

컬러 인접성과 클러스터링 기법을 이용한 객체 기반 영상 검색

이 형 진^{*} · 박 기 태^{**} · 문 영 식^{***}

요 약

본 논문은 컬러 인접성과 클러스터링 기법을 이용한 객체 기반 영상 검색 기법을 제안한다. 컬러 인접성이란 영상내의 서로 이웃한 영역에서 나타나는 컬러의 특징값을 말하고, 영상 데이터베이스로부터 사용자가 찾고자하는 영역과 유사한 후보 영역들을 우선 추출하는데 사용된다. 또한 클러스터링 기법은 후보 영역들 가운데 객체가 존재하는 영역만을 추출하는데 사용되고, 질의 영상과 데이터베이스 영상 사이의 유사도 측정을 위하여 히스토그램 인터섹션(histogram intersection) 방법이 사용된다. 제안하는 방법에서 사용되는 영상의 컬러쌍 정보는 객체의 이동, 회전, 그리고 크기 변화에 강건하며, 실험을 통하여 제안하는 방법이 기존의 방법보다 우수함을 확인하였다.

Object-Based Image Retrieval Using Color Adjacency and Clustering Method

Hyung-Jin Lee^{*} · Ki-Tae Park^{**} · Young-Shik Moon^{***}

ABSTRACT

This paper proposes an object-based image retrieval scheme using color adjacency and clustering method. Color adjacency features in boundary regions are utilized to extract candidate blocks of interest from image database and a clustering method is used to extract the regions of interest(ROI) from candidate blocks of interest. To measure the similarity between the query and database images, the histogram intersection technique is used. The color pair information used in the proposed method is robust against translation, rotation, and scaling. Consequently, experimental results have shown that the proposed scheme is superior to existing methods in terms of ANMRR.

키워드 : 컬러 인접성(color adjacency), 객체 추출(object extraction), 영상 검색(image retrieval), 클러스터링(clustering), 히스토그램 인터섹션(histogram intersection)

1. 서 론

최근 컴퓨터와 통신 기술의 발달로 인해 다양한 멀티미디어 데이터가 증가하고 있고, 디지털 방송 관리, 주문형 비디오, 디지털 라이브러리, 의료 진단 시스템, 홈쇼핑 등의 다양한 분야에서 활용되고 있다. 하지만 다양한 데이터와 포맷, 큰 용량, 그리고 의미 있는 데이터의 자동 추출 등은 기존의 데이터베이스 시스템으로는 한계가 있다. 따라서 이러한 멀티미디어 정보들의 다양한 활용을 위해서는 데이터를 효과적으로 저장하고 검색, 색인 하는 기술이 선행되어야 하며, 이에 관한 다양한 방법들이 활발하게 연구되고 있다. 특히, 영상 혹은 비디오 데이터베이스로부터 원하는 영상을 찾고자 하는 영상 검색(Image Retrieval) 기법은 그 중 가장 활발히 연구되고 있는 분야중 하나이

다[1].

영상 검색은 크게 텍스트 기반 영상 검색(Text-based Image Retrieval)과 내용 기반 영상 검색(Content-based Image Retrieval)으로 분류된다. 텍스트 기반 영상 검색은 주석이나 키워드에 의존하는 검색 방법으로 주석 자체의 데이터양이 작아 저장 및 관리가 용이하고 고속 검색이 가능한 장점이 있다. 그러나 수작업에 의해 영상의 내용을 주석으로 표현해야하는 번거로움이 있고 표현하는 사람의 관점에 따라 영상의 내용이 다르게 표현될 수 있어 정보의 일관성과 객관성이 보장되지 않는다. 이에 반해 내용 기반 영상 검색은 영상의 컬러(Color), 모양(Shape), 질감(Texture) 등의 내용적인 특징이나 속성을 이용하여 영상을 표현하고 이것을 기반으로 검색을 수행한다. 그러므로 영상의 시각적인 특성을 기반으로 데이터베이스를 손쉽게 구축할 수 있고, 효율적인 관리 및 검색이 가능하므로 현재 많은 시스템에서 내용 기반 영상 검색을 사용한다.

최근 영상의 객체 정보에 대한 관심이 증가하고 있으며,

* 이 논문은 한국과학재단 목적기초연구 지원으로 수행되었음.

* 교신저자 : ysmoon@cse.hanyang.ac.kr

† 정 회 원 : 모비넷피아(주)

** 준 회 원 : 한양대학교 대학원 컴퓨터공학과

*** 정 회 원 : 한양대학교 컴퓨터공학과 교수

논문접수 : 2004년 9월 17일, 심사완료 : 2005년 2월 2일

이를 이용한 다양한 서비스를 제공하기 위한 연구가 진행 중이다. 객체 정보를 이용한 서비스의 예로서, 차세대 대화형 TV는 객체 단위로 영상의 내용을 색인하고, 이 정보를 이용하여 영상에 포함된 상품, 인물 등의 부가적인 객체 정보를 이용한 서비스를 제공 할 수 있다. 이러한 객체 추출 및 검색에 관한 연구는 영상내의 에지(Edge), 모양(Shape)과 같은 기하학적인 정보를 이용하는 방법과 영상내의 국소적인 컬러(Color) 정보를 이용하여 복잡한 배경에서 관심 객체를 추출하는 방법 등이 있다.

