

칼라 영상에서의 중심 객체 추출에 관한 연구

김성영[†] · 박창민^{**} · 권규복^{***} · 김민환^{****}

요 약

본 논문에서는 영상에 포함된 중심 객체를 추출하는 방법에 대해 제시한다. 중심 객체는 촬영의 중심이 되어 영상의 중앙 부분에 비교적 큰 면적을 차지하는 객체로 정의하는데 영상 내용에 대한 중요한 정보를 제공한다. 중심 객체 추출을 위해 우선 입력 영상에 대해 해상도를 줄여가며 영상 분할하고 분할된 결과에 대해 계층적 영역 병합을 수행함으로써 객체가 많은 수의 영역으로 세분화되어 영상 분할되는 것을 방지할 수 있도록 하였다. 분할된 각 영역은 영상의 경계와 접하는 경계 영역과 그 외의 비경계 영역으로 분류하였다. 비경계 영역은 중심 객체에 해당될 가능성이 있는 영역으로써, 이들 중에서 영상 중심 부근에서 가장 큰 크기를 차지하는 영역이 핵심객체영역으로 선택된다. 또한 경계 영역 중에서 영상의 네 모서리에 인접하는 영역은 핵심배경영역으로 선택되어 핵심객체영역과 함께 중심 객체 추출에 이용된다. 각 비경계 영역은 핵심배경영역 및 핵심객체영역과 칼라 분포 유사도를 비교하여 배경영역과 전경영역으로 분류된다. 핵심객체영역 및 핵심객체영역과 연결성을 가지는 전경영역이 최종 중심 객체로 선택된다. 본 논문에서 제안된 방법은 비교적 복잡한 배경을 갖는 영상에 대해서도 어느 정도 만족할 만한 결과를 얻을 수 있었다.

A Study on Extraction of Central Objects in Color Images

Sung-Young Kim[†], Chang-Min Park^{**}, Kyu-Bok Kwen^{***}
and Min-Hwan Kim^{****}

ABSTRACT

An extraction method of central objects in the color images is proposed, in this paper. A central object is defined as a comparatively large object at center of the image. First of all, an input image and its decreased resolution images are segmented. Segmented regions are classified as the outer or the inner region. The outer region is adjacent to boundaries of the image and the other is not. Each region is merged with its neighbors when adjacent regions are included by a same region in the decreased resolution image. Then core object regions and core background regions are selected from the inner region and the outer region respectively. Core object regions are the representative regions for the object and are selected by using the information about the region size and location. Each inner region is classified into foreground or background regions by comparing values of a color histogram intersection of the inner region against the core object region and the core background regions. The core object region and foreground regions consist of the central object in the image.

Key words: 객체 추출, 칼라 영상 분할, 영상 분류, 내용기반영상검색, 데이터베이스 인덱싱

본 연구는 한국과학재단 목적기초연구(R01-2000 00275) 지원으로 수행되었음.

[†] 창원전문대학 멀티미디어과

^{**} 성심외국어대학 경영정보시스템전공

^{***} 한국 정보통신 교육원 강사

^{****} 부산대학교 컴퓨터공학과

1. 서 론

영상을 입력으로 이용하는 다양한 응용 분야들이 최근 증가하고 있다. 특히 영상 검색은 최근 상당히 주목을 받고 있고 많은 연구가 이루어지고 있는 분야이다. 영상 검색을 위해서는 데이터베이스에 저장되는 각 영상에 대한 분류 및 주제어에 대한 결정 과정이 필수적이다. 초기 영상 검색에서는 수동으로 인덱스를 지정하는 방식을 이용하였으나 입력 작업의 비효율성과 작업자의 주관에 따라 입력 내용이 결정되는 등의 많은 어려움을 내포하고 있었다. 이에 최근에는 영상을 분석함으로써 칼라 분포나 질감 등과 같은 영상의 특징 정보를 자동으로 추출하여 검색에 이용하는 내용기반영상검색(CBIR)이 등장하게 되었다. 그런데 현재까지의 많은 방법들은 영상의 전역적인 특징 정보를 이용하고 있다. 사람은 영상에 포함된 객체 단위로 유사 영상을 검색하고자 하는 것이 일반적이므로 영상의 전역적 특징 정보를 이용하는 것보다는 객체 단위의 특징을 이용함으로써 영상 검색 효율을 증가시킬 수 있을 것이다. 영상에서 객체 영역만을 추출함으로써 영상의 자동 분류 및 영상 데이터베이스에서의 효율적인 영상 인덱싱 등에도 유용하게 사용할 수 있을 것이다.

