Texture)

李亨九\*, 尹逸東\*

(Hyung Goo Lee and Il Dong Yun)

## 요약

본 논문은 컬러 히스토그램과 “컬러 텍스처”를 이용하는 새로운 내용기반 영상 검색 기법을 제안한다. 제안하는 기법은 영상의 컬러 히스토그램을 k-means 군집화하여 얻은 컬러 벡터로 히스토그램을 대표하고, 각 대표 컬러 벡터를 중심으로 화소 색상과의 거리를 이용해 컬러 텍스처를 만든다. 그러므로, 컬러 텍스처란 영상의 컬러 히스토그램에 의해 두드러지는 텍스처 성분을 의미하며 본 논문에서는 컬러 텍스처를 Gaussian Markov Random Field (GMRF) 모델로 해석한다. 제안하는 알고리듬은 영역화와 같은 기하학적 정보를 추출하는 과정이 없으므로 고속의 검색에 적합하며, 기존의 컬러 히스토그램만을 이용한 기법이나 영상의 밝기 성분에서 나타나는 텍스처를 이용한 방법에 비해 효과적인 검색 결과를 나타낸다.

## Abstract

In this paper, a color image retrieval algorithm is proposed based on color histogram and color texture. The representative color vectors of a color image are made from k-means clustering of its color histogram, and color texture is generated by centering around the color of pixels with its color vector. Thus the color texture means texture properties emphasized by its color histogram, and it is analyzed by Gaussian Markov Random Field (GMRF) model. The proposed algorithm can work efficiently because it does not require any low level image processing such as segmentation or edge detection, so it outperforms the traditional algorithms which use color histogram only or texture properties come from image intensity.

## I. 서론

지난 몇 년간 정보 시스템 분야에서 가장 중요한 변화는 처리하고자 하는 정보가 과거와는 달리, 단순한 텍스트 정보가 아닌 영상, 비디오, 오디오 등이 결합된

디지털 멀티미디어 정보라는 것이다. 최근 들어 이런 멀티미디어 정보는 매일 엄청난 양으로 쏟아져 나오고 있으며 데이터의 양 또한 과거의 텍스트 정보와 비교할 수 없을 정도로 방대하다. 따라서 멀티미디어를 효율적으로 표현하고, 저장하고 원하는 정보를 쉽게 검색하는 방법의 필요성이 점점 증대하고 있는 실정이며, 이에 관련된 여러 다양한 방법들이 활발히 연구되고 있다. 특히, 영상 데이터 베이스나 여러 순차 영상들로부터 원하는 것을 찾고자 하는 영상 검색 방법은 영상 처리의 새로운 분야로 각광을 받고 있다<sup>[1]</sup>. 영상 검색 기법(image retrieving technique)은 전자도서관, on-line shopping, trademark searching 등으로 그 응용 범위가

\* 正會員, 韓國外國語大學校 制御計測工學科  
(Dept. of Control and Instrumentation Eng., Hankuk Univ. of F.S.)

※ 이 논문은 1997년 학술진흥재단의 공모과제 연구비에 의하여 연구되었음.

接受日字 : 1999年1月20日, 수정완료일 : 1999年7月23日

점점 확대되고 있으며, 특히 앞으로 다가올 정보화 시내에 필수적인 기술이라 할 수 있다.

이러한 영상 검색 기법은 크게 텍스트 주석에 기반한 검색 방법(index-based image retrieving)과 내용에 기반한 검색 방법(content-based image retrieving)으로 나눌 수 있다. 멀티미디어는 그 정보의 속성상 텍스트 주석으로 표현하기에는 주관적이고 모호하다. 따라서 보다 객관적 유사성인 영상의 통계적, 기하학적 특징을 사용하는 내용 기반 검색 기법이 주로 연구되고 있다<sup>[2-4]</sup>. 실제로 몇몇의 연구자들은 자신의 알고리듬을 WWW 상에 구현하여 시연을 하고 있는 실정이다<sup>[5-7]</sup>. 이러한 기법중의 일부는 코닥이나 오라클과 같은 영상이나 데이터 베이스 관련 회사에 그 검색 엔진을 이식하는 단계에 이르고 있다<sup>[7]</sup>. 그러나 불행히도 일반적인 영상 데이터 베이스에서 자동으로 영상의 종류를 분류(classify)하거나 영상 내의 객체(thing, object)를 인식하는 표준적인 기법은 제안되지 않았다. 사용자는 특정 객체를 포함하고 있는 영상을 찾는데 반해, 아직까지 내용기반 검색 기법들은 영상의 색상이나 텍스트, 또는 형태 요소 혹은 이들의 공간적 조합을 영상의 특징으로 이용할 때 때문이다.

최근 영상의 내용을 객체로 판단될 수 있는 후보 영역별 특징으로 표현하기 위한 다양한 기법이 연구되고 있다<sup>[3-4]</sup>. 그러나 이러한 접근 방식으로 일반적인 영상 데이터 베이스를 검색한 결과가 색상 특징 한 가지 기준에 의한 검색 보다 부정확한 결과를 나타내는 등 복잡한 부가 계산을 감수하고 후보 영역별 특징을 구했지만 이것이 검색 성능의 향상을 가져왔다고 결론 내릴 수 있는 일반적이고 정량적인 실험은 아직 이루어지지 않고 있다.

본 논문에서는 “컬러 텍스트”라는 일반적인 컬러 영상의 검색 기준을 제시한다. 컬러 텍스트는 영상의 대표적인 색상 군집에 의해 강조되는 텍스트로, 영상 내 객체의 표면이 나타내는 텍스트를 의미하지는 않으나 영상의 낮은 차원의 특성에 비해 보다 의미 있는 검색 기준이 된다. 특히 영역화 등의 기하학적 정보를 추출하는 과정이 없고 각 특성을 조합하기 위한 부가 연산이 없으므로 특징 추출을 위한 연산 부담이 적어 고속의 정합에 적합하다는 장점이 있다.

본 논문의 구성은 다음과 같다. 다음 2장에서는 컬러 히스토그램과 텍스트를 이용한 기준의 영상 검색 기법의 연구 동향을 살피고, 3장에서는 제안하는 내용기반

영상 검색 기법과 연구 방향을 언급하고, 4장에서는 연구 결과를 제시하며, 마지막 5장에서 결론과 함께 향후 연구 방향에 대해 기술한다.

## II. 기준의 연구

QBIC<sup>[5]</sup>, VisualSEEK<sup>[6]</sup>, Virage<sup>[7]</sup> 등의 대표적인 내용기반 검색 시스템들은 자동 영역화 등의 기능을 제공하지만 대체로 영상의 낮은 차원의 특징을 검색에 이용하고 있으며, 최근 들어 객체 차원의 특징을 이용한 검색에 대한 연구가 진행되고 있다. 고전적 객체 인식 기법은 기계의 부품과 같은 기하학적 객체나 영역화에 의한 인식에 기반 한다. 그러나 일반 영상에서는 이와 같은 제한이 존재하지 않으므로 고전적 객체 인식 기법을 일반 영상 검색에 적용하는데는 한계가 있다. 따라서 일반 영상에서 객체 차원의 특징을 구하고자 하는 최근의 연구에서는, 영상의 기하학적 특징 대신, 모든 컬러 영상에서 나타나는 특성인 색상과 텍스트를 이용해 객체로 판정될 수 있는 후보 영역을 기술한다. 그러나 이러한 접근 방법은 영역화를 수반할 뿐 아니라 텍스트 특징과 컬러 특징을 조합하기 위한 부가 연산이 따르는 단점이 있다.

본 장에서는 기준의 내용기반 검색 기법에서 색상, 텍스트 특성이 어떻게 사용되는지 살펴 보고, 색상과 텍스트를 사용하여 객체 차원의 특성을 추출하는 기법에 대해 기술한다.

### 1. 컬러 히스토그램 기반의 검색 기법

본 절에서는 내용 기반의 영상 검색 기법 중 가장 널리 이용되는 기법인 컬러 히스토그램의 분석에 대해 기술한다. 컬러 히스토그램을 이용한 방법이라는 것은 컬러 이미지의 히스토그램을 만든 후 각각의 히스토그램이 얼마나 유사한가를 고려하는 방법으로 이것은 영상의 회전(rotation)이나 이동(translation), 그리고 어느 정도의 가려짐(occlusion)에도 상당히 안정된 성능을 나타낸다. 그러나 밝기와 색의 변화 등 히스토그램 자체에 대한 변화에 상당히 민감하다는 단점이 있다. 이러한 방법을 사용하는데 있어서의 관건은 두 개의 히스토그램이 얼마나 유사한가를 판단하기 위해서 히스토그램의 어떤 특징을 사용할 것인가와 이런 특징들을 어떤 척도(measure)로 비교할 것인가라고 할 수 있다.