본 논문은 사용자가 질의한 객체의 색상 정보를 이용하여 대용량 영상 데이터베이스로부터 질의 객체를 포함하고 있는 영상을 빠른 속도로 검색하고, 검색된 영상에서 관심 객체의 위치를 찾는 방법에 대해서 제안한다. 대부분의 내용 기반 영상 검색 시스템들은 전체 영상을 검색하는 것에 초점이 맞추어져 있기 때문에, 때때로 관심 객체를 포함하는 배경 색상의 영향으로 잘못된 검색 결과를 보여주기도 한다. 그러므로 우리는 이러한 단점을 보완하기 위하여 공간 정보를 포함하는 컬러 인접성을 이용하여 객체 영역을 추출하고 검색을 수행하는 기법을 제안한다. 사용자는 영상 내에 존재하는 특정한 객체에 더 관심을 가진다는 가정을 하고, 비교되는 데이터베이스 영상 내에 주목하는 객체가 존재하면 우선적으로 유사성이 큰 것으로 판단하고, 존재하지 않으면 상대적으로 유사성이 낮은 것으로 판단한다. 이러한 것은 사용자가 주목하는 객체에 보다 민감하게 작용하는 검색을 수행하며, 사용자가 얻고자 하는 영상이나 영상내의 객체에 대하여 우선적으로 검색할 목적으로 사용이 가능하다. 유사도 측정에는 배경을 포함하는 관심 객체를 가진 질의 영상과 데이터베이스로부터 추출된 객체 영역에 대하여 수정된 히스토그램 인터섹션 기법을 사용한다.

본 논문은 2장에서 공간 정보를 나타내는 컬러쌍 추출(Color-pair Extraction) 알고리즘과 클러스터링(Clustering) 방법에 대하여 설명하고, 3장에서는 제안하는 방법의 실험 결과를 비교 분석하고, 4장에서 결론 및 향후 과제에 대하여 기술한다.

2. 제안하는 알고리즘

2.1 기존의 연구

일반적으로, 영상 검색 기법의 구조는 영상 입력 과정, 질의 과정, 그리고 검색 과정으로 나눌 수 있다. 영상의 입력 과정은 입력된 영상을 가장 잘 표현할 수 있는 특징(Feature)을 추출해서 영상 데이터베이스 내에 저장할 하는 단계이고, 영상의 질의 과정은 사용자에게 의해 찾고자 하는 영상의 입력이 주어지는 과정이다. 여기서는 입력 과정과 마찬가지로 주어진 영상의 특징을 추출하여, 데이터베이스

내의 특징과 비교함으로써 검색을 수행하고 원하는 영상을 출력으로 나타낸다.

이러한 내용 기반 검색 시스템은 각 영상이나 객체를 기술하는데 사용되는 특징의 집합을 어떻게 선택하느냐가 중요한 관건이라고 할 수 있으며, 특징 값들은 크게 컬러, 질감 그리고 모양 정보를 사용한다. 그 중 컬러 특징은 영상내에서의 이동, 회전, 크기에 불변하다는 장점을 가지고 있기 때문에 컬러를 기반으로 하는 영상 검색 기법에 대한 연구가 많이 이루어지고 있다.

기존의 대표적인 내용 기반 영상 검색 시스템으로는 QBIC, Chabot, FOCUS, C-BIRD, Blobworld, MARS, SEDL, PicToSeek 등이 있다. QBIC은 1995년에 IBM사에서 발표하였고, 데이터베이스, Query Interface, Match Engine으로 구성된다. 데이터베이스와 Query Interface는 특징벡터를 추출하여 색인과 질의를 수행하며, 색상, 질감, 모양, 다중 객체, 스케치, 위치, 사용자 정의 색상, 모양 등 다양한 특징값으로 구성된 특징벡터를 R*트리 구조로 색인한다. QBIC에서 사용되는 특징벡터는 20차원 정도로 많은 양인데, 효과적으로 색인하기 위해서 Karhunen-Loève(K-L) Transform을 이용하여 특징벡터의 주축(Principle Axes)을 색인하며, Match Engine에서는 사용자의 질의 영상과 색인된 데이터베이스 영상의 유사도 순으로 보여준다. C-BIRD(Content-based Image Retrieval in Digital Libraries)는 영상 검색의 접근 방법으로 조명 불변 컬러 모델(Illumination Invariant Color Model)을 사용하며, 컬러를 각 채널별로 정규화 시켜서 사용한다. 검색 방법은 객체 모델을 이용한 검색과 영상 분할을 통해 얻어지는 특징점의 위치를 이용한 검색을 제공하며, 컬러 정보 색인은 객체를 이루는 컬러 영역의 중심점을 노드(Vertex)로 모델링하고 이들의 인접한 영역의 노드를 연결(Edge)하여 작성한다. 질감 정보 색인은 영상내의 에지 정보를 이용하고, NTSC방식의 밝기신호(Luminance)로부터 추출하며, 경계선 지도(Edge Map)와 기울기 방향(Gradient Direction)을 이용한다. 모양 정보 색인은 Generalized Hough Transform 으로 추출해 색인하고, 컬러 정보에 대한 유사도 측정은 Swain과 Ballard가 제안한 컬러 히스토그램 인터섹션(Color Histogram Intersection)을 사용한다. Blobworld는 픽셀들의 컬러 및 질감이 유사한 부분의 집합을 'blob'이라고 정의하고 영상을 blob들의 집합으로 단순화 시켜 이들의 컬러 분포와 질감 특징에 따라서 검색하는 방법이다. Blob 단위로의 영상 분할은 $L*a*b*$ 컬러 모델에서 영상이 가우시안 믹스처 모델(Gaussian Mixture Model)을 따른다고 가정하고, EM(Expectation Minimization) 알고리즘을 이용해 수행한다. 이 방법은 영상을 분할하여, blob단위로 작성하는데, 많은 시간이 걸리므로 실시간 처리에는 적합하지 않다. MARS(Multimedia Analysis and