MPEG-4[1]에서는 기존의 MPEG-1, 2에서는 없었던 객체 기반 부호화 개념과 이에 따른 VOP(video object plane)라는 새로운 개념을 도입하였다. 이 개념은 압축하고자 하는 동영상을 이전처럼 화소들의 집합으로 생각하는 것이 아니라 서로 다른 층에 각각 놓여있는 객체들의 집합으로 간주하여 서로 다른 객체를 분리해 부호화하는 것을 포함하고 있다. 이와 같이 영상에 대한 사용이 증가하면서 영상으로부터 객체를 분리해야 되는 필요성도 증가하고 있다.

영상에 포함된 객체들을 추출하는 것은 결코 쉽지 않은 작업이다. 영상에 포함된 객체는 유사한 색이나 질감으로 구성되는 경우(그림1(a))도 있지만 상이한 색과 질감으로 구성되는 경우도 많다. 또한 복잡한 배경에 객체가 포함(그림1(b))되어 객체 영역 판단에 어려움을 주는 경우도 상당수 존재한다. 그리고 배경 영상과 같이 뚜렷한 객체가 없는 경우도 있다(그림1(c)). 사람은 오랜 기간동안 축적된 다양한 지식을 사용하여 영상에 포함된 객체를 비교적 쉽게 구분하여 인식할 수 있지만, 사람이 지니고 있는 고수준(high level)의 지식을 컴퓨터로 재현하여 사람과 유



(a) 단순 배경 (b) 복잡한 배경 (c) 원경

그림 1. 다양한 영상의 예

사한 작업을 수행하도록 하기 위해서는 많은 어려움이 따른다.

칼라영상으로부터 전경과 배경을 분리하기 위한 이전 시도들이 있었다. Huang *et al.* [2,3]은 영상의 전경과 배경을 자동으로 분리하는 방법을 개발하였다. 이들이 제안한 방법은 영상의 코너나 비교적 평탄한 배경에서 일정한 거리만큼 떨어진 영역을 전경으로 분리하는 방법이다. 이 방법은 경계 정보를 이용한 접근 방법으로써 비교적 변화가 심한 배경을 가진 영상으로부터 전경 영역만을 추출하기에는 어려움이 있다. Lu와 Guo[4]는 효율적인 영상 검색을 위해 배경 영역을 제거하기 위한 시도를 수행하였다. 이들은 배경에 대한 위치적 분석을 통해 배경으로 분류되는 확실한 영역만을 제거함으로써 배경에 포함되는 많은 영역들이 남게 되는 문제를 가지고 있다.

알려진 객체 정보를 이용하여 복잡한 영상으로부터 주어진 객체 모델과 유사한 특징을 갖는 영역만을 추출하는 방법[5-8]도 개발되었다. Cheng *et al.* [5]은 입력 영상을 분할하고 이들을 객체 모델과 유사도를 비교하여 가장 유사한 영역을 찾아 낸다. 이후 찾아낸 영역을 초기 영역으로 하고 이 영역과 공간적으로 인접한 영역들을 초기영역에 더해 나가면서 객체 모델과 유사도를 계속적으로 비교함으로써 최종 객체를 추출한다. 이 방법은 객체에 대한 정보가 알려진 경우에만 적용 가능하다. 또한 추출할 객체를 스케치하여 형상 정보로 이용하고 객체의 질감 및 색상 정보는 사용자가 직접 영상에서 선택하여 이들 정보를 모두 만족하는 영역을 추출하는 방법[9]도 시도되었다.