히스토그램을 이용한 검색으로는 Swain과 Ballard가

세안한 “Histogram Intersection”이 널리 알려져 있다<sup>[8]</sup>. 이 방법은 컬러 히스토그램의 특징으로 식 (1)과 같이 히스토그램의 각 bin에 존재하는 화소의 수를 사용한다. 이 방법은 간단하고 빠르므로 실시간 구현에 적합하며, 회선, 이동, 가려짐 등의 변화에 강인한 장점이 있어, 컬러 히스토그램을 이용한 검색 방법에 관련된 논문들은 주로 이 방법에 기반하여 보다 나은 성능을 낼 수 있도록 여러 가지를 추가하거나 수정하는 방법을 택하고 있다<sup>[2]</sup>. 그러나 일반적으로 영상의 컬러 히스토그램 분포는 그림 1과 같이 특정 bin에 몰려있는 성질을 나타내므로, RGB축을 균일하게 나누는 것은 컬러 히스토그램의 통계적 특징을 잘 이용하지 못하는 것이라고 할 수 있다.

$$H_N(D, Q) = \frac{\sum_{i=0}^n \min(D_i, Q_i)}{\sum_{i=0}^n Q_i} \quad (1)$$

$D_i$  : 데이터 베이스 영상의 i번째 bin의 화소수

$Q_i$  : 입력 영상의 i번째 bin의 화소수

$D_i$  : # of the pixels in the i-th bin of a database image

$D_i$  : # of the pixels in the i-th bin of a query image

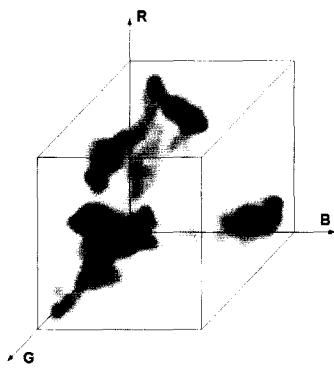


그림 1. 컬러 영상에 대한 전형적인 히스토그램 분포  
Fig. 1. A typical color histogram distribution of an image.

컬러 히스토그램을 이용하는 경우 컬러 공간상의 정확한 거리 측정을 위해 RGB 모델 이외에도 HSV, YIQ, YUV 등의 모델이 사용되고 있다. 특히 HSV 모델은 인간의 시각 능력과 유사한 특성을 갖는 것으로 알려졌는데, 각각 색상(hue), 명도(value), 채도(saturation)의 세가지 요소로 구성되어 유사한 색상 별로 쉽게 구

집화가 가능하다는 장점이 있다. 또한 사람의 인지 범위를 벗어나는 컬러 공간 상의 과잉 정보(redundancy)를 제거하고 군집화 하는 방법도 연구되고 있다<sup>[9~10]</sup>.

## 2. 텍스처 정보를 이용한 검색 기법

텍스처는 영상에서 동일한 패턴을 가지는 영역의 특성으로, 객체의 표면이 가지는 특성을 표현하고 있기 때문에 영상에게는 매우 중요한 특징이 된다. 특히 항공 사진이나 위성 사진을 분석하는 경우에 텍스처 정보가 많이 사용되어 있는데, 텍스처를 통해서 사진의 각 영역이 물의 표면, 경작지, 산 또는 도시 중에 어디에 속하는지를 분석할 수 있다<sup>[11]</sup>.

텍스처 정보를 이용한 내용 기반 검색기법에 대한 연구에는 텍스처의 gradient 히스토그램을 비교하여 텍스처의 유사도를 측정하는 방법<sup>[12]</sup>, 텍스처를 Markov나 Fractal 등으로 모델링하고 모델의 계수를 추출하여 텍스처 특성을 정의하는 방법<sup>[13]</sup>, Gabor filter나 Wavelet 변환을 통해서 텍스처 정보를 주파수 대역별로 분석하는 방법 등이 사용되고 있다<sup>[3, 14]</sup>.

## 3. 컬러와 텍스처의 조합을 이용하는 기법

색상과 텍스처 각각으로는 영상 내의 관심 있는 영역에 대한 묘사가 어렵다. 예를 들어 넓은 회색 하늘을 포함하고 있는 영상을 검색하고자 한다면, 색상 특징만 사용할 경우 검색의 범위를 효과적으로 좁히지 못한다. 영상 데이터 베이스에는 도심의 빌딩 숲, 거리 풍경, 임의의 지형 등 사용자의 관심 밖인 회색 영역이 많이 있으므로 색상만으로 원하는 영상을 찾기 어렵다. 그러나 사용자가 찾고자 하는 영역이 부드럽고 텍스처가 약한 특성을 가진다는 기술을 덜붙인다면 원하는 내용을 포함한 영상에 더욱 근접한 결과를 얻을 수 있다.

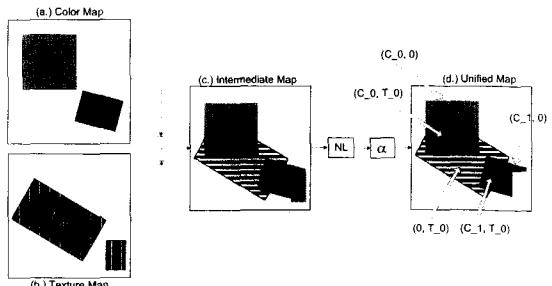


그림 2. 특징 추출을 위한 색상 map과 텍스처 map의 조합  
Fig. 2. Combination of(a.) color and (b.) texture maps at time of feature extraction.

이와 같이 컬러와 텍스처의 조합을 이용하면 보다 효과적인 검색이 가능하여 최근에는 이 두 가지를 동시에 이용하는 검색 기법이 연구되고 있다<sup>[3, 4, 14]</sup>.

Smith와 Chang<sup>[3]</sup>은 그림 2와 같이 영상의 color map과 texture map을 만들고, 두 map의 조합을 영상의 특징으로 사용하며, Belongie<sup>[4]</sup> 등은 영상의 텍스처 공간과 컬러 공간 상의 밀접성을 이용해 화소 정보를 blobworld라는 영상의 작은 영역으로 변환하여 영상의 특징을 나타낸다. Blobworld 표현은 컬러와 텍스처 특징 벡터를 Expectation Maximization (EM) 알고리듬을 이용해 영역화(segmentation)하여 얻어진다.

위와 같은 기법들은 영상의 내용을 단순히 색상이나 텍스처, 또는 형태 요소만으로 나타내지 않고 이를 이용해 객체로 판단될 수 있는 후보 영역별 특징으로 표현한다는 점에서 주목 할만 하다. 그러나 색상과 텍스처를 이용하는 기존의 기법들은, 색상과 텍스처 특성을 따로 추출하고 화소에 대한 특성인 색상과 공간적으로 이웃한 점들이 나타내는 특성인 텍스처를 조합하기 위한 많은 부가 연산이 따른다는 한계를 가지고 있다. 게다가 각 특성의 추출, 혹은 특성의 조합 과정에서도 영역화가 수반되어 연산 부담이 한층 높아진다. 그러나 이러한 기법은 객체 차원의 특징을 표현하기 위해 계산의 복잡도가 상당히 증가한데 비해, 일반적인 영상 데이터 베이스에 대한 검색에서 Swain의 Histogram intersection 보다 효율이 떨어지는 결과를 내는 등 후보 영역별 특징 표현이 검색 성능의 향상을 가져온다는 결론을 뒷받침할 일반적이고 정량적인 실험은 이루어지지 않고 있다<sup>[4]</sup>. 따라서 본 논문에서는 색상 군집에 의한 텍스처-컬러 텍스처-를 정의하여 영역화와 부가 연산이 생략되어 고속의 정합에 적합한 알고리듬을 제안한다.

### III. 제안하는 내용기반 검색 기법

영상의 특징은 그 추출 방법에 따라 검색의 성능을 한정 지으며, 추출 방법의 복잡도는 향후 시스템의 디자인에 절대적인 영향을 끼치는 요인으로 작용하므로 본 논문은 검색 효율과 계산량을 만족시키는 새로운 영상의 특징을 제안한다. 제안하는 기법은 컬러 특징과 텍스처 특징을 이용한다. 영상의 컬러 특징으로 RGB 히스토그램을 군집화하여 얻은 대표 벡터를 사용하며,

이 기법은 기존의 Histogram intersection과 같은 방법에 비해 컬러 히스토그램의 특징을 효율적으로 사용한다는 장점이 있다. 텍스처 특징으로는 영상을 대표하는 색상 군집의 텍스처, 즉 제안하는 컬러 텍스처를 이용하는데, 이는 텍스처 특징을 구하기 위한 영역화나 텍스처 특징과 색상 정보와 조합하는데 따르는 부가 연산이 필요치 않다는 장점이 있다.

