

객체의 모양과 색상특징을 이용한 내용기반 영상검색 기법

정희원 박종현*, 박순영*, 오일환*

A Content-Based Image Retrieval Technique Using the Shape and Color Features of Objects

Jong-Hyun Park*, Soon-Young Park*, Il-Whan Oh* *Regular Members*

요 약

본 논문에서는 객체들의 공간적 특성이 반영된 시각적인 특징벡터를 이용한 내용기반 영상검색 알고리즘을 제안한다. 제안된 검색 기법은 여러 색상으로 이루어진 객체들을 표현하기 위하여 가우시안 혼성 모델을 적용하며 모델의 최대유사 파라미터는 EM 알고리즘을 사용하여 추정한다. GMM을 기반으로 하여 분할된 각 객체들로부터 Fourier descriptor와 색상 히스토그램을 사용하여 모양과 색상 특징을 추출하게 된다. 영상 검색은 두 단계로 구성되는데 첫 단계에서는 공간적인 모양 특징을 추출하여 모양이 유사한 객체들을 후보 영상으로 압축하게 되며 마지막으로 객체의 색상 히스토그램에 의하여 검색이 수행된다. 실험 결과 제안된 알고리즘은 분할된 객체의 공간적, 시각적 특징을 이용하여 효율적으로 검색을 수행할 수 있음을 보여준다.

ABSTRACT

In this paper we present a content-based image retrieval algorithm using the visual feature vectors which describe the spatial characteristics of objects. The proposed technique uses the Gaussian mixture model(GMM) to represent multi-colored objects and the expectation maximization(EM) algorithm is employed to estimate the maximum likelihood(ML) parameters of the model. After image segmentation is performed based on GMM, the shape and color features are extracted from each object using Fourier descriptors and color histograms, respectively. Image retrieval consists of two steps: first, the shape-based query is carried out to find the candidate images whose objects have the similar shapes with the query image and second, the color-based query is followed. The experimental results show that the proposed algorithm is effective in image retrieving by using the spatial and visual features of segmented objects.

1. 서 론

최근 컴퓨터의 급속한 성능 향상과 네트워크의 확산으로 많은 양의 멀티미디어(multimedia) 정보를 매일 다양한 형태로 접하고 있으며 정보의 획득, 저장, 처리, 전송 등의 기술 발전으로 이를 사용자의 목적에 맞게 활용할 수 있는 기회가 증가하고 있다. 이러한 정보들은 일반 사용자와 기술을 개발하

는 전문가들에 의하여 일반화가 되었으며 방송이나 인터넷을 통하여 사람들은 매일 많은 양의 매체(masses)를 접하고 있다. 전문가들은 계속적으로 새로운 형태의 정보를 새로운 목적에 맞게 재구성하고 있으며 사용자들은 그러한 새로운 정보 형태를 이용하여 개인의 audiovisual의 기억들을 보존하기를 원하고 있다. 또한 공공접속 장치에 자신의 데이터베이스를 구축하여 모든 사람들이 효율적인 정보

* 목포대학교 전자공학과(jhpark@chungkye.mokpo.ac.kr)
논문번호 : 99113-0328, 접수일자 : 1999년 3월 28일

를 검색(retrieval)하고 공유하기를 바라고 있다^[1]. 그러나 이러한 방대한 데이터베이스로부터 원하는 정보를 찾기 위하여 효율적으로 필터링(filtering)할 수 있는 기술이 요구되고 있으며 이를 위하여 멀티미디어가 가지고 있는 속성을 효율적으로 추출하고 이를 이용한 데이터베이스의 구축 및 검색이 수행되어야 한다.

초기의 검색 시스템에서 많이 이용한 방법으로 내용을 색인한 후에 질의자가 키워드를 사용하여 검색하는 텍스트 기반 검색(text-based retrieval) 방법은 색인 구축이 수동적으로 수행되었기 때문에 시간과 인력의 소모가 많았으며 멀티미디어 데이터가 가지는 속성을 정확하게 표현하기 힘들고 또한 검색시 개인의 주관성 때문에 검색 질의어로 사용하는 키워드의 선정에도 문제점이 뒤따랐다.

반면에 내용기반 멀티미디어 검색기술(content-based multimedia retrieval technique)은 audiovisual 특징들의 내용을 키워드나 텍스트에 의해 묘사할 경우 발생하는 모호성(ambiguity)의 문제점을 해결할 수가 있는데 이 방법은 색상(color), 질감(texture), 모양(shape)과 같은 시각적 특징을 추출하여 색인 과정에 사용함으로써 데이터베이스 구축에 필요한 시간 및 인력의 소모를 줄일 수 있으며 컴퓨터 그래픽스 기술을 이용하여 질의 영상을 생성할 수 있어서 텍스트 기반보다 효율적으로 정보를 검색할 수 있다.

가장 잘 알려진 내용기반 영상검색 시스템은 IBM의 QBIC^[2]로 영상의 전체적인 색상과 질감 정보를 기반으로 정지영상 및 동영상 검색을 수행할 수 있었으며, Berkely 대학의 CBIR 시스템^[3]은 영상내 동질의 색상과 질감 특징에 의해 자동적으로 객체를 분할하여 검색을 수행하였다. 그리고 Columbia 대학의 Visual SEEK와 WebSEEK^[4]는 인터넷 상에서 영상을 검색할 수 있는 사용자 인터페이스를 제공하여 영상 정보를 검색할 수 있는 시스템으로 영상이 가지고 있는 객체의 시각적인 특성과 공간적인 상관관계를 잘 묘사 할 수 있는 인터페이스를 제공하고 있으며, MIT의 Photobook 시스템^[5]은 영상의 통계적인 특성에 기반을 두고 있으며 모양, 질감 특징 벡터를 추출하여 검색을 수행하는데 얼굴 인식 분야에 응용하고 있다.