또한 영상 분할을 통해 대략적인 객체에 해당되는 영역을 추출하는 방법들도 소개되었다[10,11]. 이들 방법들은 주로 내용기반영상검색에서 객체단위의 질의를 수행하기 위해 영상 분할과정에서 뚜렷하게 구별되는 특징을 가지는 영역 위주의 영상 분할이 가능하도록 하여 분할된 각 영역이 영상 검색에 이용

되도록 하고 있다. 그런데 이들 방법에서 분할된 영역은 객체 영역과는 어느 정도 차이가 있고 각 영역이 영상의 중심 객체인지 아니면 보조 객체인지에 대한 판단은 사람이 결정해야 한다. 따라서 영상 검색 단계에서는 질의 영상에서 기준 영역에 대한 선택 과정이 필요하게 되고 데이터베이스의 각 영상과의 정합과정에서 많은 비교가 발생하게 된다.

본 논문에서는 중심 객체를 정의하고 영상에 존재하는 중심 객체를 자동화된 방법으로 추출하는 방법을 제안한다. 영상 내에는 보는 관점에 따라 많은 수의 객체가 존재할 수 있다. 사람은 포함된 객체들의 중요 정도를 판단하여 의미있는 객체를 선택할 수 있지만 영상 처리를 통해서는 의미있는 객체에 대한 판단은 쉽지 않다. 본 논문에서는 중심 객체를 다음과 같이 정의한다. 중심 객체는 (1) 사진 촬영의 주요 대상이 되는 피사체이고, (2) 대체적으로 영상의 가운데 위치에 존재하며, (3) 영상에서 비교적 큰 영역을 차지한다. 영상으로부터 중심 객체를 추출하기에 앞서 영상에 중심 객체가 포함되어 있는지 여부를 판단하는 작업이 선행되어야 한다. 그런데 이 작업은 많은 연구가 수행되어야 되는 또다른 분야이므로 본 논문에서는 중심 객체가 포함된 영상을 대상으로 중심 객체를 추출하기 위한 시도를 하였다.

중심 객체 추출을 위해 본 논문에서는 우선 영상 분할을 수행하고 계층적 영역 병합 과정을 거쳐 영상 분할과정에서 발생하는 객체가 여러 영역으로 세분화되는 현상을 어느 정도 극복할 수 있도록 하였다. 분할된 영역들은 위치와 크기 정보를 이용하여 경계 영역 및 비경계 영역으로 분류된다. 경계 영역은 영상의 경계와 인접한 영역으로써 주로 배경에 해당될 가능성이 높은 영역이고 비경계 영역은 영상의 중심 위치에 존재하는 영역으로써 객체에 해당될 가능성이 높은 영역이다. 경계 영역과 비경계 영역으로부터 배경 및 중심 객체의 중요한 특징을 대표하는 핵심배경영역과 핵심객체영역을 각각 선택한다. 중심 객체는 관심 영역 윈도우를 이용하여 창에 포함되는 객체 후보영역들에 대해 핵심객체영역 및 핵심배경영역과 칼라 분포 유사도를 비교하여 최종 결정할 수 있도록 한다. 객체에 존재하는 그림자, 하이라이트 영역 등은 객체 단위의 영역 추출을 어렵게 하는 요소이므로 영역간의 유사도 비교시에 밝기 성분을 제외한 칼라 정보만을 이용하도록 하였다.

2장에서는 영상 분할 및 계층적인 영역 병합 방법에 대해 소개하고, 3장에서는 분할된 결과로부터 핵심객체영역 및 핵심배경영역을 선택하고 이들로부터 최종 객체 영역을 추출하는 방법에 대해 다룬다. 이어서 4장에서는 실험 결과를 보이며, 5장에서는 결론 및 향후 연구 방향에 대해 설명한다.

2. 영상 분할 및 계층적 영역 병합

2.1 영상 분할

영상으로부터 객체에 해당되는 영역을 추출하기 위해서는 영상 분할과정이 필요하다. 지금까지 다양한 영상 분할 방법들이 소개되었으며 많은 방법들이 어느 정도의 만족스러운 결과를 제공한다. 따라서 이들 영상 분할 방법 중에서 적절한 처리 시간 내에 칼라와 질감을 함께 고려 가능한 방법이면 중간 단계에서 영역 병합 과정을 거치게 되므로 분할 방법에 따라 최종 중심 객체 추출 결과는 큰 영향을 받지 않는다.