본 장에서는 영상의 색상 분포를 군집화하여 대표 벡터를 구하는 방법, 컬러 텍스처를 구하는 사영 함수, 텍스처 특징 기술 방법, 유사도 계산을 위한 척도에 대해 각각 설명하고 마지막으로 제안하는 검색 기법의 전체적 구성을 제시한다.

#### 1. 컬러 히스토그램에서 대표 컬러 추출

앞에서 설명했듯이 일반적으로 영상의 컬러 히스토그램 분포는 특정 bin에 몰려있는 성질을 나타내므로, RGB축을 균일하게 나누는 것은 히스토그램의 통계적 특징을 잘 이용하지 못하는 것이라고 할 수 있다. 따라서 본 논문에서는 히스토그램의 통계적 특성을 보다 효율적으로 나타내기 위해, 컬러 히스토그램 공간에 k-means 군집화 기법을 적용하여 영상을 대표하는 컬러를 추출하고, 이 대표 컬러를 입력 영상의 컬러 히스토그램 특성으로 사용한다.



(227, 221, 180) (185, 60, 37)  
 (118, 33, 15) (73, 85, 87)  
 (176, 149, 136) (25, 33, 67)  
 (218, 203, 100) (82, 66, 48)  
 (58, 43, 32) (56, 55, 55)  
 (135, 110, 93) (25, 17, 7)  
 (65, 73, 76) (212, 132, 74)  
 (41, 34, 24) (198, 174, 159)  
 (9, 13, 35) (153, 134, 111)  
 (59, 63, 71) (114, 85, 65)

그림 3. Fruit과 이 영상을 대표하는 20개의 컬러 벡터  
 Fig. 3. The 20 representative color vectors of a fruit image.

k-means 알고리듬은  $k$  개의 초기 군집 중심에 대해 군집화를 수행하고 각 군집 영역에 속하는 모든 점으로부터 군집 중심까지의 거리의 제곱의 합을 최소로 하는 새로운 군집 중심을 정의하는 것을 일정한 군집 중심을 얻을 때까지 반복하는 알고리듬이다. 이렇게 구해진 군집 중심이 영상의 히스토그램 분포를 대표하는 컬러 벡터가 되며, 이는 일정한 크기로 나누어진 bin의 화소수로 유사도를 계산하는 Histogram intersection에 비해 간결하고 효과적인 영상 기술을 가능하게 한다. 본 논문에서는 영상의 대표 벡터의 수를 20, 15, 10, 5 개로 바꾸어가며 세 가지 데이터 베이스에서의 검색을 수행하여 영상의 표현하는 적절한 벡터 수를 조사했다.

## 2. 제안하는 컬러 텍스처

본 논문에서 제안하는 컬러 텍스처란, 영상의 밝기 성분에서 나타나는 텍스처와 달리, 영상의 각 화소 값과 영상을 대표하는 컬러 사이의 벡터 거리를 0부터 255 범위의 밝기로 사영하여 얻어진 gray 영상을 말하며, 거리 함수는 식 (2)와 같다.

$$dist = \sqrt{(R_n - r)^2 + (G_n - g)^2 + (B_n - b)^2} \quad (2)$$

$R, G, B$  : 대표 벡터  $r, g, b$  : 화소값

사영 함수는 255에서 0까지의 밝기 범위와 클러스터의 중심에서 경계까지의 거리 범위를 선형 사영하는 함수로 정의했다. 따라서, 화소의  $r, g, b$ 가 대표 컬러와 같을 경우의 벡터 거리는 0이며, 이 화소는 컬러 텍스처에서 밝기 255인 화소가 된다. 그리고 벡터 거리가 커져 클러스터의 경계에 이르면 사영 함수에 의해 그 화소는 컬러 텍스처에서 가장 어두운 밝기 0인 화소에 대응된다.

그림 4(a), (b)는 그림 3의 Fruit 영상의 대표 컬러 벡터 중 (25, 33, 67)과 (59, 63, 71)이 대표하는 색상 군집의 특성을 나타내는 그림으로, 군집 중심에서 가장 거리가 먼 화소까지의 거리를 40 단계로 나누고 각 단계에 속하는 화소의 수를 표시한 히스토그램이다. 각 대표 컬러는 영상의 컬러 히스토그램을 k-means 군집화하여 얻은 것이므로, 그림 4(a), (b)와 같이 대표 컬러와 거리가 멀어질수록 화소의 수가 적어지는 히스토그램을 얻게 되며, 이는 그림 4(c)와 같이 개념적으로 나타낼 수 있다. 이러한 분포를 가지는 군집에서 최대 거리를 군집의 경계로 삼고 위의 사영 함수를 적용하게 되면, 그림 5처럼 한가지 색상에 대한 컬러 텍스처

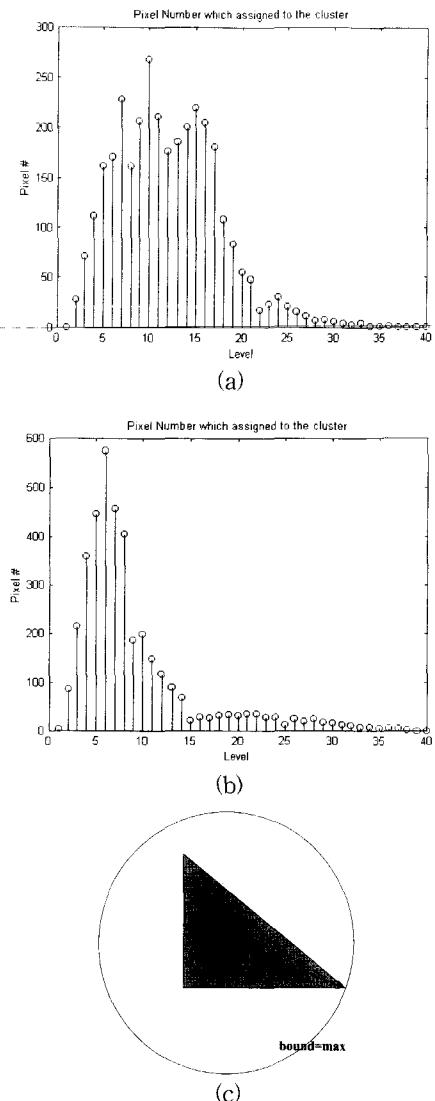


그림 4. k-means 알고리듬으로 구한 Fruit 영상에 대한 군집의 특성

- (a) 군집 중심 (25, 33, 67)부터의 거리에 따른 화소 분포, 최대거리를 40단계로 나누었다.
- (b) 군집 중심 (59, 63, 71) 부터의 거리에 따른 화소 분포.
- (c) 개념적 표현 : 군집 중심으로부터 멀어질수록 화소의 수가 적어진다.

Fig. 4. Cluster characteristics of Fruit image using k-means algorithm.

- (a) the count of pixel distribution from cluster center (25, 33, 67) to bound, the maximum distance quantized in to level
- (b) the count of pixel distribution from cluster center (59, 63, 71) to bound
- (c) Conceptual expression : the count of pixel decreases as far as the distance from the center.