일반적으로 내용기반 검색은 영상 전체에 대한 특징 벡터를 추출하여 검색에 이용하는 방법과 영상내 객체들을 분할한 후에 각 객체에 대한 특징 벡터를 이용하는 방법이 있다. 그러나 전체적인 영

상에서의 특징 추출은 영상내 객체들이 포함하는 공간적인 특징 정보의 독립성이 흐려지기 때문에 검색에 있어서 효율성이 떨어지는 단점이 있다.

그러므로 본 논문에서는 영상내의 여러 색상으로 이루어진 객체들에 GMM(Gaussian mixture model)을 적용하여 모델링하고 EM(expectation maximization) 알고리즘^[6]을 이용하여 추정된 통계적인 분포 특성에 따라서 동질의 영역을 포함하는 샘플들의 군(cluster)으로 객체를 분할한 후에 각 객체들로부터 공간적인 특징 정보가 반영된 시각적인 특징 벡터를 추출하여 내용기반 영상검색을 수행하는 방법을 제안한다. 특히 회전(rotation), 이동(translation), 축척(scale) 등의 기하학적인 변환에 불변한 Fourier descriptor를 분할된 객체에 적용하여 공간적인 모양특징을 추출하여서 모양이 유사한 객체들을 후보 영상으로 압축한 후에 선별된 후보 영상들로부터 객체에 포함된 시각적인 색상 히스토그램을 추출하여 내용기반 영상검색을 수행하는 방법을 기술한다. 본 논문의 구성은 II장에서 GMM을 이용한 영상내 객체 분할에 대하여 기술하고, III장에서는 분할된 객체들로부터 추출된 공간적인 특징정보가 반영된 시각적인 특징벡터를 이용한 내용기반 영상검색에 대하여 설명한다. IV장에서는 제안된 방법과 기존의 방법을 적용한 실험 결과에 대하여 논하였으며, 마지막으로 V장은 제안된 방법에 대한 결론으로 구성되어 있다.

II. GMM을 이용한 객체 분할

영상에 포함된 객체의 분할은 화소들이 가지고 있는 특성을 이용하여 균일한 영역(homogeneous region)으로 나누어 영상의 데이터를 분석하고 처리하는 것으로 영상처리 및 컴퓨터 비전 분야에서 많이 응용되고 있다. 본 논문에서는 내용기반 영상검색에 있어서 영상이 가지고 있는 전체적인 속성(attribute)을 이용할 경우 영상내 포함된 객체의 독립적인 속성이 희석되기 때문에 통계적인 특성을 기반으로 객체를 분할하고 객체가 가지고 있는 모양 및 색상 성분을 추출하여 내용기반 영상검색에 이용하고자 한다.

영상내의 화소들은 시공간적으로 독립적이고 동일하게 분포(independent and identically distributed : i.i.d.)되었으며 K 개의 객체로 구성되었다고 가정하면 각 객체의 화소들은 식 (1)과 같이 K 개의 다

른 확률 밀도 함수(probability density function: pdf)로 구성된 Gaussian mixture model(GMM)로 표현될 수 있다^[7].

$$g(x|\Phi) = \sum_{i=1}^K \alpha_i f_i(x|\theta_i) \quad (1)$$

여기서 $f_i(x|\theta_i)$ 는 n 차원 벡터의 mean μ_i 과 $n \times n$ 행렬의 covariance Σ_i 에 의한 i 번째 객체의 확률 밀도 함수를 나타내며 α_i 는 영상내 객체의 weight로 $\sum_{i=1}^K \alpha_i = 1$ 이고 Φ 는 $i=1, 2, \dots, K$ 일 때 GMM의 $\alpha_i, \mu_i, \Sigma_i$ 파라미터를 포함하는 벡터이다. 본 연구에서는 R.G.B로 입력된 영상을 H.S.I(Hue-Saturation-Intensity) 색상 모델로 변환한 후 각 화소에 대한 H와 S성분을 이용하여 GMM을 적용한다. θ_i 에 의한 2D 색상 공간의 다차원 확률 밀도 함수 f_i 는 다음과 같이 나타낼 수 있다.

$$f_i(x|\theta_i) = \frac{1}{2\pi \det \Sigma_i^{1/2}} e^{-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)} \quad (2)$$

GMM을 영상내의 화소들에 이상적으로 적합(fitting) 시키기 위한 척도는 식 (3)과 같은 N 개의 전체화소에 대한 total log likelihood로 측정될 수 있는데 이 값을 최대가 되도록 하는 파라미터 Φ 의 추정엔 EM 알고리즘을 적용하여 구할 수가 있다.

$$\mathcal{L}(X|\Phi) = \log L(X|\Phi) = \sum_{j=1}^N \log g(x_j|\Phi) \quad (3)$$

EM 알고리즘에 의한 파라미터의 total log likelihood 추정에서는 미리 주어진 초기치 Φ^0 에 대해서 다음과 같이 expectation step(E-step)과 maximization step(M-step)을 반복적으로 적용하여 파라미터 Φ 를 추정할 수가 있으며 EM 알고리즘 순서도는 그림 1에 나타나 있다.

E-step : 식 (4)처럼 가능한 데이터 X 와 현단계 파라미터 Φ^p 의 조건 아래서 완전한 데이터 Y 에 대한 log likelihood의 기대치를 계산한다.

$$Q(\Phi, \Phi^p) = E(\log g(Y|\Phi)|X, \Phi^p) \quad (4)$$

M-step : $Q(\Phi, \Phi^p)$ 를 최대로 할 수 있는 Φ 를 추정하여 새로운 단계인 Φ^{p+1} 로 한다.