Retrieval System)는 컬러, 질감 모양에 대한 검색과 특징값의 배치에 대한 검색이 가능하다. 특징값 배치에 대한 검색을 위해 HSV 컬러모델에서 2D 색상 히스토그램과 거침, 대조, 방향성(CCD : Coarseness, Contrast, Directionality), 그리고 모양에 대한 특징(MFD : Modified Fourier Descriptor)을 조합하여 배치에 대한 특징값으로 사용한다. 검색 방법은 Fuzzy Boolean 검색 방법을 사용한다. PicToSeek은 컬러 분포, 채도, 농도의 변화, 배경 등을 특징값으로 색인하여 사용한다. 물체의 반사도를 고려해 조명에 불변한 컬러와 질감, 모양 정보를 이용하여 영상 검색을 한다. 유사도 측정은 히스토그램 Cross-Correlation에서 좋은 결과를 보이고 있다.

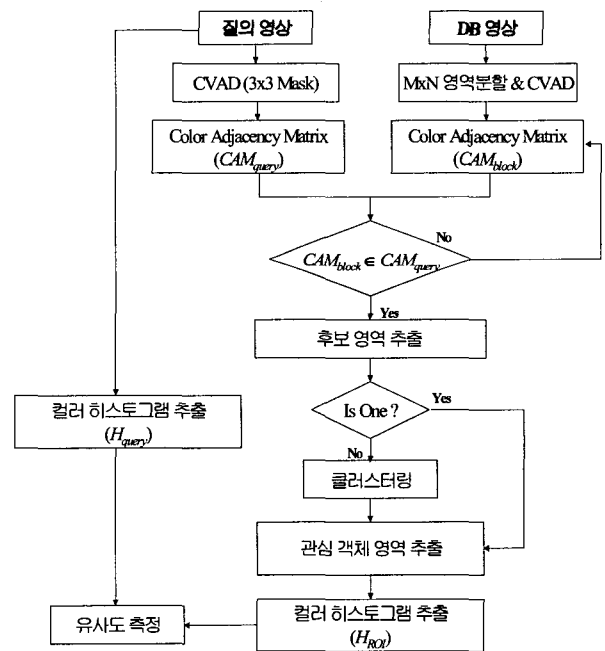
이러한 각각의 시스템들은 나름대로의 장점이 있지만, 몇 가지의 단점을 가지고 있다. QBIC의 경우, 사용자가 다양한 형태의 특징 값들을 사용할 수 있지만, 영상 분할에 사용자의 수작업이 필요하다는 단점을 가지고 있다. Netra는 분할된 영상을 이용하여 영상을 검색하지만, 실제적으로 분할될 영역 수를 사용자가 임의로 넣어 주어야 한다는 단점이 있다. VisualSEEk은 영상 내의 객체 또는 영역의 공간 정보를 이용한 검색에 적합하지만, 컬러 이외에 중요한 질감이나 모양 정보들을 사용하지 않고 있다. 마지막으로 Blobworld는 컬러와 질감, 위치 정보를 결합하여 좋은 영상 분할 결과를 보여주지만, 한 영상을 분할하는데 너무 많은 시간(약5분)이 걸리는 단점이 있다.

위에서 언급한 방법들은 영상 전체의 특징값을 추출해서 유사도를 구하여 검색하는 시스템들이다. 하지만 이 시스템들은 관심 있는 실제 영역(찾고자 하는 객체)을 찾기 위해서 객체의 배경까지 포함하게 되므로 검색 시 많은 오류 결과가 나타난다. 이러한 단점을 보완하기 위하여 Das는 객체(관심 영역)를 포함하는 부영상(subimage)을 질의 영상으로 만들고 데이터베이스 영상에 대해서 객체만을 추출하여 유사도를 구하는 영상 검색 시스템을 제안하였다[2]. 이 시스템은 두 가지 단계로 나뉘어져 있는데, 첫 번째 단계는 Peak Detection 단계로써 데이터베이스 영상에서 질의로 주어진 객체가 후보 영역을 찾아내는 단계이고, 두 번째 단계는 질의 영상과 데이터베이스 영상에서 찾아낸 후보 영역간의 유사도를 구하는 단계이다. 또한 컬러들의 공간적인 관계를 기술하기 위해서 Spatial Proximity Graph(SPG)를 구성하여, 각 영역들의 컬러 피크(Color Peak)는 하나의 노드로 구성하고, 예지는 두 컬러 피크가 인접하고 있으면 결합으로써 2차원 행렬을 구성한다. 마지막으로 질의 영상에서의 피크들과 데이터베이스 영상들에 대해서 찾아진 후보 영역들에 대한 피크들을 L1 Distance를 사용하여 유사도를 구하고, Graph Intersection 기법을 사용하여 매칭(Matching) 한다.