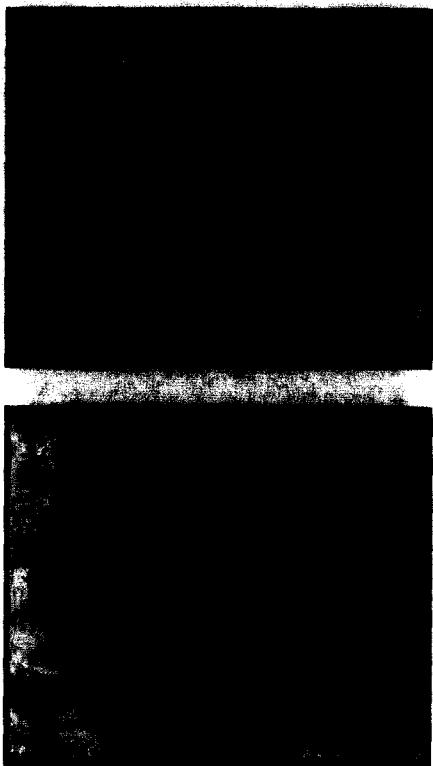
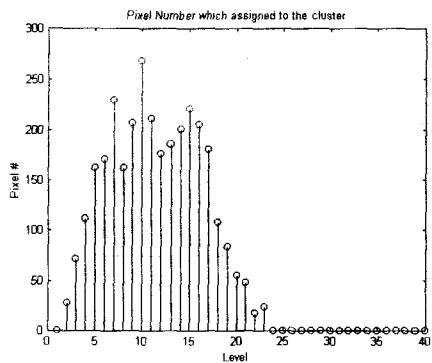


그림 5. Fruit 영상의 컬러 텍스처, 최대거리를 군집 경계로 사용한 경우, 컬러 (25, 33, 67)과 (59, 63, 71)에 대해

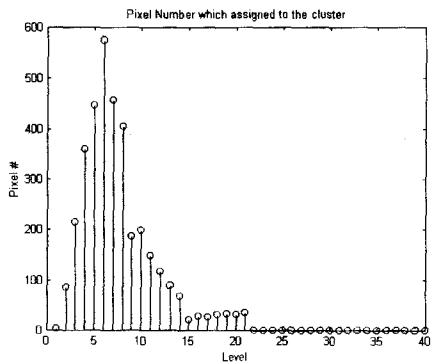
Fig. 5. Color textures of the fruit image, use maximum distance as cluster bound, for color vector (25, 33, 67) & (59, 63, 71).

에 직류 값이 다른 몇 개의 텍스처 영역이 나타나게 된다. 높은 직류 값을 가지는 텍스처 영역은 대표 컬러 벡터에 의해 강조되는 텍스처 영역이며, 낮은 직류 값의 영역은 군집 중심과 거리가 먼 화소들이 나타내는 텍스처 영역이다. 제안하는 기법은 컬러 텍스처를 평균 0인 프로세스로 가정하고 텍스처 특징을 구하므로 특정 직류 성분을 가지는 텍스처 영역을 제거해도 해석에 영향을 주지 않는다. 따라서 본 논문에서는 그림6과 같이 군집 중심과 거리가 먼 10%의 화소를 제거하여 군집 경계를 정하므로, 낮은 직류 값을 가지는 영역을 제거하여 대표 벡터에 충실한 컬러 텍스처를 얻는다.

그림 5와 그림 7은 Fruit 영상의 대표 벡터 20개 중 (25, 33, 67)과 (59, 63, 71)에 대한 컬러 텍스처이다. 그림5는 최대 거리를 경계로 이용하여 컬러 텍스처를 구한 경우로, 컬러 텍스처에 몇 개의 서로 다른 직류 값을 가지는 영역이 나타남을 볼 수 있다. 그림7의 경우는 군집 중심으로부터 거리가 먼 10%의 화소를 제거함



(a)



(b)

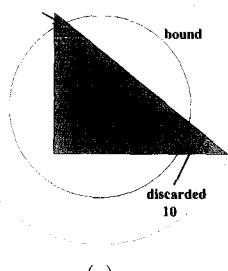


그림 6. 군집 중심으로부터 거리가 먼 10%의 화소를 제거하여 군집의 경계를 정한다.

(a) 그림 4(a)에서 경계부분 10%를 제거한 화소 분포

(b) 그림 4(b)에서 경계부분 10%를 제거한 화소 분포

(c) 경계 화소 제거의 개념도

Fig. 6. Determine cluster bound by discarding 10% pixels which are farthest from center.

(a) The pixel distribution which discards 10% boundary points of Fig. 4(a).

(b) The pixel distribution which discards 10% boundary points of Fig. 4(b).

(c) Conceptual representation of boundary points discarding.



그림 7. Fruit 영상의 컬러 텍스처, 거리가 먼 10%의 화소를 제거하여 결정한 경계를 사용한 경우

Fig. 7. Color textures of Fruit image, the case of determining cluster bound by discarding farthest 10% pixels.

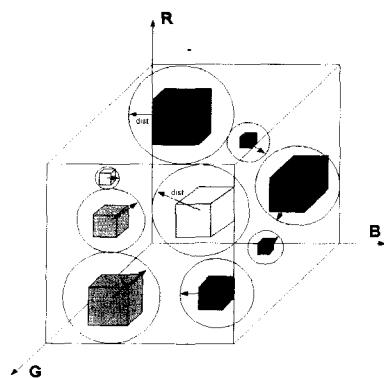
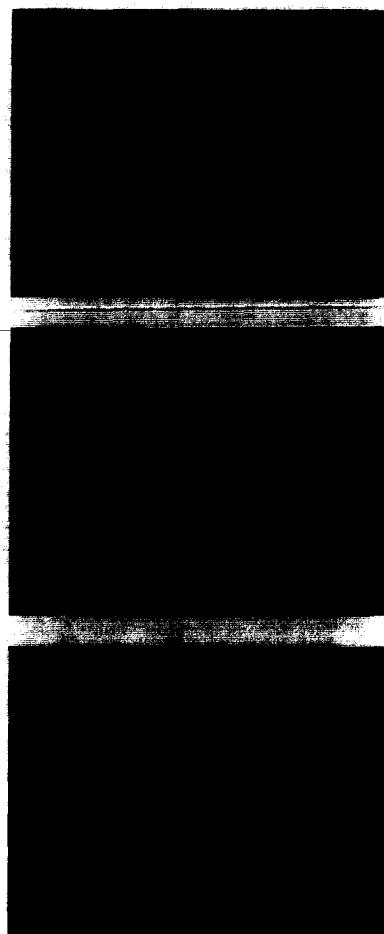


그림 8. 그림 6(c)의 RGB 컬러 공간 상의 표현

Fig. 8. RGB color space representation of Fig. 6(c).

으로 경계 거리가 최대거리의 대략 45% 가량 감소되었고, 이 경계와 위의 사영 함수를 사용하여 한 가지 색상 군집에 의해 두드러지는 텍스처를 얻을 수 있었다.

### 3. 텍스처 특징 추출

Random field 모델은 영상의 화소들을 불규칙 변수



그림 9. Fruit 영상으로부터 얻은 대표적인 컬러 텍스처 5개

Fig. 9. Five sample color textures from a fruit image.

(random variable, r.v.)로 보고 확률적인 방법으로 영상 신호를 처리하는 기법이다. 디지털 영상은 직각 격자  $L = \{(i, j) | 1 \leq i, j \leq M\}$  위에서 정의되는 확률 과정(process)이며, 이때 화소  $x(i, j)$ 는 영상의 왼쪽 위에서  $i$ 번째 행,  $j$ 번째 열에서의 한 실현(realization)이다. 기호를 간단히 하기 위해  $s = j + (i - 1) \times M$ 라고 하면  $x(i, j)$ 는,  $x(s)$ ,  $s = 1, \dots, N$ ,  $N = M \times M$ 로 나타낼 수 있다.

$x(s)$ 가 식 (3)과 같은 조건부 확률 특성을 가지면 이웃 시스템  $B_s$ 에 대하여 Strict Sense Markov(SSM) 조건을 만족한다고 한다. SSM 조건을 이용하여 화소  $x(s)$ 를 식 (4)에서처럼 이웃들의 선형 조합과 오차의 합으로 모델링 할 수 있으며, 이때 오차 항은 식 (5), (6)과 같이 직교 조건을 만족해야 한다. 식에서 영상 신호  $x(s)$ 와 모델링 오차  $e(s)$ 는 평균이 0인 신호라고 가정한다.

$$P(x(s) | x(\hat{s}), \hat{s} \neq s) = P(x(s) | x(\hat{s}), \hat{s} \in B_s) \quad (3)$$

$B_s = \{s + r : r \in B\}$ 이며  $B$ 는 이웃 시스템으로 직각 격자  $L$ 의 유한 부분집합이다.

$$x(s) = \sum_{r \in N} b_r x(s + r) + \sqrt{v} e(s) \quad (4)$$

$$E[e(s)e(\hat{s})] = \begin{cases} \sigma^2 & \text{만약 } s = \hat{s} \\ \sigma^2 \beta_m & \text{만약 } s - \hat{s} = m \in B \\ 0 & \text{그외 모든 경우} \end{cases} \quad (5)$$

$$E[e(s)e(\hat{s})] = \begin{cases} \sigma^2 & \text{만약 } s = \hat{s} \\ 0 & \text{그외 모든 경우} \end{cases} \quad (6)$$