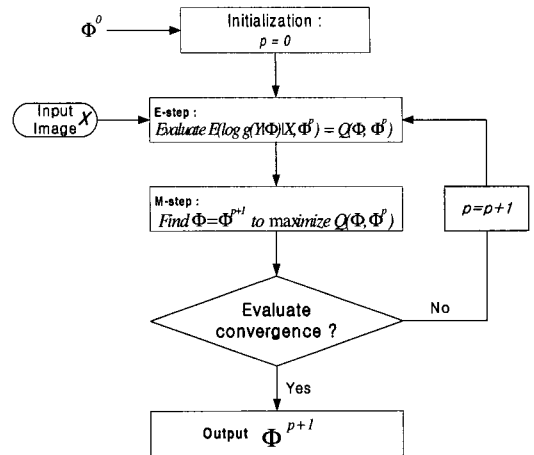


그림 1. EM 알고리즘 기반의 파라미터 추정 순서도

여기서 Gaussian mixture의 파라미터 추정을 위한 EM 알고리즘은 $Q(\cdot)$ 함수가 연속적인 경우 발산하지 않고 최대값으로 수렴하기 때문에 log likelihood 방정식의 편도함수를 0으로 취하여 유도할 수 있으며 그 결과는 다음과 같다^[8].

$$\begin{aligned} \alpha_i^{p+1} &= \frac{1}{N} \sum_{j=1}^N p(i|x_j, \Phi^p) \\ \mu_i^{p+1} &= \frac{\sum_{j=1}^N p(i|x_j, \Phi^p) x_j}{\sum_{j=1}^N p(i|x_j, \Phi^p)} \\ \Sigma_i^{p+1} &= \frac{\sum_{j=1}^N p(i|x_j, \Phi^p) (x_j - \mu_i^p)(x_j - \mu_i^p)^T}{\sum_{j=1}^N p(i|x_j, \Phi^p)} \end{aligned} \quad (5)$$

여기서 N 은 화소의 전체 개수이며

$p(i|x, \Phi) = \frac{\alpha_i f_i(x|\theta_i)}{\sum_{k=1}^K \alpha_k f_k(x|\theta_k)}$ 로 Φ 가 주어졌을 때 i 번째 가우시안 함수에 대한 화소 x_j 가 적합될 확률이다.

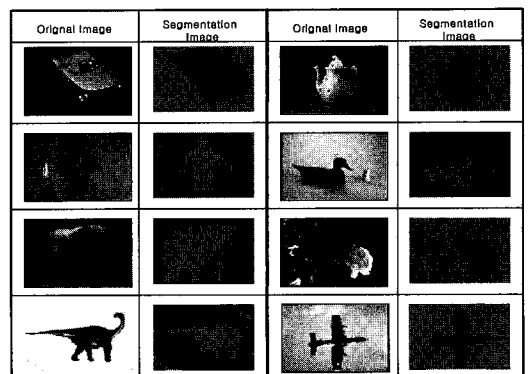


그림 2. GMM에 의한 영상의 객체 분할 결과

따라서 객체의 분할은 각 화소에 대한 $f(i, x_j, \phi)$ 를 계산하여 가장 높은 확률 밀도를 갖는 가우시안 함수 i 로 모델링 되어진 그룹에 화소를 할당함으로써 영상내의 동질의 특성을 갖는 객체들을 분할할 수 있다.

그림 2는 GMM을 적용하여 영상내의 객체가 분할된 결과를 보여주고 있는데 원래의 영상에 대하여 2개의 객체로 분할된 결과가 잘 나타나 있다.

III. 객체분할을 이용한 내용기반 영상검색

1. 제안된 SCBQ 시스템의 기본 구조

본 논문에서 제안된 SCBQ(Shape and Color Based Query) 시스템은 앞장에서 기술된 GMM 알고리즘을 이용하여 분할된 객체들로부터 모양과 색상 정보를 추출하여 검색을 하도록 설계되었는데 그림 3에 블록 다이어그램이 나타나 있다.

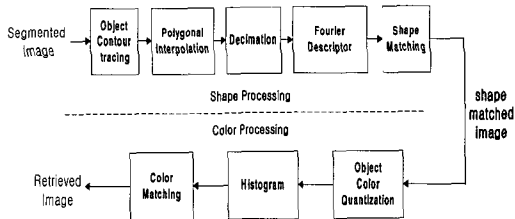


그림 3. 모양과 색상이기반 질의 시스템의 블록 다이어그램

검색과정은 크게 두 단계로 구성되어 있는데 제1 단계에서는 분할된 객체들에 Fourier descriptor를 적용하여 추출된 모양 특징을 이용하여 객체의 모양이 유사한 후보 영상을 선별하였으며 제2단계에서는 선별된 후보 영상들에 색상 특징을 이용한 유사성 정합을 하여 내용기반 영상검색을 수행하게 된다.

2. 객체의 윤곽점 추적

분할된 영상에 morphology operator를 적용하여 객체 주위의 잡음과 같은 고립점들을 제거한 후에 객체의 윤곽을 위치 함수로써 표현하기 위하여 객체의 외부 테두리를 4-connected neigh-borhood 방법을 사용하여 연결하고 각 점들의 좌표값을 복소수 함수로 나타내었다. 다음으로 영상내 객체의 윤곽점들을 일정한 개수로 정규화하기 위하여 윤곽점을 polygonal interpolation에 의하여 연속적인 함

수로 보간을 하였으며 연속된 함수들로부터 등간격의 N 개 윤곽점들로 샘플링 하였다. 그림 4는 객체의 외부 테두리와 영역의 위치 함수를 얻기 위한 절차를 보여주고 있는데 그림 (a)와 같은 분할된 객체로부터 윤곽점을 추출한 결과가 그림 (b)에 나타나 있으며 정규화를 위하여 윤곽점의 샘플 개수를 polygonal interpolation을 적용하여 연속적인 함수로 보간을 한 결과와 등간격에 의해 N 개의 윤곽점들로 샘플링한 결과가 그림 (c)와 (d)에 나타나 있다.