본 논문에서는 영상 분할 방법으로 Deng *et al.* [12]이 제안한 JSEG 영상 분할 방법을 이용하였다. 이들이 제안한 JSEG 방법은 J-value의 계산을 기반으로 두고 있는 방법으로 J-value는 영역 경계일수록 값이 크고, 영역 내부일수록 값이 작도록 정의된다. JSEG 방법에서는 영상 분할 결과에 영향을 미치는 임계값이 세가지 존재한다. 본 논문에서는 미리 정해진 디폴트 값으로 임계값을 지정해 사용하도록 한다.

2.2 분할된 영역 분류

분할된 각 영역은 영상의 경계에 접하는 경계 영역(Boundary Region, BR) 및 접하지 않는 비경계 영역(Non-Boundary Region, NBR)으로 분류된다(수식 (1)).

$$\frac{L_{Ri}}{L} < T_1 \quad \text{AND} \quad \frac{S_{Ri}}{S} > T_2 \quad (i=1,2,\dots,n) \quad (1)$$

n 은 영상 내 분할된 영역의 총 개수이며, Ri 는 영상 내의 i 번째 영역이다. L , L_{Ri} 는 각각 영상 전체 둘레 길이 및 Ri 가 영상 경계에 접하는 길이를 나타낸다. 그리고 S , S_{Ri} 는 각각 영상 전체 크기 및 Ri 의 크기이고, T_1 , T_2 는 수식에 사용된 임계값이다. 일정 크기 이상을 가지며 영상의 경계에 낮은 비율로 접하는

영역을 비경계 영역으로 정의하며 그렇지 않은 영역을 경계 영역으로 정의한다. 본 논문에서 중심 객체를 영상의 중심 부근에 위치하고 어느 정도 크기가 큰 객체로 정의하였다. 따라서 분할된 각 영역의 위치 및 크기는 중심 객체 추출에 중요한 정보를 제공한다. BR은 영상의 경계와 인접한 영역이므로 주로 영상의 배경에 해당될 가능성이 높은 영역이고, 반면 NBR은 중심 객체에 해당될 가능성이 높은 영역이다.

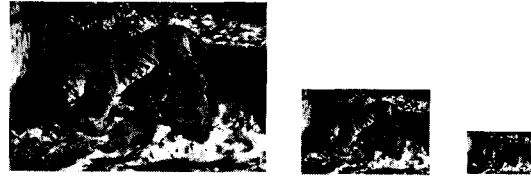
2.3 계층적 영역 병합

JSEG 방법은 질감에 비교적 무관하게 영상을 분할하지만 객체를 추출하기 위해서는 객체 영역이 다소 세분화되어 분할되는 경향이 있다. 또한 그림자, 하이라이트 등으로 인해 발생하는 영역 세분화는 피할 수 없는 문제이다. 본 논문에서는 JSEG의 영상 분할 결과를 보완하여 영상 내부의 사소한 색상 및 질감 변화, 그림자 및 하이라이트로 생긴 변화 등은 어느 정도 흡수하며 중요한 변화가 발생하는 영역 단위의 분할 결과를 얻기 위해 계층적 영역 병합을 수행하였다. 계층적 영역 병합은 입력 영상에 대해 해상도를 달리하며 계층적인 영상 집합을 생성하고, 각 해상도 영상에 대한 영상 분할을 수행한 후 그 결과를 최종 조합함으로써 이루어진다.

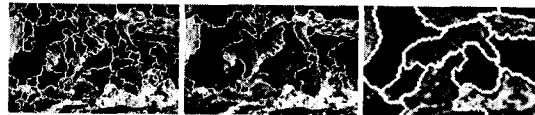
그림 2(a)는 입력 영상의 가로, 세로를 절반씩 줄여가며 만든 계층적인 영상 집합이고, 그림 2(b)는 각 해상도 영상을 분할한 후 동일한 크기로 표현한 것이다. 해상도를 줄이는 과정은 두 번만 수행하는데 더 이상 해상도를 줄이게 되면 경우에 따라 객체에 대한 정보가 거의 유실되어 잘못된 병합 결과를 초래할 수 있게 된다. 저해상도에서 영상 분할된 결과에는 세부적인 변화 및 그림자 등이 어느 정도 흡수되며 포괄적인 객체의 형태로 영상 분할된 것을 볼 수 있다. 계층적인 영역 병합은 저해상도에서의 포괄적인 객체 영역 정보와 원 영상에서의 세부적인 영역 정보를 함께 조합함으로써 중요한 영역 단위의 분할 결과를 얻을 수 있도록 한다.