여기서  $e(s)$ 가 식 (7)과 같이  $x(\hat{s})$ 와 상호 배타적이고, 평균이 0, 분산이 1인 신호일 때 이를 Wide Sense Markov(WSM) 모델이라 하고 만약  $e(s)$ 가 가우시안 분포라면 그 때 양방향 WSM 모델을 Gaussian Markov Random Field(GMRF)라 한다.

$$\begin{aligned} E[e(s) | \text{every } x(\hat{s}), \hat{s} \neq s] &= 0 \\ E[e(s)] &= 0 \\ E[e^2(s)] &= 1 \end{aligned} \quad (7)$$

식 (4)는  $Q(s)$ 와  $b$ 를 정의하여 아래와 같이 다시 쓸 수 있다.

$$\begin{aligned} x(s) &= \sum_{i=1}^k b_i [\sum_{m \in N_i} x(s + m)] + \sqrt{v} e(s), \quad s \in B_s \\ &= Q^T(s)b + \sqrt{v} e(s) \end{aligned} \quad (8)$$

여기서

$$Q(s) = \begin{pmatrix} \sum_{m \in N_1} x(s + m) \\ \sum_{m \in N_2} x(s + m) \\ \vdots \\ \sum_{m \in N_k} x(s + m) \end{pmatrix}, \quad b = (b_1, b_2, b_3, \dots, b_k)^T$$

식 (8)의  $Q(s)$  양변에 를 곱하면 식 (6)의 직교 조건에 의하여 가우시안 잡음이 제거되고 아래 식 (9)를 얻을 수 있다.

$$\begin{aligned} Q(s)x(s) &= Q(s)Q^T(s)b \\ E[Q(s)x(s)] &= E[Q(s)Q^T(s)]b \end{aligned} \quad (9)$$

그리고 식 (10)과 같이 식 (9)를 최소 자승 추정(Least Square Estimation)하여 얻어낸 계수  $b_r$ 을 텍스처 계수로 이용한다.

$$b = [\sum_{s \in A} Q(s)Q^T(s)]^{-1} [\sum_{s \in A} Q(s)x(s)] \quad (10)$$

여기서

$$A = \{s \in A ; B_s \subseteq A\}$$

#### 4. 유사도

유사성(similarity)의 척도는  $N$ 개의 입력 영상의 대표 벡터들에 대해 가장 가까운 모델의 대표 벡터를 찾아 거리의 합을 구하는 것이다. 즉, 입력 영상의  $N$ 개의 컬러 벡터 가운데  $i$ 번째 컬러 및 텍스처 벡터  $\vec{ci}_i, \vec{ti}_i$ 와 가장 거리가 가까운 모델 벡터  $\vec{cm}_k, \vec{tm}_k$ 를 찾아 이 둘 사이의 거리를 구하고 이런 과정을  $N$ 개의 입력 벡터에 대해 반복하여 거리의 합을 구하는 것이다. 그러나, 각 대표 벡터들에 속하는 화소의 수에 따라 가중치를 주기 위해 식 (11)와 같은 척도를 사용한다. 또한 두 대표 벡터가 포함하는 화소가 많을수록 가중치를 두기 위해  $\frac{I_i + M_k}{2}$ 를 곱하고, 두 대표 벡터가 포함하는 화소의 수의 차가 적을수록 가중치를 두기

위해  $\frac{\min(I_i, M_k)}{\max(I_i, M_k)}$  를 곱하였다. 그리고 컬러 벡터와 텍스처 벡터의 scale 차이를 보정하기 위해 가중치  $w$  를 곱하였다. 식 (12), 식 (13), 식 (14)는 각각 컬러 벡터만 이용할 경우, 컬러 텍스처만 이용할 경우, 영상의 밝기 성분에서 나타나는 텍스처만을 이용한 경우 적용되는 유사성의 척도이다. 식 (12)와 식 (13)은 입력 영상의  $N$ 개의 컬러 벡터 가운데  $i$ 번째 컬러 또는 텍스처 벡터와 가장 거리가 가까운 모델 벡터를 찾아 이들 사이의 거리를 구하고 이런 과정을  $N$ 개의 입력 벡터에 대해 반복하여 거리의 합을 구하는 것이다. 그리고 영상의 밝기 성분의 텍스처는 영상에서 하나의 벡터로 구해지므로 이들의 유사도는 벡터의 거리를 이용했다.

$$\text{Similarity} = \sum_{i=1}^N \frac{I_i + M_k}{2} \times \frac{\min(I_i, M_k)}{\max(I_i, M_k)} \times (|\vec{ci}_i - \vec{cm}_k| + |\vec{ti}_i - \vec{tm}_k|) \quad (11)$$

$$\text{Similarity} = \sum_{i=1}^N \frac{I_i + M_k}{2} \times \frac{\min(I_i, M_k)}{\max(I_i, M_k)} \times (|\vec{ci}_i - \vec{cm}_k|) \quad (12)$$

$$\text{Similarity} = \sum_{i=1}^N \frac{I_i + M_k}{2} \times \frac{\min(I_i, M_k)}{\max(I_i, M_k)} \times (|\vec{ti}_i - \vec{tm}_k|) \quad (13)$$

$$\text{Similarity} = |\vec{ti} - \vec{tm}| \quad (14)$$

입력 영상	$I_i$	( $i$ 번째 군집의 화소수) / (영상 전체의 화소수)
	$\vec{ci}_i$	입력 영상의 $i$ 번째 컬러 벡터
모델 영상	$M_k$	( $k$ 번째 군집의 화소수) / (영상 전체의 화소수)
	$\vec{cm}_k$	모델 영상의 $N$ 개의 컬러 벡터 가운데 입력 영상의 $i$ 번째 컬러 벡터와의 거리를 최소화하는 벡터
	$\vec{tm}_i$	모델 영상의 $N$ 개의 컬러 벡터 가운데 입력 영상의 $i$ 번째 텍스처 벡터와의 거리를 최소화하는 벡터

### 5. 전체적 구성

본 과제에서는 컬러 히스토그램 공간에 k-means 군집화 알고리듬을 적용하여 구한 대표 컬러 벡터로 영상의 히스토그램 공간을 나타낸다. 그리고 영상의 각 화소값과 대표 컬러 사이의 벡터 거리를 밝기로 사용

하여 Gray 영상을 만들면 영상을 대표하는 RGB값에 의해 두드러지는 텍스처 성분을 얻을 수 있다. 이렇게 구해진 대표 벡터와 텍스처 계수로 영상을 표현하고 이들의 정합을 통해 영상을 검색하는 방법을 연구의 기본 방향으로 한다. 본 기법은 “색상에 의한” 텍스처 특성을 이용하므로 Smith<sup>[3]</sup>나 Belongie<sup>[4]</sup> 등의 방법과 달리 색상 정보와 텍스처 정보로 후보 영역 특징을 표현하기 위한 영역화 등의 부가 연산이 없어 고속의 정합에 적합하다.

이러한 제안하는 영상의 대표 벡터들만을 비교하는 방법의 장점을 살펴보면 다음과 같다. 첫째로는 RGB축의 레벨에 따라 bin을 나누어 bin에 속하는 화소의 수로서 히스토그램을 나타내는 방법보다 히스토그램의 특징을 잘 반영할 수 있어 적은 수의 벡터로 히스토그램을 효율적으로 표현한다는 것이다. 둘째로, 제안하는 기법은 히스토그램에 의해 두드러지는 텍스처의 계수를 이용하므로, 컬러 히스토그램만을 사용한 경우나, 영상의 밝기 정보를 텍스처로 이용한 경우 보다 더 풍부한 정보를 포함하게 되며, 컬러 정보와 텍스처 정보를 이용해 객체로 판단될 수 있는 후보 영역의 특징을 표현하기 위한 영역화, 경계 검출이 따르지 않으므로 고속의 정합에 적합하다는 장점이 있다.

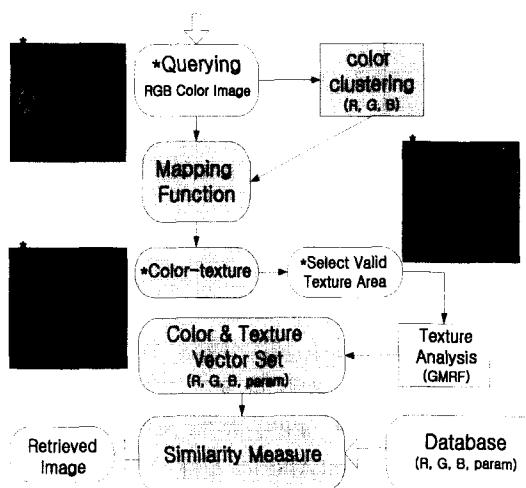


그림 10. 제안하는 기법의 전체적 구성

Fig. 10. Overall scheme of the proposed algorithm.