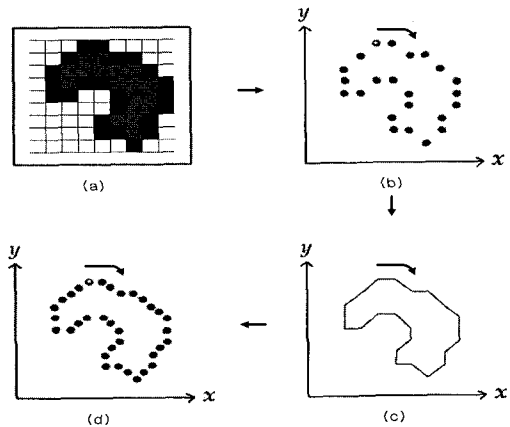


그림 4. 객체의 외부 테두리와 영역의 위치 함수를 얻기 위한 방법의 예
(a) 객체가 포함된 영상,
(b) 영역의 테두리,
(c) polygonal 보간,
(d) 균일화된 윤곽의 샘플

3. Fourier Descriptor에 의한 모양특징 추출

내용기반 영상검색에서 영상이 가지고 있는 속성을 이용하여 추출된 모양 특징은 색상이나 질감의 특징에 비해 가장 응용이 어려운 부분 중의 하나이다. 실제로 이러한 객체의 모양에 의한 유사성 검사는 어려움이 많이 있으므로 최근에는 통계적인 특성을 이용하여 효율성을 높일 수 있는 방법들을 연구하고 있다.

본 논문에서는 패턴인식(pattern recognition), 모양 근사화(shape approximation), 그래픽 애니메이션(graphic animation)에서 영상내의 객체의 특징 정보의 추출을 위하여 일반적으로 사용되는 Fourier descriptor를 사용하였다. Fourier descriptor는 영상에서 객체의 회전, 이동, 축척 등에 불변하다는 장점을 가지고 있어 효율적으로 객체의 모양 특징을

추출할 수가 있으며 식 (6)과 같이 DFT(Discrete Fourier Transform)에 의하여 Fourier descriptor를 계산할 수가 있다^[9].

$$Z_k = \frac{1}{N} \sum_{n=0}^{N-1} z_n e^{-j(2\pi/N)kn} \quad (6)$$

여기서 z_n 은 객체 윤곽점의 위치를 나타내며 N 은 전체 point의 수이다.

위와 같이 영상내 객체의 윤곽선으로부터 구하여진 Fourier descriptor로부터 객체의 회전, 이동, 축척에 불변한 모양 특징을 추출하기 위하여 본 논문에서는 Wallace와 Wintz가 제안한 Fourier descriptor 정규화 알고리즘을 사용하였다^[10]. 본 알고리즘은 객체의 위치에 대한 정규화를 위하여 먼저 Fourier descriptor의 파라미터 $Z_0 = 0$ 으로 하였으며, Z_1 의 크기와 위상을 이용하여 축척과 회전에 대한 정규화를 할 수 있다. 그림 5는 Fourier descriptor 정규화를 위한 전체적인 순서도를 보여 주고 있다.

4. 색상 특징 추출

색상은 영상들로부터 정보를 추출하는데 있어서 매우 중요한 특징으로 히스토그램을 이용한 내용기반 영상검색에 많이 이용하고 있다^[11]. 본 논문에서는 그림 6에서와 같이 R.G.B 세 개의 색상으로 이루어진 입력 영상의 각 화소를 인간의 시각 인지와 밀접한 관계가 있는 H.S.I 모델로 변환하여 객체들로부터 영상내 명도, 크기, 그리고 위치의 변화에 어느 정도 무관하게 적용 가능한 색상 정보를 추출하여 검색에 이용하였다.

H.S.I 모델에서 H는 색상의 분포에 따라 360°의 각도를 이루고 있으며 S는 H에 대한 색상의 세기, 즉 채도를 나타낸다. 채도 값이 1에 가까울수록 높은 순도의 색상 값을 가지며 I 값은 명도로 색상의 명암을 나타내고 I 값이 1에 가까울수록 밝은 색에 가까워진다.

질의 영상과 데이터베이스 영상의 유사성을 비교하기 위하여 색상 특징 정보는 H.S.I 색상 공간에 대하여 intensity 성분에 관계없이 hue는 6개, saturation은 2개의 그룹(bin)으로 분할하여서 총 12개 그룹으로 양자화 하였으며, 특히 Hue 성분은 인간의 시각체계에 더 민감하게 반응을 하기 때문에 세밀하게 분할하였다. GMM에 의하여 분할된 각각의 객체들로부터 색상 히스토그램을 추출하여서 공

간적 정보가 반영된 시각적 특징 벡터로 이용한다.