계층적 영역 병합은 다음과 같은 과정으로 수행된다(그림 3).

- ① 계층적 영상 집합(I_0, I_1, I_2 ; I_0 은 원 영상, I_2 는 최저해상도 영상) 생성



(a) 해상도를 줄여 생성한 계층적 영상 집합



(b) 해상도별 영상 분할 결과

그림 2. 해상도에 따른 영상 분할 결과

- ② 각 영상의 영상 분할된 각 영역에 대해 인접성을 조사하여 영역 인접 그래프(Region Adjacent Graph, RAG) 생성
- ③ I_0 을 P_1 , I_1 을 P_2 로 설정
- ④ P_1 의 인접 영역(R_i, R_j)에 대해 다음의 두 가지 조건을 조사하여 모두 만족하면 병합
 $Attr(R_i) = Attr(R_j)$ ($Attr(R) \in \{BR, NBR\}$)
 $Ru(R_i) = Ru(R_j)$ ($Ru(R)$ 은 P_2 영상에서 P_1 의 영역 R 을 포함하는 영역)
- ⑤ P_1 의 모든 인접하는 영역에 대해 병합 과정 반복 수행하여 R_1 생성
- ⑥ R_1 을 P_1 , I_2 을 P_2 로 설정하고 ④, ⑤번 과정을 수행하여 최종 병합 결과 생성

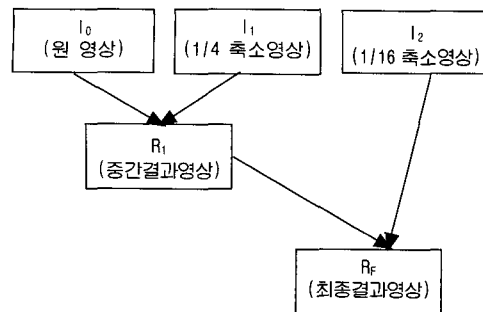
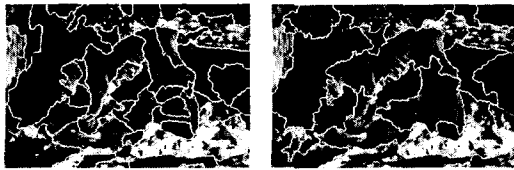


그림 3. 계층적 영역 병합

그림 4(a)는 JSEG 방법을 적용하여 구해진 영상 분할 결과로써 세가지 임계값을 수동으로 적절히 조절해서 영상에 포함된 객체가 최대한 잘 추출되도록 만든 결과 영상이며, 그림 4(b)는 계층적 영역 병합



(a) JSEG 영상분할 결과 (b) 계층적 영역 병합 결과
그림 4. 영상 분할 결과와 계층적 영역 병합 결과 비교

과정을 거쳐 최종 병합된 결과이다. JSEG 영상 분할 결과는 아무리 사람이 임계값을 적절히 설정한다고 하더라도 영상 분할이 상당히 세분화되어 객체 및 배경 영역이 많은 수의 영역들로 나누어져 분할된 것을 볼 수 있다. 이렇게 과잉 분할된 결과는 객체 추출에 어려움을 초래한다. 하지만 계층적 영역 병합 과정을 수행함으로써 의미 있는 영역 단위의 분할이 이루어져 객체 영역 추출에 많은 이득을 얻을 수 있다.