### IV. 실험 결과

실험에는 Histogram intersection을 제안한 Swain과

Ballard가 사용한 데이터 베이스와 Virage 데이터 베이스 그리고 MIT 대학의 Vision and Modeling Group이 제공하는 VisTex 데이터 베이스가 사용되었다. Swain과 Virage 데이터 베이스의 각 영상은 크기가 작고 많은 면적의 단일색 배경으로 이루어졌으며 영상에 텍스트 특성이 잘 나타나지 않는다. 따라서 본 실험에서는 Swain과 Virage 데이터 베이스에서 컬러 색상 특징만을 이용한 검색을 수행하였다. 텍스트 특성 실험에는 텍스트 성분이 잘 나타나는 컬러 영상들로 구성된 VisTex 데이터 베이스를 사용하였다. 현재 Reference textures와 Texture scenes가 제공되는데, Reference texture의 132개의 영상 데이터 베이스에서 실험을 수행하였다. 이 데이터 베이스에는 각 범주(category)마다 여러 다양한 영상을 보유하고 있어 제안하는 기법의 효용성을 실험하기에 적절하다.

실험을 수행한 시스템의 사양은 아래 표 1과 같다.

표 1. 검색을 수행한 시스템 사양

Table 1. System specification.

System specificatoin	
CPU	Intel Pentium2 celeron 333MHz
Memory	64MBytes
OS	windows NT 4.0 workstation

### 1. 컬러 정보만을 이용한 검색 실험

- Swain의 알고리듬과 군집화하여 얻은 대표 벡터를 이용한 방법의 비교
- 영상을 표현하는 적절한 컬러 벡터 수 조사

이번 절에서는 본 논문에서 제안하는 영상의 색상 특성 표현 방법의 검색 효율과 영상을 대표하는 적절한 대표 벡터 수를 알아보기 위한 실험을 수행한다.

우선 표 2에서와 같이 Swain 데이터 베이스에서 Histogram intersection과 k-means 군집화 하여 얻은 대표 컬러 벡터를 이용하는 방법을 비교했다. Swain은 66개의 데이터 베이스에 대해 32개의 입력을 주어 얼마나 잘 찾는가를 실험하였다<sup>[8]</sup>. 입력 영상은 모델 영상 내의 물체를 회전하거나, 늘리거나, 일부분을 가리거나, 혹은 안보이던 일부분이 보이는 등의 변형을 하였다. 대표 컬러 벡터를 이용하는 경우는 전자에서 사용된 32개의 영상을 입력으로 군집의 수가 20, 15, 10, 5인 네 가지 경우에 대해 검색을 수행하였다. 표 2의

각 순위에 쓰여진 숫자는 모델 영상  $a_n$ 에 변형을 가한  $a'_n$ 을 입력 영상으로 검색을 수행했을 경우 영상  $a_n$ 이 검색되어진 순위를 누적한 값이다.

실험 결과 군집의 수가 20, 15, 10인 경우 제안하는 기법은 Histogram intersection 보다 효과적인 검색 결과를 나타냈으며, 10개 가량의 대표 컬러 벡터 만으로도 영상의 색상 특징을 효율적으로 표현했다. 특히 이 기법은 작은 크기의 영상 검색에서 상대적으로 보다 높은 검색 성능을 나타냈다.

표 2. Swain 알고리듬과 군집화하여 얻은 대표 벡터를 이용한 방법의 비교(Swain 데이터 베이스)

Table 2. Comparison of Histogram intersection and the propose color vector representation(on Swain database) Database size : 66, #, of inputs : 32

Size	Method	Correct Match Placement							
		1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	6 <sup>th</sup>	7 <sup>th</sup>	8 <sup>th</sup>
	Histogram Intersection	29	3						
128	vector #: 20	31	1						
	vector #: 15	31			1				
	vector #: 10	30			1				1
	vector #: 5	25	4		1				1
	Histogram Intersection	24	7	1					
90	vector #: 20	31	1						
	vector #: 15	29	1	1	1				
	vector #: 10	30	1	1					
	vector #: 5	25	5			1		1	
	Histogram Intersection	24	7	1					
32	vector #: 20	31	1						
	vector #: 15	29	1	1	1				
	vector #: 10	30	1	1					
	vector #: 5	25	5			1		1	
	Histogram Intersection	24	7	1					
23	vector #: 20	31	1						
	vector #: 15	29	1	1	1				
	vector #: 10	30	1	1					
	vector #: 5	25	5			1		1	
	Histogram Intersection	24	7	1					

적절한 컬러 벡터 수 결정을 위한 보다 일반적인 실험을 위해 992개의 Virage 데이터 베이스에서 군집화가 20, 15, 10, 5인 경우에 대한 검색을 수행했다. 이 경우 역시 입력 영상으로 모델 영상에 각종 변형을 가한 영상을 사용했다. 그러나 Swain 데이터 베이스 실험에서 사용한 가려짐, 이동, 회전, 크기 변화 등의 변형 이외에 흐림(blur), 선명(sharpen), 단면화, 모자이크, 분열 및 영상의 색상 히스토그램을 크게 바꾸는 잡음 추가, 고주파 성분 제거 등의 변화를 가했다. 변형의 정도가 커져 Swain 데이터 베이스에서 대표 컬러 벡터를 이용해 수행한 실험에 비해 검색 결과의 분산이 커졌으므로, 상위 rank의 결과를 비교하여 성능을 평가하지 않고 8<sup>th</sup> rank 내에 찾지 못한 영상 개수를 비교한다. 표 3은 Virage 데이터 베이스에서의 검색 결과이고, 마지-

막 열은 영상의 왜곡이 너무 커서 찾지 못한 영상의 개수이다. 표 3의 각 순위에 쓰여진 숫자는 모델 영상  $a_n$ 에 변형을 가한  $a'_n$ 을 입력 영상으로 검색을 수행했을 경우 영상  $a_n$ 이 검색되어진 순위를 누적한 값이다. 실험 결과 군집의 수가 10일 경우, 찾지 못한 영상의 수가 군집 수 20일 때와 거의 같아 10개 가량의 대표 컬러 벡터 만으로도 영상의 색상 특징을 효율적으로 표현함을 알 수 있다.

표 3. 대표 컬러 벡터를 이용한 실험 결과  
(Virage 데이터 베이스) 데이터 베이스  
크기 : 992, 입력 영상 개수 : 50, 영상  
크기 : 96×64

Table 3. Correct Match Placement of the proposed color vector representation(on Virage database) Database size : 992, #. of inputs : 50, image xiae : 96×64.

Method	Vector#	Correct Match Placement							
		1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	6 <sup>th</sup>	7 <sup>th</sup>	8 <sup>th</sup>
Color vector	20	33	3			1			13
	15	33	2		1	1	1		12
	10	28	4			3		1	14
	5	27	2			3		1	17

색상을 군집화 하는데 필요한 시간은 표 4와 같다. 본 실험에서 군집화는 식 (15)를 만족할 때까지 반복한다. 여기서 왜곡(distortion)은 영상의 training vector와 codebook의 차의 절대치의 합이다. 식 (15)는 전단계와 현단계의 왜곡의 차를 현단계 왜곡으로 정규화 한 값이  $10^{-2}$ 보다 작아질 때 까지 군집화를 반복하게 되는 조건이다. 따라서 같은 크기의 영상이라도 영상마다 군집화에 소요되는 시간이 다르게 나타난다. 그리고 하나의 영상의 군집화에 걸리는 시간도 시스템의 상태에 따라서 조금씩 변화한다. 그러므로 같은 크기의 영상 20개의 군집화에 소요되는 시간의 평균 값을 기록하였고 각 영상의 군집화 시간은 5회 측정치의 평균을 사용했다. QueryPerformanceCounter, QueryPerformanceFrequency 함수를 이용하여 수행 시간을 측정하고  $10^{-6}$  초 단위까지 표기한다. 연산 과정 중 전역 변수로 미리 할당해 놓을 경우 반복할 필요가 없는 변수 선언이나 메모리 할당에 걸리는 시간은 제외했다.

$$\frac{\text{prior\_distortion} - \text{current\_distortion}}{\text{current\_distortion}} < 10^{-2} \quad (15)$$

표 4. 영상의 크기와 군집화에 소요되는 시간  
Table 4. K-means clustering performance.