2.2 제안하는 방법

본 논문은 질의 영상과 데이터베이스로부터 추출된 관심 영역을 비교하는 객체 기반 영상 검색을 제안한다. 제안하는 방법은 (그림 1)과 같이 크게 4개의 단계로 구성되어 있다.

첫 번째 단계에서, 질의 영상을 위한 Color Adjacency Matrix(CAM_{query})는 컬러 벡터 각 거리(Color Vector Angular Distance : CVAD)를 이용하여 구성한다. 또한 데이터베이스 내의 영상들은 M x N 개의 영역으로 분할하고 질의 영상에서와 마찬가지로 CVAD를 이용하여 각 영역별 CAM_{block} 을 구성한다. 두 번째 단계에서, CAM_{query} 에 나타나는 컬러 쌍이 CAM_{block} 에도 존재하며 컬러쌍의 개수가 일정치 이상이면, 데이터베이스 영상의 현재 영역은 후보 영역이라고 판단하여 추출하게 된다. 세 번째 단계에서, 만약에 데이터베이스 영상에 다수의 후보 영역들이 존재하면, 클러스터링 알고리즘을 적용하여 최종 영역을 추출하게 된다. 마지막으로, 질의 영상과 데이터베이스 영상에서 추출된 최종 관심 영역의 유사도 측정은 수정된 컬러 히스토그램 인터섹션 기법(Modified Color Histogram Intersection)을 사용하여 구하게 된다.



(그림 1) 제안하는 기법의 순서도

가. Color Adjacency Matrix(CAM)

일반적으로 객체는 몇 개의 평탄한 컬러 영역을 가지고 있고 이러한 평탄 컬러 영역들은 서로 이웃하고 있으므로, 이들 평탄 영역 사이에서 에지(Edge)에 대한 정보, 즉 컬러쌍(Color-pair)에 대한 정보를 추출할 수 있다. 또한 이러한 것들은 관심 객체에 대한 공간적인 정보로 이

용될 수 있다.

컬러쌍을 추출하기 위하여 우리는 각각의 화소에 3x3 마스크를 적용한다. 마스크 내의 중심 화소와 나머지 8개의 이웃 화소 사이의 색상 거리(Color Distance)가 최대가 되는 화소를 컬러쌍으로 정의하고, 색상 거리가 일정치 이상이 되었을 때 에지 부분이라고 정의한다. Color Vector Angular Distance(CVAD)는 RGB 색 공간에서 컬러 벡터의 좌표에 따르는 색도에 기반한 각(Angle) 측정 방법으로 명암도의 변화에 강건한 특성을 가진다. 따라서 본 논문에서의 컬러쌍은 Androutsos가 제안한, 벡터의 각과 거리를 모두 고려한 CVAD 측정 방법을 사용하여 구한다[3]. 식 (1)은 CVAD 계산 방법을 나타낸다.

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

각도 크기

\vec{x}_i 와 \vec{x}_j 는 중심화소와 이웃화소의 3차 컬러 벡터를 나타내고, 각 부분에서 $2/\pi$ 의 정규화 인자는 최대각이 $\pi/2$ 이내에서만 존재한다는 것을 나타낸다. 또한 정규화 인자는 최대 거리 각이 (255, 255, 255)이라는 것이고 벡터의 거리는 $\sqrt{3 \cdot 255^2}$ 이다.

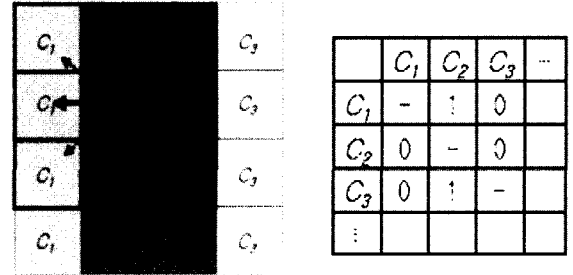
다음으로, 이렇게 구하여진 컬러쌍은 Color Adjacent Matrix(CAM)를 구성하는데 사용되며, 컬러쌍은 에지 부분의 정보를 나타내는데 사용되므로 모든 컬러쌍을 고려하는 것이 아니라 컬러쌍의 거리 값이 일정치 이상이 되는 것만을 선택하여 CAM을 구성한다. 3x3 마스크를 모두 적용하고 난 후 구성된 CAM은 식 (2)에 의하여 각 컬러쌍의 개수가 일정치 이상이 되는 것에는 1을 그렇지 않은 것에는 0을 할당하여 단순화 시킨 후, 객체의 후보 영역 추출을 위한 특징 값으로 사용한다.

$$CAM(C_i, C_j) = \begin{cases} 1 & \text{if } N(C_i, C_j) \geq Th \\ 0 & \text{other} \end{cases} \quad (2)$$

$N(C_i, C_j)$ 은 C_i 와 C_j 의 컬러쌍을 가진 화소수를 나타내고 Th 는 문턱치 값이다. (그림 2)는 컬러쌍 추출과 식(2)에 의해 CAM이 구성되는 것을 보여준다. (그림 2)에서 3X3 마스크의 중심 화소와 나머지 8개의 이웃 화소들 사이의 CVAD를 각각 구하고 난 다음 CVAD 측정값들 중에서 가장 큰 값이 문턱치 이상일 때 그 쌍을 중심화소의 컬러쌍으로 정의한다.