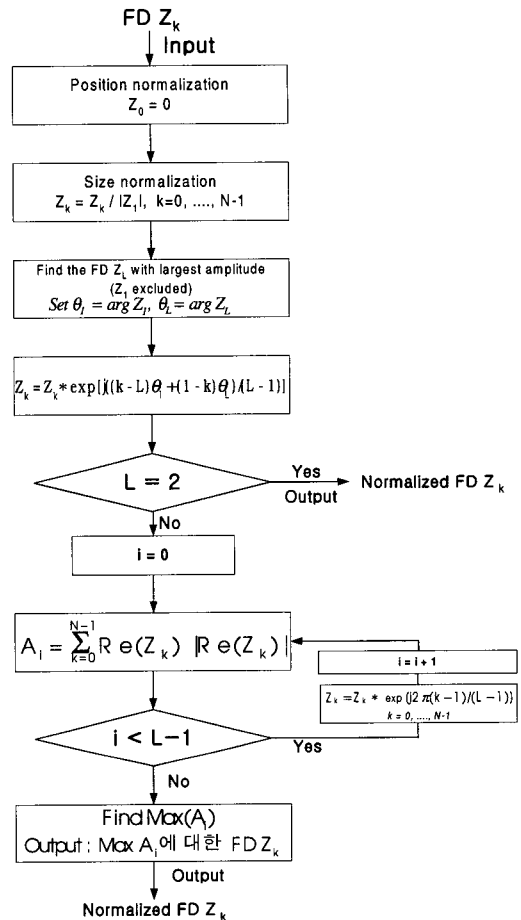


그림 5. Fourier descriptor 정규화 과정의 순서도

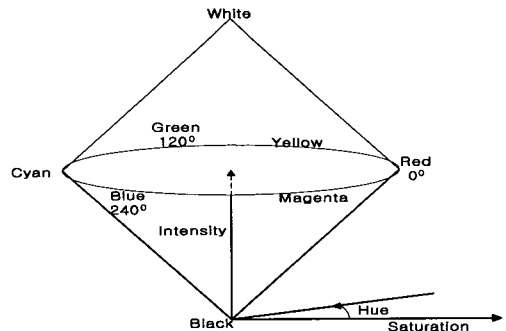


그림 6. H.S.I 색상 공간

5. 색상과 모양 특징의 유사성에 의한 정합
내용기반 영상검색은 그림 7에서와 같이 여러 개의 객체들로 구성되어 있는 영상으로부터 공간적인

모양 특징과 시각적인 특징을 추출한 후에 객체들의 모양에 따른 상대적인 관계를 고려하면서 질의 영상과 데이터베이스에 저장된 영상들 사이의 유사성을 비교하여 정합을 하게 된다. 즉, Fourier descriptor에 의해 생성된 특징 벡터를 이용한 객체의 모양 정합을 통하여 모양이 비슷한 객체들을 내포하는 영상을 선별하였으며 다음으로는 질의 영상과 데이터베이스에 있는 후보 영상에 대하여 공간상에서 상대적으로 모양에 따른 유사도가 높은 객체들끼리 색상 정합을 수행한다. 유사성은 질의 영상과 데이터베이스 영상들로부터 추출된 특징 벡터들 사이의 차(difference)에 의해 측정되는데 본 논문에서는 식 (7)과 같은 유클리디언 거리 측정 방법을 사용하며 정합 결과는 그 차이가 적을수록 영상들간의 유사성이 커지게 된다.

$$D_{qd} = \sqrt{(\bar{f}_q - \bar{f}_d)^2} = \sqrt{\sum_{i=1}^N (\lambda_i^q - \lambda_i^d)^2} \quad (7)$$

여기서 \bar{f}_q 는 질의 영상의 특징 벡터를 \bar{f}_d 는 데이터베이스 영상의 특징 벡터를 나타내고 N 은 정규화된 Fourier descriptor에 의해 추출한 모양 특징 벡터 및 객체의 색상 히스토그램에 해당하는 그룹의 개수이며, λ_i 는 i 번째에 해당하는 특징 벡터를 나타낸다.

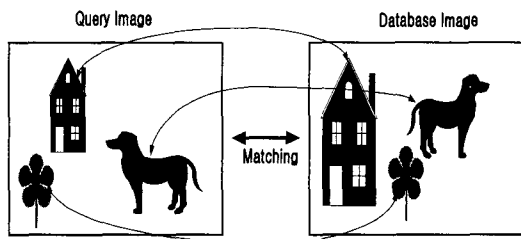


그림 7. 모양과 색상 특징에 의한 정합

IV. 실험 및 결과

GMM을 사용하여 분할된 객체들로부터 색상과 모양의 복합적인 특징을 이용한 내용기반 영상검색을 수행하는 본 알고리즘의 효율성을 실험하기 위해 영상내 객체들의 특징을 추출하여 질의 영상과 데이터베이스 영상 사이의 유사성에 의한 정합을 하였다. 실험을 하기 위한 시스템으로 Pentium PC를 사용하며 검색 알고리즘은 visual C++ 5.0을 이용하

여 구현하였다. 실험을 위해 사용한 자연 영상(natural image)은 각각 다른 구조와 색상으로 구성된 Corel Photo CD 영상 235장을 혼합하여 데이터베이스로 사용하였고 데이터베이스의 일부를 그림 8에 나타내었다.

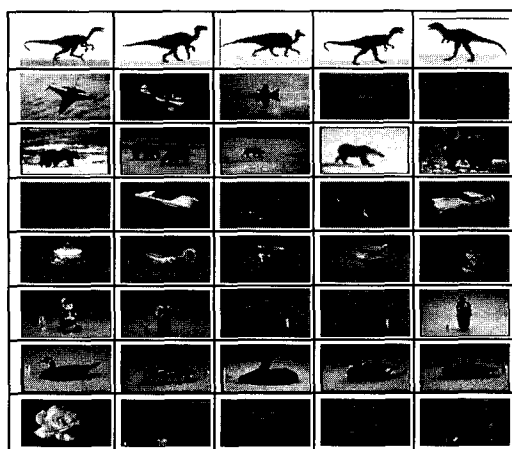


그림 8. 실험에 사용된 database의 일부

데이터베이스 영상은 120×80의 화소들로 이루어진 256 칼라 영상으로 dinosaur, airplane, car, china, bear, rose, doll, duck_doll의 총 8개 class로 구성되어 있으며, 한 개의 class는 30개의 유사한 영상들로 구성되어 있다. 단, rose class의 경우는 25개의 영상들로 이루어져 있다. 질의 영상은 class 내의 영상들로부터 랜덤하게 한 개씩 선택된 후에 데이터베이스에는 포함되지 않는다.

실험에 사용된 영상들은 대부분 1개의 객체로 구성되었기 때문에 GMM을 적용하여 객체와 배경 두 집단으로 분할하였으며 제안된 SCBQ 시스템의 동작 특성을 분석하기 위하여 다음과 같이 3종류의 실험을 수행하였다. 즉, 그림 3과 같이 제 1단계 객체의 모양 특징 정합을 수행한 후에 유사도가 높은 순서로 정렬(sorting)된 결과에 대하여 제 2단계 색상특징 정합을 직렬(cascade) 형태로 수행한 원래의 검색방법인 forward SCBQ, forward SCBQ의 역순(reverse order) 동작방법으로 먼저 색상 특징 정합을 수행하고 다음으로 모양 특징 정합을 직렬 형태로 수행하는 방법인 backward SCBQ, 그리고 가중치를 두어서 모양 특징 정합과 색상 특징 정합을 병렬적으로 동시에 정합하는 weighted SCBQ 등 3종류의 방법이다. 실질적으로 영상내에 객체들이 2개 이상일 경우에는 모양의 유사성에 따른 객체들