Size (color depth)	Vector#	Implementing time for clustering
$128 \times 90$ (16M)	20	0.851725 sec
	15	0.527954 sec
	10	0.278065 sec
	5	0.095770 sec
$512 \times 512$ (16M)	20	1.008166 sec
	15	0.608303 sec
	10	0.315545 sec
	5	0.105104 sec

컬러 벡터로 영상의 컬러 히스토그램을 대표하는 방법은 데이터 베이스의 크기, 영상의 크기 등의 조건에 따라 10개 내외의 적절한 군집수 선택이 가능하고 영상 데이터 베이스를 보다 효율적으로 인덱싱할 수 있으므로 대형 데이터 베이스에서 검색에 소요되는 시간을 줄일 수 있는 장점이 있다. 또한 영상의 팔레트 정보를 이용하면 보다 빠른 군집화가 가능하다.

## 2. 컬러 히스토그램과 컬러 텍스처를 이용한 검색 실험

### - VisTex 데이터 베이스에 대한 실험

이번 절에서는 컬러 텍스처를 이용한 검색 기법의 효용성을 알아보기 위한 실험을 수행했으며, VisTex 데이터 베이스를 사용했다. 본 실험에서는 표 2와 표 3의 Correct Match Placement가 아닌 Group Match Placement를 구하고 이로써 성능을 평가한다. 표 6의 Group Match Placement의 각 순위에 쓰여진 숫자는 모델 영상 그룹  $A_n$ 의 멤버 영상인 을 입력으로 검색을 수행했을 때 그룹  $A_n$ 의 다른 멤버 영상들이 검색되어진 순위를 누적한 값이다. 검색에는 변형을 하지 않은 모델 영상을 입력으로 사용했다. 따라서 어떤 방법으로 검색을 수행해도 1<sup>st</sup> rank에는 입력 영상과 동일한 모델 영상이 찾아지므로 1<sup>st</sup> rank는 성능 평가에서 제외했다.

표 5의 영상 그룹은 VisTex 데이터 베이스의 범주를 따라 나누었다. 예를 들어 Fabric1은 VisTex의 Fabric

범주에 속하는 영상 가운데 대나무 격자 텍스처 영상 그룹으로 묶은 것이고 Paintings는 VisTex의 Paintings 범주에 속하는 영상 가운데 하나의 영상으로부터 파생된 영상을 모은 것이다, Water는 VisTex의 Water 범주를 그대로 사용한 것이다.

표 5. 입력 영상 그룹

Table 5. Input image groups.

Group name	#. of members
Fabric 1	4
Fabric 2	2
Fabric 3	3
Food	5
Grass	2
Leaves	4
Paintings	5
Tiles 1	2
Tiles 2	4
Water	8

이 절의 실험은 10개의 그룹 각각의 입력 영상에 대해, 대표 벡터 수 5, 20일 때 대표 컬러 벡터를 이용한 기법, 영상의 밝기 성분의 텍스처를 이용한 기법, 군집 수 5일 때 컬러 텍스처를 이용하는 기법, 군집 수 5일 때 컬러 텍스처와 대표 컬러 벡터를 이용한 기법, 이 5 가지 방법으로 검색을 수행했다. 그리고 검색된 순위에 따라 가중치를 주기 위해 2<sup>nd</sup> rank에는 7점을, 8<sup>th</sup> rank에는 1점을 부여하고 총점로 성능을 평가한다. 텍스처 계수를 구할 때는 2차 GMRF 마스크를 사용했으며, 유사

도는 검색 방법에 따라 식 (11)부터 식 (14)를 적용하여 계산했고, 식 (11)의 가중치  $w$ 는 1000으로 했다.

네번째 행의 군집수가 5일 때 컬러 텍스처만을 이용한 검색 결과를 보면, 두번째 행의 군집 수가 20일 때 대표 컬러 벡터를 이용해 검색한 것과 비슷한 성능을 보이며, 세번째 행의 영상의 밝기 성분의 텍스처를 이용한 검색에 비해 월등한 성능을 보여 컬러 텍스처의 정보량이 밝기 성분의 텍스처 보다 더 많음을 알 수 있다. 또 다섯번째 행의 군집수 5일 때 컬러 텍스처와 대표 컬러 벡터를 이용한 검색은 다른 방법들에 비해 향상된 성능을 보인다.

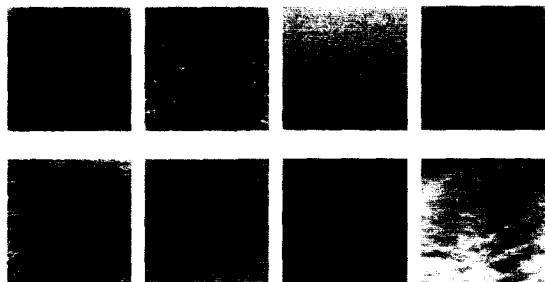
그림 11은 VisTex 데이터 베이스에서 Water 그룹의 멤버 영상을 입력으로 표 6의 각 행에 적용된 방법으로 검색을 수행한 예이다. 그림 11의 첫째 행은 Water 그룹의 8개의 멤버 영상이고, 첫째 행의 세 번째 영상이 입력이다. 입력에 어떤 변형도 가지지 않았으므로 모든 방법의 1<sup>st</sup> rank는 입력과 같은 영상이 찾아지게 되며, 이는 검색 성능 평가에서 제외한다. 그림 11의 예에서 대표 컬러 벡터만을 이용한 방법은 Water 그룹의 멤버 영상을 하나도 찾지 못했다. 그리고 영상의 밝기 성분의 텍스처를 이용한 검색에서 2<sup>nd</sup>, 6<sup>th</sup> rank에서, 컬러 텍스처를 이용한 검색에서는 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, 7<sup>th</sup> rank에서 멤버 영상을 찾아냈다. 그리고 컬러 텍스처만 이용하는 경우와 컬러 벡터를 같이 이용하는 경우는 색상 분포에 따라 검색된 영상의 순위가 바뀌었다.

표 7은 2차 GMRF 마스크를 사용할 경우 한 장의 컬러 텍스처 계수를 계산할 때 걸리는 시간이고, 영상의 크기에 따라 증가 되었다. QueryPerformanceCounter,

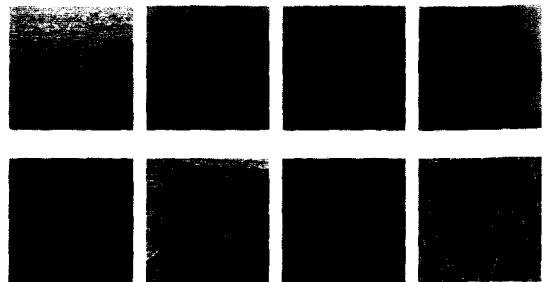
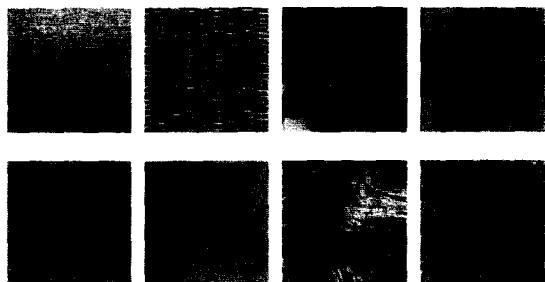
표 6. 컬러 텍스처를 이용한 검색 기법의 효용성 실험(Vis Tex 데이터 베이스) 데이터 베이스 크기 : 132, 입력 영상 개수 : 10, 영상 크기 : 512×512

Table 6. Retrieving result(on Vis Tex database) Database size : 132, #. of inputs : 10, image size : 512×512.