같은 방법으로 모든 화소에 대해서 적용하면 전체 컬러쌍 정보를 추출할 수 있다. 그러나, 본 논문에서는 추출된

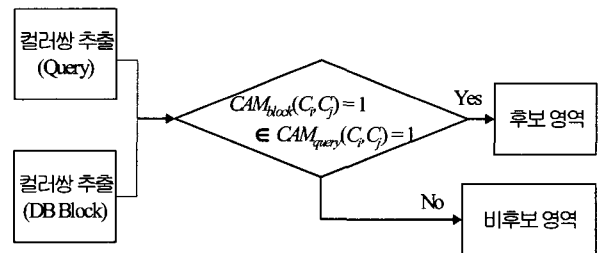
모든 컬러쌍을 이용하는 것이 아니라, 에지 정보를 표현할 수 있는 C_i 와 C_j 의 컬러쌍의 화소수가 일정치 이상인 컬러쌍들만을 선별해서 Color Adjacency Matrix로 구성하게 되는데, 컬러쌍의 화소의 개수를 이용하지 않고 1로 단순화하여 에지 정보를 가진 컬러쌍으로 표현한다.



(a) 컬러쌍 추출 (b) Color Adjacency Matrix
(그림 2) 컬러쌍과 CAM의 추출

나. 후보 영역 추출

제안하는 영상 검색 방법은 질의 영상과 데이터베이스 영상에서 관심 객체가 있을 만한 영역을 추출하여 둘 사이의 유사도를 측정한다. 이것은 기존의 연구에서 많이 사용되는 영상 전체에 대한 유사도를 측정하는 것이 아니라 질의 영상의 객체와 데이터베이스 영상의 객체를 추출하여 유사도를 측정하는 것으로, 우선 임의의 데이터베이스 영상 내에 질의로 주어진 객체 정보가 포함되어 있는지를 알아야 한다.



(그림 3) 후보 영역 추출

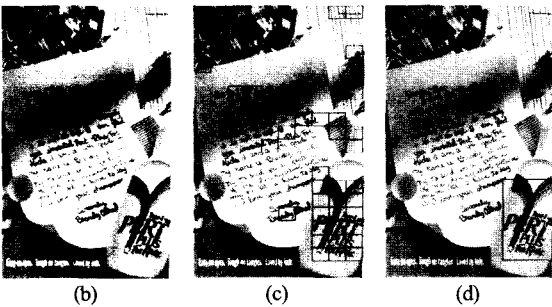
질의 영상으로부터 구성된 CAM_{query} 은 데이터베이스 영상에서 후보 영역을 결정하기 위하여 사용된다. 먼저, 임의의 데이터베이스 영상에 대하여 M x N개의 영역들로 겹치지 않게 분할하고, 각각의 영역들에 대해서 3x3 마스크를 적용하여 질의 영상에서의 컬러쌍 추출과 마찬가지로 CVAD를 측정, 문턱치보다 큰 값들에 대해서만 컬러쌍 정보를 추출한다. 이렇게 구하여진 컬러쌍 정보를 이용하여 각 영역에 해당하는 CAM_{block} 을 구성한다. 만약 CAM_{block} 에서의 컬러쌍들이 CAM_{query} 에도 존재하면, 그 영역은 일단 후보 영역으로 고려하게 되고 모든 영역에 대하여 수행하고 나면,

하나의 데이터베이스 영상은 관심 객체가 있을 만한 영역(후보 영역)과 그렇지 않은 영역으로 나눌 수 있게 된다. 이처럼 컬러쌍이 존재하는 영역들만 남겨두고, 컬러쌍이 존재하지 않는 영역들은 제거함으로써 질의 영상에 대한 객체가 존재할 수 있는 후보 영역을 추출한다. (그림 3)은 데이터베이스 영상으로부터 후보 영역이 어떻게 추출되어지는가를 보여주고 있다.

다. 클러스터링 알고리즘

후보 영역 구성 시, 모든 후보 영역들이 관심 객체를 포함하고 있는 것이 아니기 때문에, 객체가 존재하는 영역들만을 선택하도록 하는 과정이 필요하게 된다. 따라서 이러한 후보 영역들 중 이웃하는 영역들은 그룹화 될 수 있도록 영역들 사이의 클러스터링(Clustering) 과정이 필요하게 된다. 본 논문에서는 클러스터링을 위하여 후보 영역의 상하좌우 4방향에 인접한 다른 후보 영역이 존재 할 경우, 하나의 레이블로 할당하여 클러스터링하는 4-Neighbor를 고려한 Labeling 방법을 이용한다[4]. 결과적으로 클러스터링 후 구하여진 여러 개의 클러스터는 각 클러스터마다 하나 이상의 후보 영역들을 가지고 있게 된다. 따라서 그들 중 최대 컬러쌍의 개수를 가지는 후보 영역을 포함하는 클러스터를 관심 객체가 존재할 만한 최종 영역으로 판단하여 질의 영상과 유사도 매칭을 하게 된다. (그림 4)는 이러한 후보 영역들과 객체 영역의 추출 결과를 보여준다.