의 상대적인 정합을 먼저 수행하여 후보영상을 추출하기 때문에 원래의 방법인 forward SCBQ만을 적용할 수 있다. 또한 본 논문에서 제안된 방법의 효율성을 실험하기 위하여 Swain & Ballard가 제안한 histogram intersection 방법^[11]과 비교하였다. Histogram intersection 방법은 데이터베이스 영상의 히스토그램을 D_i 라 놓고 질의 영상의 히스토그램을 Q 라 하였을 때 각 색상 red, green, blue에 대하여 16개의 그룹으로 분할하여 전체적으로는 $N=4096$ 개의 그룹을 생성하였다. 그리고 다음 식 (8)과 같이 두 영상 사이의 그룹에 해당하는 히스토그램의 최소 값을 정규화 하여서 질의 영상과 데이터베이스 영상의 정합을 하였다. 질의 영상과 데이터베이스 영상의 $H(\cdot)$ 의 값이 1에 근접할수록 유사성이 높게 나타나며 1보다 작을수록 유사성은 떨어진다.

$$H(D, Q) = \frac{\sum_{i=1}^N \min(D_i, Q_i)}{\sum_{i=1}^N D_i} \quad (8)$$

그림 9, 10, 11은 기존의 Swain & Ballard가 제안한 histogram intersection 방법을 사용하여서 dinosaur, airplane, car 영상을 질의로 하여 유사성이 높은 순서대로 상위 20개의 검색된 결과를 보여 주고 있다. 이 방법은 현재의 내용기반 영상검색 시스템에서 가장 많이 응용하는 히스토그램에 대한 대표적인 방법으로 영상의 전체적인 색상 특징을 이용하기 때문에 객체와 배경이 비슷한 색상을 많이 가지고 있는 경우는 검색의 효율성이 우수하나 객체가 상대적으로 배경에 비하여 작은 영역으로 구성되어 있는 경우에는 영상내 객체가 가지고 있는 속성이 배경에 의하여 무시되기 때문에 객체의 특징 정보를 표현하기에는 효율적이지 못함을 볼 수가 있다.

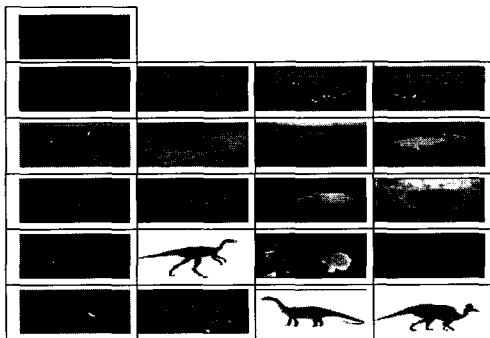


그림 9. Histogram intersection 방법에 의한 Dinosaur 영상의 질의 결과

특히 airplane의 경우 질의영상과 전체적인 색상이 유사한 china 그리고 doll 클래스의 영상들이 우선 순위에 나타나는 결과를 관찰할 수 있다.

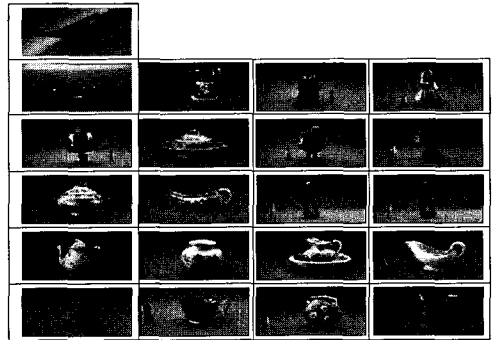


그림 10. Histogram intersection 방법에 의한 Airplane 영상의 질의 결과



그림 11. Histogram intersection 방법에 의한 Car 영상의 질의 결과

그림 12, 13, 14는 forward SCBQ 방법으로 1단계에서 Fourier descriptor를 이용하여 추출된 모양 특징을 질의 영상과 데이터베이스 영상에 적용하여 서로의 객체 모양이 가장 유사한 90개의 후보 영상을 선별한 후에 객체에 해당하는 색상 히스토그램을 계산하여 유사성이 높은 순서대로 상위 20개의 검색된 결과를 보여준다. 그림에서 볼 수 있듯이 객체의 모양 특징을 이용하여 선별된 후보 영상에 객체의 색상 특징을 고려함으로써 영상의 많은 부분을 차지하는 배경 색상과는 상관없이 질의 영상에 포함된 객체의 특성을 잘 반영하여 같은 종류의 많은 영상이 우선 순위가 높게 나타났다. 그러므로 모양과 색상의 복합적인 특징을 검색에 이용함으로써 목적에 맞는 영상을 효율적으로 필터링할 수 있어 검색에 있어서 효율적이며 능률적임을 관찰 할 수 있다.