3. 중심 객체 추출

3.1 핵심배경영역

객체를 촬영한 영상에는 배경이 함께 포함된다. 배경 영역은 중심 객체 추출에 매우 중요한 역할을 담당한다. 객체 영역을 추출하기 위해서 객체에 대한 특징만을 사용해서는 정확한 결과를 얻기가 힘들게 되고, 객체를 둘러싸는 배경에 대한 정보도 함께 사용해야 비교적 용이하고 정확하게 객체 영역을 추출할 수 있다. 그런데 배경 영역을 정의하는 것 또한 쉬운 일이 아니다. 동영상에서는 연속된 몇 개의 영상에서 거의 변화가 없는 영역을 배경으로 정의할 수 있지만 정지 영상에서는 이런 정보를 이용할 수 없다. 또한 배경과 전경에 대한 경계가 모호한 경우도 흔하다.

본 논문에서는 중심 객체가 영상의 중심 부근에 위치한다고 정의하였는데, 이 정의에 기반을 두면 영상의 경계는 배경 선택에 아주 중요한 실마리를 제공한다. 특히 영상의 네 모서리는 배경 영역과 아주 밀접한 관계가 있다. 이에, 본 논문에서는 배경 영역을 대표하는 핵심배경영역(core background region, CBR)을 선택하되, 영상의 네 모서리에 인접하거나 주어진 임계값보다 큰 L_{Ri}/L 수치(수식 (1) 참조)를 갖는 BR을 CBR로 선택한다(그림 5).

3.2 핵심객체영역

분할된 영역들 중에서 중심 객체의 대표적인 특징

을 지니고 있는 영역을 핵심객체영역(core object region, COR)으로 선택한다. 핵심객체영역은 영상 내에서 중심 객체일 가능성이 가장 높은 영역이 되도록 설정된다. 선택된 핵심객체영역은 핵심배경영역과 함께 중심 객체 추출의 기준으로 사용된다. 핵심객체영역을 선택하기 위해 관심 영역 윈도우(Attention Window, AW)를 설정한다. 중심 객체는 영상의 중심 부근에 비교적 큰 비율을 차지하므로 AW는 중심 객체가 포함될 가능성이 높은 영상의 중심을 기준으로 영상의 1/4 크기를 차지하는 사각형 형태로 설정하였다(그림 5). AW가 설정된 후 핵심객체영역은 수식 (2)를 사용하여 추출한다.

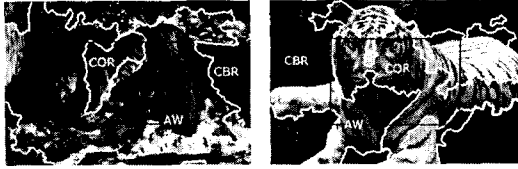
$$COR = R_i \\ i = ARG(Max(NR_1, NR_2, \dots, NR_n)) \quad (2)$$

NR_1, NR_2, NR_n 은 관심 영역 창에 포함되는 각 NBR들의 크기를 나타내며, ARG는 가장 큰 값의 인덱스를 구하는 함수이다. 수식 (2)를 사용함으로써 AW 내부에 포함된 NBR 중 가장 넓은 면적을 가지는 NBR이 핵심객체영역으로 선택되며, 이것은 영상 내에 객체가 포함될 경우 영상의 중심부에 비교적 큰 크기를 차지하고 있음을 반영한다.

그런데 객체가 근접 촬영되는 것과 같이 경우에 따라 객체가 NBR이 아닌 BR로 구성되는 경우도 발생하게 된다. 이런 경우에는 핵심객체영역을 BR에서 선택해야 되므로 최종 핵심객체영역을 선택하기 위해서는 수식 (2)로 구해진 영역과 AW 내에 포함되는 최대 크기의 BR에 대해 수식 (3)을 적용한다. 수식 (3)을 만족하는 경우에는 AW 내에 포함된 최대 크기의 BR을 핵심객체영역으로 선택하고 그렇지 않은 경우에는 수식 (2)로 구해진 NBR을 핵심객체영역으로 선택한다.