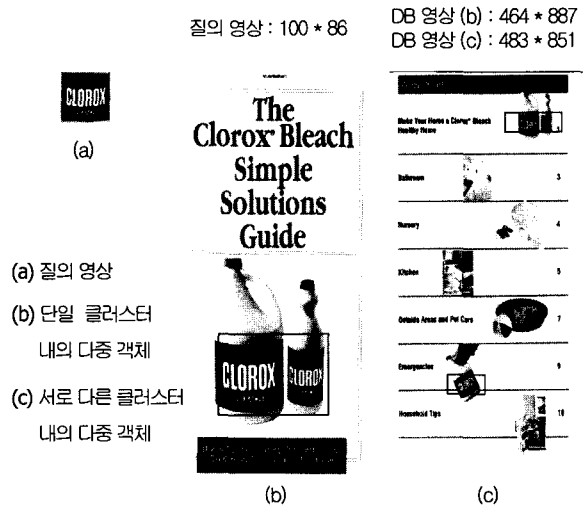
- (a) 질의 영상
- (b) DB 영상 질의 영상 : 172 * 175
- (c) 후보 영역들 DB 영상 : 727 * 1077
- (d) 클러스터링 후 최종 관심 영역



(그림 4) 후보 영역들과 객체 영역의 추출 결과

하지만 영상 내에 찾고자 하는 관심 객체와 유사한 배경 컬러가 널리 퍼져있는 경우, 최종 관심 영역으로 추출된 해당 클러스터가 실제 객체의 크기보다 훨씬 크게 구성되는 경우가 있게 된다. 이러한 경우는 최종적으로 질의 영상과 데이터베이스 영상 사이의 유사도 측정 시 그 값이 떨어지는 요인이 되므로, 클러스터내의 불필요한 영역들을

제거하는 분할(Split) 과정이 필요하게 되고, 이렇게 함으로 클러스터가 커지는 것을 방지할 수 있게 된다. 이러한 클러스터 분할 과정은 후보 영역의 위치(x, y)를 고려한 정렬 과정을 거치게 된다. 관심 객체가 속한 클러스터는 최대 컬러쌍의 개수를 가지는 영역을 중심으로 후보 영역들이 상하좌우로 퍼져 있으므로, 단순히 클러스터 내의 각 영역의 위치인 (x+y)의 값으로 정렬하는 것 보다 (x²+y²) 값으로 정렬하는 것이 더 바람직하게 된다. 이렇게 해당 클러스터를 정렬하고 난 후, 클러스터 내의 후보 영역들이 가지는 컬러쌍 개수의 최대, 최소를 고려하여 문턱치(Threshold) 이하인 영역을 기준으로 클러스터를 분리하게 된다. 또한 클러스터 내의 후보 영역 개수와 질의 영상의 후보 영역 개수도 함께 고려하는 방법으로 클러스터 분할 과정을 하게 된다.



(그림 5) 다중 객체 추출에 대한 그림 예

또한 영상에 존재하는 다중 객체도 고려하여야 한다. 다중 객체는 다음과 같은 두 가지 경우가 존재하는데, 첫번째는 다중 객체가 하나의 클러스터 내에 존재하는 것으로 그럴 경우 다중 객체를 추출하는데 문제가 없게 된다. 두번째는 다중 객체가 서로 다른 클러스터에 존재할 경우로, 이를 위하여 데이터베이스 영상에서 최대 컬러쌍 개수를 (CP_{main}) 가지는 영역을 추출한 후, 추출된 영역을 제외한 다른 영역에서도 최대 컬러쌍 개수를(CP_{other}) 가지는 영역을 추출한다. 다중 객체를 추출하기 위하여 CP_{main}과 CP_{other}의 차이가 10% 이내에 존재 할 경우 CP_{other}을 포함하는 영역을 또 다른 객체가 존재하는 영역으로 판단하여 다중 객체를 추출한다. (그림 5)는 이러한 다중 객체의 추출 결과를 보여준다.

라. 유사도 측정

본 논문에서 히스토그램 인터섹션(Histogram Intersection)

기법은 질의 영상과 데이터베이스로부터 추출된 객체 영역 사이의 유사도를 측정하는데 이용된다. Swain과 Ballard가 이러한 히스토그램 인터섹션 기법을 제안했다 [5]. 하지만 기존의 히스토그램 인터섹션 기법은 영상의 전체적인 스케일이 서로 다른 경우에는 같은 영상이라도 유사도가 낮아지는 문제점이 있다. 예를 들면, 같은 그림이 2배로 커지면 면적은 4배로 증가되므로 유사도가 낮아지게 된다. 따라서 본 논문에서는 찾고자하는 객체의 크기가 질의 영상의 크기와 다를 경우를 고려하여, 식 (3)과 같이 각각의 영상 전체 면적으로 정규화 한 확률 값을 사용하는 수정된 히스토그램 인터섹션 (Modified Histogram Intersection) 기법으로 질의 영상과 객체 영역 사이의 유사도를 측정한다[6].