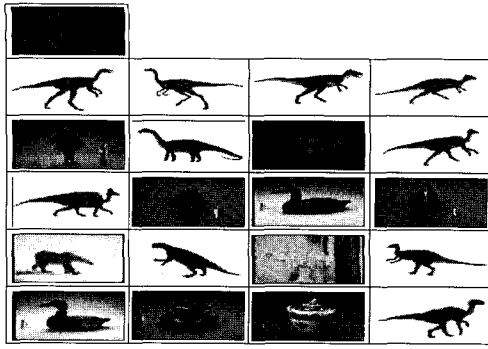


그림 12. Forward SCBQ에 의한 Dinosaur영상의 질의 결과

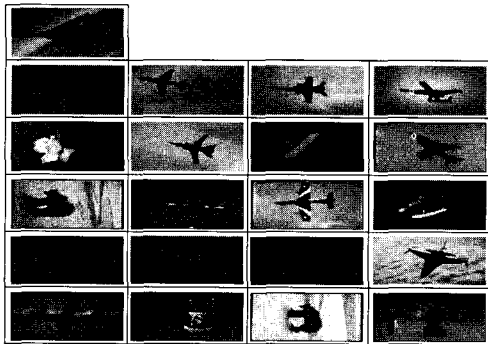


그림 13. Forward SCBQ에 의한 Airplane 영상의 질의 결과

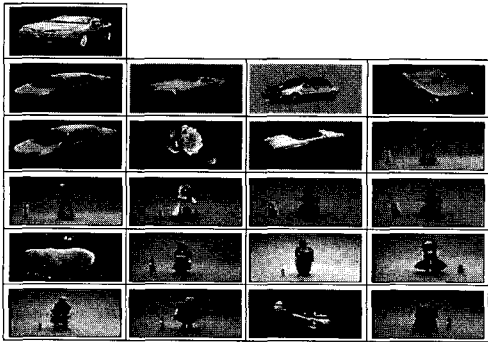


그림 14. Forward SCBQ에 의한 Car 영상의 질의 결과

표 1은 histogram intersection, forward SCBQ, weighted SCBQ 그리고 backward SCBQ의 4가지 질의방법을 사용하여 유사성에 의한 검색을 수행한 후에 class-confusion matrix 형태⁶⁾로 변환시킨 결과를 보여주고 있다. 왼쪽 column에 있는 영상들은 실험에 사용된 질의 영상들로서 데이터베이스에는 포함되지 않으며 각각의 검색을 수행한 후에 유사

성이 높은 30개의 영상을 데이터베이스로부터 추출 하여서 각 class에 소속된 영상들의 개수를 조사하는 실험 방법을 질의 영상을 바꾸어 가면서 5번 수행한 후에 평균값을 백분율로 환산한 결과가 오른쪽 column 들에 나타나 있다. 예를들어 forward SCBQ의 경우 dinosaur는 45%가 같은 class의 영상 들로부터 정확히 분류되었으며 그 외는 1%의 airplane, 24%의 china, 12%의 doll, 8%의 duck_doll, 2%의 rose 그리고 8%의 bear 클래스로부터 잘못 분류되었음을 나타내고 있다. weighted SCBQ의 경우 모양과 색상에 $w_s=0.8$, $w_c=0.2$ 의 가중치를 주어 검색을 수행한 결과이며 backward SCBQ에서는 1단계에서 객체의 색상 히스토그램에 의해 90개의 후보 영상을 압축하고 마지막으로 압축된 후보 영상으로부터 객체의 모양 정보에 의해 상위 30개의 영상을 검색하여 실험한 결과를 나타낸 것이다.

결국 분할된 객체들의 모양과 색상의 특징을 정합에 이용할 경우 평균적으로 45% 이상의 검색율을 나타냈으나 histogram intersection 방법처럼 전체 영상에 대한 특징을 사용한 경우 31% 정도로 검색되어 제안된 방법의 효율성이 입증되었으며 그 이유는 앞에서 기술한 바와 같이 객체들의 특징 정보를 독립적으로 사용하였기 때문이다.

표 1. Corel Photo images의 유사성에 의한 검색 결과

(a) Histogram Intersection(단위: %)

class	classification							
	a	b	c	d	e	f	g	h
a)Dinosaur	21	3	4	29	3	1	4	36
b)Airplane	2	12	15	16	8	11	8	28
c)China	3	3	56	15	3	1	2	17
d)Car	2	13	11	28	8	1	13	24
e)Doll	3	3	7	20	31	7	17	14
f)Duck_doll	9	1	22	1	4	38	14	12
g)Rose	2	14	1	41	4		35	1
h)Bear	6		23	3	9	26	8	26

(b) Forward SCBQ(단위: %)

class	classification							
	a	b	c	d	e	f	g	h
a)Dinosaur	45	1	24		12	8	2	8
b)Airplane	12	61	5		4		4	14
c)China	29		30	2	15	8	3	13
d)Car		6	1	49	36	1	4	3
e)Doll		8	1	9	67	5		9
f)Duck_doll	11	8	5		14	51	1	8
g)Rose	10	24	3		9	3	35	18
h)Bear	6	20	5	4	25	11	2	27

(c) Weighted SCBQ(ws=0.8,wc=0.2)(단위: %)

class	classification							
	a	b	c	d	e	f	g	h
a)Dinosaur	45	2	23		11	8	3	3
b)Airplane	13	52	7		6		7	16
c)China	31	2	28	2	9	9	5	13
d)Car	1	2	1	52	39	1	4	1
e)Doll		2	1	19	69	4		4
f)Duck_doll	11	11	6		9	42	1	9
g)Rose	10	17	8	4	5		31	18
h)Bear	9	23	12	3	19	8	2	23

(d) Backward SCBQ(단위: %)

class	classification							
	a	b	c	d	e	f	g	h
a)Dinosaur	40	2	25		13	11	3	7
b)Airplane	13	57	6		4	1	4	15
c)China	30	1	24	1	13	11	4	15
d)Car				62	39		1	
e)Doll		1	1	27	67	3		2
f)Duck_doll	13	5	5		18	52	2	4
g)Rose	10	23	2	3	10	1	37	14
h)Bear	7	23	7	3	25	7	2	27

그림 15는 Markus Stricker와 Markus Orengo가 제안한 방법^[12]을 응용하여 4가지 질의 방법의 성능을 평가한 결과로서 검색된 영상들로부터 질의 영상과 유사한 영상을 순서대로 정렬한 후에 다른 클래스로부터 검색된 첫 번째 영상을 false 영상으로 간주하였으며 이 false 영상의 순위가 낮을수록 더 효율성이 떨어진다고 판단 할 수가 있다. 전반적으로 제안된 질의 방법들의 경우 false 영상 순위가 평균적으로 5, 6번째 나타났으나 histogram intersection 방법은 1, 2번째 나타나고 있으며 따라서 false 영상의 검출측면에서도 본 제안된 방법들이 신뢰성 있게 동작할 수 있음을 알 수 있다.