$$BR'_{max} > (BR_{max} - BR'_{max}) \quad (3)$$

BR_{max} 는 BR들 중에서 AW에 가장 큰 비율로 포함되는 영역의 크기를 나타내며, BR'_{max} 는 이 BR이 AW 내에 포함되는 크기이다. 수식 (3)은 객체가 영상의 대부분 영역을 차지하고 있을 경우에도 AW 내에 포함되는 핵심객체영역의 크기가 AW 내에 포함되지 않은 영역의 크기보다 일반적으로 더 크다는 의미를 포함한다. 그림 5는 선택된 핵심객체영역을 나타내고 있다. 그림 5(a)는 NBR중에서 핵심객체영역이



(a) COR이 NBR 중에서 선택 (b) COR이 BR 중에서 선택
 그림 5. 핵심배경영역(CBR), 핵심객체영역(COR) 및 관심 영역 참(AW)

선택된 결과이고, 그림 5(b)는 BR중에서 핵심객체영역이 선택된 결과를 나타낸다.

상당히 차이가 나는 칼라나 질감으로 객체가 구성 되어 있을 경우에는 하나의 핵심객체영역만을 이용해서는 좋은 추출 결과를 얻을 수 없다. 본 논문에서는 두 개까지의 핵심객체영역을 선택할 수 있도록 하였다. 두 번째 핵심객체영역은 첫 번째 핵심객체영역으로부터 영역을 확장하여 핵심객체영역을 추출한 후 나머지 NBR중에서 추출된 핵심 객체와 인접하면서 가장 큰 크기를 갖는 영역으로 선택한다. 첫 번째 핵심객체영역은 반드시 선택되어야 하지만 두 번째 핵심객체영역은 다음과 같은 경우 생략될 수 있다. 두 번째 핵심객체영역으로 선택된 영역이 1) 핵심배경영역과 유사하거나 2) 크기가 영상 크기의 일정치 이하인 경우에는 두 번째 핵심객체영역을 생략한다.

영상 내에서 객체와 배경을 대표하는 영역들이 추출된 후에 이들 정보를 기반으로 최종 객체 영역들을 추출하기 위해서 영상 내에서 객체를 포함하고 있을 가능성이 높은 영역들을 우선 추출한다. 이러한 영역들을 객체후보영역으로 정의하였다. 객체후보영역을 설정함으로써 객체 영역 추출 단계에서 비교 영역의 수를 줄일 수 있을 뿐만 아니라 상관없는 영역이 객체 영역으로 병합될 가능성을 줄일 수 있다. 객체후보 영역은 AW에 속하는 모든 NBR로 설정하였다(그림 6). 이때 NBR들 중에 크기가 작은 것은 인접한 영역들 중 크기가 가장 큰 영역에 병합하여 제거하였다.

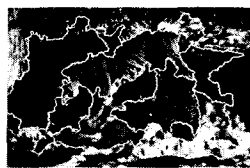


그림 6. 객체 후보 영역

3.3 중심 객체 추출

각 객체후보영역에 대해 핵심객체영역 및 핵심배경영역과 칼라 분포 유사도를 비교하여 핵심객체영역에 좀더 유사한 칼라 분포를 갖는 영역들을 최종 객체 영역으로 추출한다. 이때 칼라 분포 유사도는 히스토그램 인터섹션[13]을 사용하였다. 최종 객체 영역은 다음과 같은 과정으로 선택된다.

- ① $REF \leftarrow$ 핵심객체영역
- ② 객체후보영역 중에서 REF 에 인접한 영역 R_i 선택
- ③ R_i 가 CBR 과 인접한 경우, $HI(R_i, COR) >= T_1$ & $HI(R_i, COR) >= HI(R_i, CBR)$ 이면 (HI : Histogram Intersection, COR : 핵심객체영역, CBR : 핵심배경영역) $REF \leftarrow REF \cup R_i$
 R_i 가 CBR 과 인접하지 않은 경우 $HI(R_i, COR) >= HI(R_i, CBR)$ 이면 $REF \leftarrow REF \cap R_i$
- ④ 모든 인접 영역에 대해 번 ③과정 수행
- ⑤ REF 에 변화가 없으면 두 번째 핵심객체영역 선택하여 ①번 과정부터 반복 수행하고 더 이상 사용할 핵심객체영역이 없으면 REF 에 포함된 영역을 최종 객체 영역으로 설정하고 종료
- ⑥ 변화된 REF 에 대해 ②번 과정부터 반복 수행