$$SH(H_{query}, H_{ROI}) = \sum_{j=1}^n \min \left(\frac{H_{query}^j}{S_{query}}, \frac{H_{ROI}^j}{S_{ROI}} \right) \quad (3)$$

식 (3)에서 $SH(H_{query}, H_{ROI})$ 는 질의 영상과 객체 영역의 유사도를 어떻게 계산하는지 나타내고 있다. H_{query} 와 H_{ROI} 는 질의 영상의 히스토그램과 데이터베이스 영상으로부터 구하여진 객체 영역의 히스토그램을 나타내고 있다. 또한 S_{query} 와 S_{ROI} 는 질의 영상과 객체 영역의 크기를 나타낸 것이다. n 은 히스토그램 bin의 개수를 나타낸다.

3. 실험 결과

본 논문에서 제안하는 방법은 영상 검색을 위하여 FOCUS 데이터베이스 광고 영상을 사용하고, 26개의 질의 영상을 사용하여 검색 성능을 평가한다. (그림 6)은 실험에 사용된 질의 영상의 일부이다.



(그림 6) 실험에 사용된 질의 영상

제안하는 방법에 대한 객관적인 평가 기준으로써, 현재

MPEG-7 컬러 및 질감에 대한 성능 평가 방법인 ANMRR 지표를 사용한다[7]. 그리고 관심 객체 추출의 성능을 평가하기 위하여 객체 추출의 성공 및 실패에 대한 평가도 같이 진행한다.

3.1 검색 성능 평가(ANMRR)

식(4)는 검색 성능 평가를 위해 사용하는 ANMRR (Average Normalized Modified Retrieval Rank)에 대한 식이다.

$$ANMRR = \frac{1}{Q} \sum_{q=1}^Q NMRR(q)$$

$$NMRR(q) = \frac{MRR(q)}{K + 0.5 - 0.5 * NG(q)}$$

$$MRR(q) = AVR(q) - 0.5 - \frac{NG(q)}{2} \quad (4)$$

$$K = \min(4 * NG(q), 2 * GTM)$$

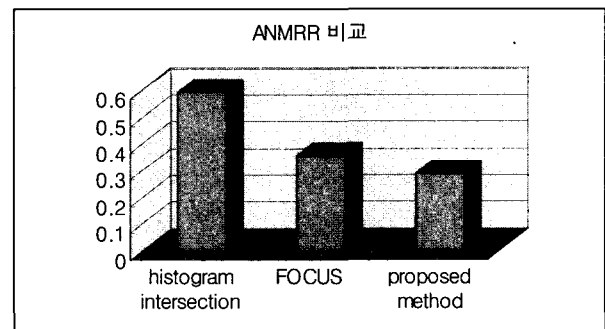
$$AVR(q) = \sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)}$$

여기서 $NG(q)$ 는 카테고리별 영상 개수를 의미하고, GTM 은 $NG(q)$ 중 가장 큰 수를 의미한다. 위 식은 평균 순위를 정규화한 수식으로써 항상 0에서 1사이의 값을 가지며, ANMRR은 낮은 값일수록 좋은 결과를 나타낸다.

<표 1>과 (그림 7)은 컬러 히스토그램 인터섹션 기법을 사용하여 ANMRR를 비교한 결과로써 기존의 방법인 배경을 포함하는 전체 영상에 대한 검색 결과와 Das[2]가 제안한 FOCUS에 대한 검색 결과, 그리고 제안하는 방법인 객체가 존재할 수 있는 후보 영역을 추출한 영상에 대한 검색 결과를 나타내고 있다.

<표 1> ANMRR 검색 정확도 비교

구분	ANMRR
Histogram Intersection	0.593
FOCUS	0.352
제안하는 방법	0.286



(그림 7) ANMRR 검색 정확도 비교

세 가지 방법을 비교한 결과 제안하는 방법이 Das가 제안한 FOCUS 방법보다 18.8%정도의 성능 향상을 보여 우수한 검색 성능을 보여주고 있다.

3.2 객체 추출 성공률/실패율 및 시간

관심 객체 추출에 대한 성능 평가는 기존의 WHI(Weighted Histogram Intersection) 방법과 비교한다[8]. 비교를 위해 FOCUS DB에서 사용된 26개의 질의 영상과 상품을 포함하고 있는 56장의 데이터베이스 영상을 사용하고, 포함된 총 관심 객체의 개수는 76개이다.

제안하는 방법의 추출 성능을 <표 2>에 나타내고 있다. 추출 성능 비교는 사용한 76개 관심 객체에 대해 제안하는 방법이 81.6%로 WHI 방법보다 10.6%의 성능 향상을 보이고 있다.

<표 2>에서, 추출 성공은 하나의 관심 객체 영역에 안에 하나의 객체가 포함된 경우이다. 추출 실패인 경우는 하나의 관심 객체 영역 안에 2개 이상의 객체를 포함하는 경우와 추출된 관심 객체의 영역에 비해 실제 객체가 존재하는 영역이 1/2보다 작은 경우를 포함한다. 추출 시간은 26개의 질의로 각각 관심 객체를 포함하고 있는 56장의 데이터베이스 영상에 대해 수행한 평균 추출 시간을 의미한다.