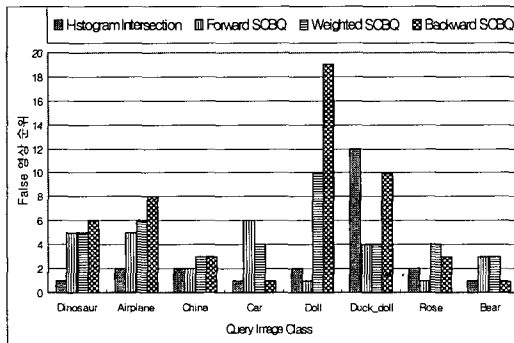


그림 15. Histogram intersection과 SCBQ 방법들의 false 영상 순위

V. 결론

본 논문에서는 영상을 구성하고 있는 객체들의 특성을 이용한 내용기반 영상검색 방법의 효율성을 증대시키기 위하여 영상내 여러 색상으로 이루어진 자연 영상에 GMM을 적용하여 객체들을 모델링하였으며 각 객체들의 데이터는 EM 알고리즘을 적용하여 추정된 통계적인 분포 특성에 따라서 동질의 영역을 포함하는 샘플들의 군으로 분할하였다. 분할된 객체들로부터 회전, 이동, 축척에 불변한 모양 특징 벡터를 추출하기 위하여 Fourier descriptor를 사용하였으며 색상 특징 정보는 Hue와 Saturation 성분을 그룹으로 나누어 특징을 추출하였다. 유사성 비교는 분할된 객체들로부터 특징 벡터를 추출하여 영상내 객체들의 공간상의 상대적인 관계를 고려하면서 객체가 가지고 있는 독립적인 특징 정보를 검색에 이용함으로써 기존의 영상 전체에 대한 특징 벡터를 추출하여 검색에 이용하는 방법보다 효율적이며 우수한 검색 결과를 보여주고 있다.

제안된 forward SCBQ, backward SCBQ와 weighted SCBQ, 그리고 기존의 histogram intersection 방법을 자연영상에 적용하여 실험한 결과 전반적으로 forward SCBQ 방법이 우수하였으며 앞으로 영상내에 여러 객체들이 존재할 경우 더욱 효율적으로 본 알고리즘이 동작할 수 있음을 예측할 수 있었다. 따라서 제안된 SCBQ 시스템은 비디오 데이터의 검색과 또한 현재 진행되고 있는 멀티미디어 정보의 효율적인 검색을 지원하기 위한 MPEG-7 등에 응용될 수 있을 것이다.

참고 문헌

- [1] M. Ibrahim Sezan and Richard J.Qia-n. MPEG-7 standardization Activities. *Proceedings ICIP98*, vol 3, Oct. 1998.
- [2] J. Ashley, R. Barber, M. Flickner, J. Hafner, D. Lee, W. Niblack, and D. Petkovic. Automatic and semiautomatic methods for image annotation and retrieval in QBIC. *Technical Report RJ 9951-87910, IBM Almaden Research Center*, April 1995.
- [3] H. Greenspan, S. Belongie, C. Carson and J. Malik. Recognition of images in large data-

bases using color and texture. *In CVPR '97*, 1997.

[4] John R. Smith and Shi-Fu Chang, Visual-SEEK: a fully automated content based image query system, *ACM Multimedia 96*, November 20, 1996.

[5] A. Pentland, R. Picard, and S. Sclaroff. Photo-book: Content-based manipulation of image databases. *Int. J. Comp. Vis.*, 18(3): 233-254, 1996.

[6] 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. *IEEE Computer Society*, pp. 675-682, Jan 1998.

[7] A. P. Dempster, N. M. Laird, and D. J. Rubin, "Maximum likelihood from incomplete data via the EM algorithm, *J. Royal Soc. Statist. Series B*, no. 1, pp. 1-38, 1977.

[8] Patricia McKenzie, Michael Alder. Initializing the EM Algorithm for use in Gaussian Mixture Modelling. *Pattern Recognition in Practice IV, Vlieland*, 1994.

[9] F. Marques, B. Llorens, and A. Gasull, Prediction of Image Partitions Using Fourier Descriptors: Application to Segmentation-Based Coding Schemes, *IEEE Trans. Image Processing*, Vol.7, No.4, April 1998.

[10] Ferdinand van der Heijden, *Image Based Measurement Systems*. WILEY.

[11] Babu M. Mehtre, Mohan S. Kankanhalli, A Desai Narasimhalu, Guo Chang Man. Color matching for image Retrieval. *Pattern Recognition Letters*. V16. No.3, pp. 325-331, Mar. 1995.

[12] Markus Stricker and Markus Orengo, "Similarity of Color Images", *SPIE 95, San Jose*.

박 종 현(Jong-Hyun Park)

정회원

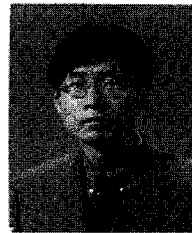


1995년 2월 : 목포대학교 전자공학과 학사
1997년 8월 : 목포대학교 전자공학과 석사
1998년 3월~현재 : 목포대학교 전자공학과 박사과정

<주관심 분야> 영상신호처리, 멀티미디어 정보검색, 영상압축

박 순 영(Soon-Young Park)

정회원



1982년 2월 : 연세대학교 전자공학과 학사
1984년 2월 : 연세대학교 전자공학과 석사
1989년 1월 : State Univ. of New York at Buffalo, 전기 및 컴퓨터공학과 박사

1989년~1990년 : State Univ. of New York at Buffalo, 전기및컴퓨터공학과, Post Doctor

1990년~현재 : 목포대학교 전자공학과 부교수

<주관심 분야> 디지털신호 및 영상처리, 멀티미디어 정보검색, 컴퓨터 비전

오 일 환(Il-Whan Oh)

정회원



1982년 2월 : 연세대학교 전기공학과 학사
1985년 2월 : Univ. of Texas, 전기전자공학과 석사
1988년 1월 : Univ. of Rhode Island, 전기전자공학과 박사

1990년~현재 : 목포대학교 전자공학과 부교수

<주관심 분야> 광신호처리, 광통신