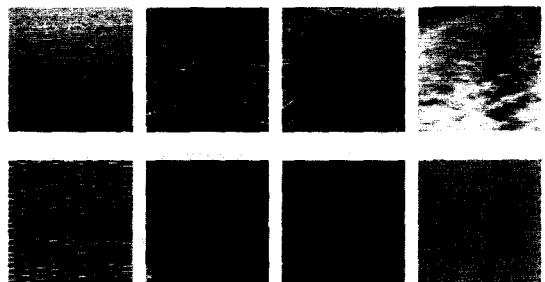
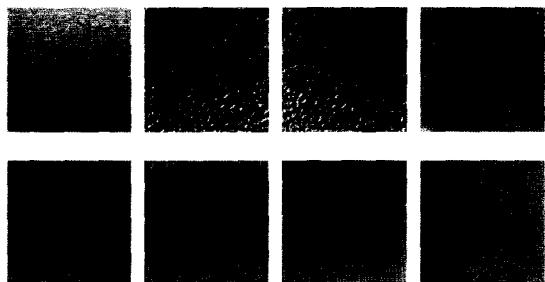
Method	Group Match Placement									Total ranked	Total score
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	6 <sup>th</sup>	7 <sup>th</sup>	8 <sup>th</sup>	8 <sup>th</sup> <		
	7points	6points	5points	4points	3points	2points	1points	0points			
Color vector only (vector # : 5)	10	3	2				1			6	35
Color vector only (vector # : 20)	10	8	2				2			12	72
Mono texture only (mask : 2 <sup>nd</sup> order)	10	3	1	1		1				6	30
Color vector only (vector # : 5, mask : 2 <sup>nd</sup> order)	10	6	2	3			1			12	71
Color vector only (vector # : 5, mask : 2 <sup>nd</sup> order)	10	8	4	3	1		1	1		18	102



Model Image Group : Water (#. of members : 8)

[3] Mono texture only (mask : 2<sup>nd</sup> order)

[1] Color vector only (vector # : 5)

[4] Color texture only (vector # : 5, mask : 2<sup>nd</sup> order)

[2] Color vector only (vector # : 20)

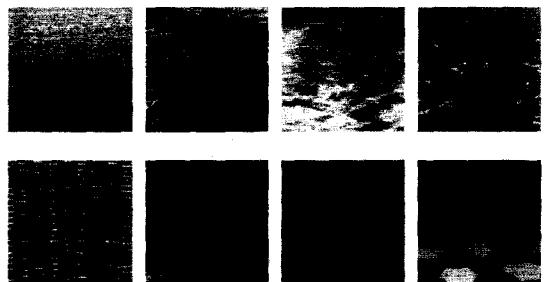
[5] Color texture with color vector (vector # : 5, mask : 2<sup>nd</sup> order)

그림 11. Vis Tex 데이터 베이스에 몇 가지 방법을 적용하여 검색을 수행한 예 (영상 크기 : 512×512)

Fig. 11. Retrieval results which some different methods are applied to Vis tex database. (image size : 512×512)

QueryPerformanceFrequency 함수를 통하여 시간을 측정하였고, 각 영상에 대해 10회 측정한 평균치를  $10^{-6}$  초 단위까지 표기한다. 본 실험에서 쓰인 영상의 크기는  $512 \times 512$ 이고 군집수는 5이므로 컬러 텍스처 계수를 계산하는데 2.6초 가량의 시간이 소요된다. 그러나 텍스처 계수 계산 시 텍스처가 강하게 나타나는, 예를 들어,  $222 \times 223$  정도의 영역에서 계수를 구하게 되면 시간이 크게 단축될 수 있다. 이렇게 할 경우 군집화하고 텍스처 계수를 계산하는데 0.5초 가량의 시간이 소요되며 이는 대표 벡터 수를 20개로 군집화하는 시간 보다 적다.

표 7. 컬러 텍스처 계수 계산에 소요되는 시간  
Table 7. The performance of calculating the color texture.

Size(color depth)	Implementing time for analysing a color texture (mask : 2 <sup>nd</sup> order)
$128 \times 90$ (16M)	0.017661 sec
$222 \times 223$ (16M)	0.078667 sec
$338 \times 250$ (16M)	0.138858 sec
$370 \times 512$ (16M)	0.303332 sec
$512 \times 512$ (16M)	0.524624 sec

## V. 결 론

본 연구에서는 컬러 히스토그램을 대표하는 컬러 벡터와 컬러 텍스처를 영상의 특징으로 제안하였다. 컬러와 텍스처를 이용하는 기준의 기법들이 자체 차원의 기술을 위해 색상과 텍스처 특성을 따로 추출하고 이를 조합하기 위한 많은 부가 연산이 따름에 비해, 제안하는 기법은 색상 군집에 의한 텍스처를 사용하므로 영역화 등의 기하학적 정보를 추출하기 위한 과정 및 부가 연산이 필요 없고 고속의 정합에 적합한 장점을 가진다.

본 논문에서는 영상의 컬러 히스토그램을 k-means 알고리듬으로 군집화하여 얻은 군집의 중심을 대표 컬러 벡터로 사용하였으며, 10개 가량의 대표 컬러 벡터로 영상을 효율적으로 표현할 수 있었다. 특히 컬러 벡터를 이용한 검색은 작은 크기의 영상 검색에서 Histogram intersection에 비해 상대적으로 높은 검색 성능을 나타냈다. 그리고 같은 그룹의 영상을 얼마나 잘 찾는지를 실험하여 컬러 텍스처가 영상의 밝기 성분의 텍스처보다 정보량이 많음을 확인했으며, 컬러 텍스처와 대표 컬러 벡터를 이용한 검색은 보다 향상된 성능을 나타냈다.

컬러 텍스처는 영상을 기술하는 효율적 특성으로 향후 영상의 색상 특성에 따른 적절한 벡터 수 결정, 사람의 인지도에 기반한 컬러 히스토그램 군집화, HSV 색상 공간에서의 군집화, 이에 따른 새로운 사영 합수, wavelet이나 Gabor filter, local activity spectrum을 이용한 컬러 텍스처 특징 추출 등의 연구가 진행될 것이다.

## 참 고 문 헌

- [1] Dan Schonfeld and Dan Lelescu, "VORTEX: Video Retrieval and Tracking from Compressed Multimedia Databases," *International Conference on Image Processing*, vol. 3, WA04.02, 1998.
- [2] I. K. Park, I. D. Yun, and S. U. Lee, "Color Image Retrieving Using Hybrid Graph Representation," *Image and Vision Computing*, vol 17, no. 7, pp. 465~474, May 1999.
- [3] J.R Smith and S. -F. Chang, "Automated image retrieval using color and texture," Technical Report TR 414-95-20, Columbia University, 1995.
- [4] Serge Belongie, Chad Carson, Hayit Greenspan and Jitendra Malik, "Color- and Texture-based Image Segmentation Using EM and Its Application to Content-based Image Retrieval," *International Conference on Computer Vision*, 1998.
- [5] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, "The QBIC Project: Querying Images by Content Using Color, Texture, and Shape," *SPIE* vol. 1908, pp. 173~187, 1993.
- [6] J. R. Smith and S. -F. Chang, "VisualSEEK: a fully automated content-based image query system," *Proceedings of ACM International Conference on Multimedia*, Boston, MA, Nov. 1996.
- [7] <http://www.virage.com>, Virage Inc.
- [8] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11~32, Nov. 1991.
- [9] D. Androutsos, K. N. Plataniotis and A. N. Venetsanopoulos, "Distance Measures for Color Image Retrieval," *International Conference on Image Processing*, vol. 2 of 3, pp. 770~774, 1998.
- [10] Yihong Gong, Guido Proietti, and Christos Faloutsos, "Image Indexing and Retrieval Based on Human Perceptual Color Clustering," *Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 578~583, 1998.
- [11] Chung-Sheng Li and Vittorio Castelli, "Deriving Texture Feature Set for Content-based Retrieval of Satellite Image Database," *International Conference on Image Processing*, 1997.
- [12] Bo Tao & Bradley W. Dickinson, "Recognition and Retrieval of Textured Images Using Gradient Indexing," *International Conference on Image Processing*, vol. 1 of 3, 1998.
- [13] R. Chellappa and S. Chatterjee, "Classifica-

- tion of Textures Using Gaussian-Markov Random Fields," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-33, no. 4, pp. 959, 1985.
- [14] T. Randen and J. H. Husoy, "Image Content Search by Color and Texture Properties," *International Conference on Image Processing*, vol. 1, pp. 580-583, 1997.
- [15] Vision and Modeling Group, MIT Media Laboratory,<http://vismod.www.media.mit.edu/vismod/>

---

 저자 소개
 

---

**李亨九(正會員)**

1997년 2월 한국외국어대학교 제어계측공학과 학사  
1999년 8월 한국외국어대학교 전자제어공학부 석사 현  
재 (주) 세연 연구원


**尹逸東(正會員)**

1989년 2월 서울대 세어계측공학  
학사 1991년 2월 서울대 제어계측  
공학 석사 1996년 8월 서울대 제어  
계측공학 박사 1997년 2월 대우전  
자 영상연구소 선임연구원 현재 한  
국외국어대학교 전자제어공학부 조  
교수 주관심분야는 영상신호처리, 멀티미디어 데이터의  
내용기반 검색, 3차원 데이터 취득 및 모델링