<표 2> 관심 영역 추출 성능 비교

객체 추출 방법	추출 성공률	추출 실패율	추출 시간
WHI	54개(71.0%)	22개(28.9%)	4.9 sec
제안하는 방법	62개 (81.6%)	14개(18.4%)	0.6 sec

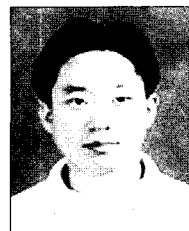
4. 결 론

본 논문에서는 컬러 영상에서 인접한 영역의 컬러쌍 정보와 후보 영역들 간의 클러스터링 과정을 통하여 관심 영역에 대한 객체 기반 영상 검색을 수행하는 시스템을 제안한다. 질의 영상의 각 화소들에 대해 주변 화소들과의 색상 차를 구하고, 그 중 가장 큰 값을 일정한 문턱치와 비교하여 문턱치보다 큰 경우에는 평탄 컬러 영역들의 경계에 위치하고 있으므로, 두 컬러에 대한 컬러쌍을 추출하고 히스토그램을 구성한다. 같은 방법으로 데이터베이스 영상은 일정한 크기의 영역으로 분할하여 각각 컬러쌍을 추출하고, 질의 영상에서 추출한 컬러쌍 정보가 존재하는 후보 영역들을 추출한다. 마지막으로 후보 영역들에 대한 클러스터링 과정을 통하여 최종적인 관심 객체 영역을 추출하고 유사도를 측정함으로써 기존 검색 방법에서 배경으로 인한 오류 검색을 줄였고, 객체의 이동, 회전, 크기변화에도 강건한 검색결과를 얻는다.

더 좋은 검색 성능을 얻기 위해서는 객체에 대한 모양 및 질감 정보에 대한 고려가 필요하며, 영상 내에 존재하는 다중 객체에 대한 연구에서는 근접한 객체가 서로 분리될 수 있도록 하는 연구가 진행되어야 한다. 또한, 특정한 광고 영상만을 대상으로 하는 것이 아닌 모든 영상에 대해서도 검색 성능을 높일 수 있는 연구가 진행되어야 한다.

참 고 문 헌

- [1] V. Castelli, L. D. Bergman, "Image Databases : Search and Retrieval of Digital Imagery," WILEY, 2002.
- [2] M.Das, E.M.Riseman, B.A.Draper, "FOCUS : Searching for multi-colored objects in a diverse image database," *International Conference Computer Vision and Pattern Recognition*, pp.756-761, 1997.
- [3] D. Androutsos, K. N. Plataniotis, and A. N. Venetsanopoulos, "A Vector Angular Distance Measure for Indexing and Retrieval of Color," *Proc. Storage & Retrieval for Image and Video Databases VII*, SPIE-3656, San Jose, USA, pp.604-613, 1999.
- [4] R.C.Gonzalez, R.E.Woods, *Digital image processing*, Prentice-Hall, 2001
- [5] M. Swain, and D. Ballard, "Color Indexing," *International Journal of Computer Vision*, Vol.II -32, pp.11-32, 1991.
- [6] J. B. Oh, and Y. S. Moon, "Content-Based Image Retrieval Based on Scale-Space Theory," *IEICE Trans. Fundamental*, June, 1999.
- [7] ISO/IEC JTC1/SC29/WG1/ "Core Experiment on MPEG-7 Color and Texture Descriptors," Doc. N2819, *MPEG Vancouver Meeting*, July, 1999.
- [8] F. Ennesser, G. Medioni, "Finding waldo, or focus of attention using local color information," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.17, No.8, pp.805-809, Aug. 1995



이 형 진

e-mail : hjlee@cse.hanyang.ac.kr

2002년 한양대학교 전자계산학과 (공학사)

2004년 한양대학교 대학원 컴퓨터공학과 (공학석사)

2004년~현재 : 모비넷피아 연구원

관심분야 : 내용기반 영상검색, 패턴인식, MPEG-7, TV-Anytime 등



박기태

e-mail : parkkt@cse.hanyang.ac.kr
2000년 한양대학교 전자계산학과(공학사)
2002년 한양대학교 대학원 컴퓨터공학과
(공학석사)
2002년~현재 한양대학교 대학원 컴퓨터
공학과(박사과정)

관심분야 : 내용기반 영상검색, 패턴인식, MPEG-7, TV-Anytime
등



문영식

e-mail : ysmoon@cse.hanyang.ac.kr
1980년 서울대학교 공과대학 전자공학과
(공학사)
1982년 한국과학기술원 전기 및 전자공학과
(공학석사)
1990년 University of California at Irvine
Dept. of Electrical and Computer
Engr.(공학박사)

1982년~1985년 한국전자통신연구소 연구원
1989년~1990년 Inno Vision Medical 선임연구원
1990년~1992년 생산기술연구원 선임연구원
1992년~현재 한양대학교 컴퓨터공학과 부교수
관심분야 : 영상처리, 패턴인식, Image/Video Indexing, MPEG-7,
TV-Anytime 등