객체에 포함된 그림자, 하이라이트 등은 객체 추출에 많은 어려움을 초래한다. 계층적 영역 병합을 통해 그림자, 하이라이트 영역으로 발생하는 어려움을 어느 정도 극복하였지만 이들로 인해 객체에 해당되는 영역이 완전히 추출되지 못하는 어려움은 여전히 존재한다. 객체 영역이 그림자 등으로 인해 잘못 추출되는 것을 최소화하기 위하여 본 논문에서는 객체후보영역에 대한 칼라 유사도를 계산할 때 밝기 성분은 제외하고 칼라 성분만을 사용하였다.

4. 실험 및 토의

본 논문에서 사용된 테스트 영상은 코렐 포토 CD 및 인터넷을 통하여 객체가 포함된 것을 무작위로 획득하여 사용하였다. 영상으로부터 최종 추출된 객체 영역은 그림 7과 같다. 복잡한 배경에 상이한 칼라 및 질감으로 구성된 객체라도 사람이 인식하는 것과

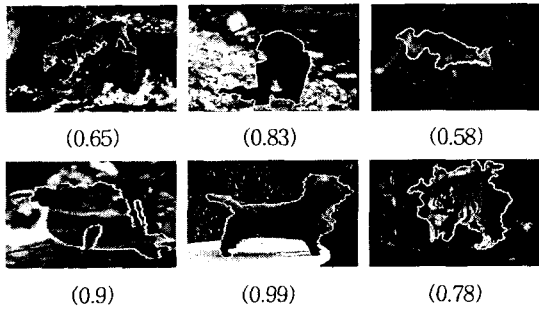


그림 7. 객체 추출 결과(괄호 내부는 추출 정확도)

유사하게 추출된 것을 볼 수 있다. 500개 이상의 영상을 대상으로 실험한 결과 배경이 단순한 영상뿐만 아니라 복잡한 배경을 가지는 영상에서도 비교적 높은 객체 추출 성공률을 나타내었다.

추출된 결과에 대해 정성적인 분석을 하기 어려웠던 논문에서는 주어진 영상에 대해 사람이 수동으로 중심 객체를 추출한 결과와 제안된 방법으로 추출된 결과를 정량적으로 분석하였다. 그림 8은 무작위로 선택된 100개의 영상에 대해 사람이 추출한 결과와 제안된 방법으로 추출된 결과를 비교하여 제안된 방법의 정확도를 표시한 그래프이다. 정확도는 다음과 같이 정의된다.

$$A = \frac{S_G - (S_u + S_o)}{S_G} \quad (3)$$

G 는 사람에 의해 수동으로 추출된 객체 영역이며 E 는 제안된 방법으로 추출된 객체이다. S_G 는 G 의 크기이며 S_E 는 $G \cap E$ 의 크기이다. 정확도는 제안된 방법으로 추출된 객체가 수동으로 추출된 객체와 정확하게 일치하는 경우에 최대 1이 되며, 객체 영역이 일부 추출되지 못하거나 배경이 포함되어 추출되면 정확도는 낮아지게 된다. 정확도는 픽셀단위로 계산

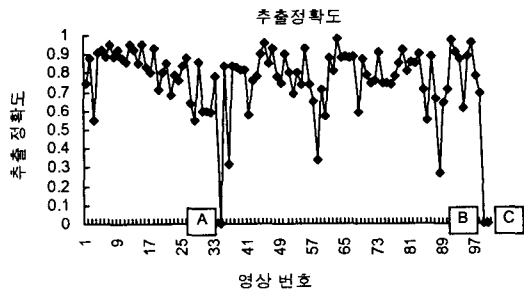


그림 8. 추출 정확도

되므로 제안된 방법으로 추출된 객체가 실제 객체와 유사하더라도 정확도가 비교적 낮은 값으로 계산되는 경우가 빈번하다. 이와 같은 현상으로 인해 정확도가 0.8이상이면 객체가 상당히 정확하게 추출된 것으로 간주할 수 있다. 제안된 방법으로 추출된 객체는 평균 0.7정도의 정확도를 가진다.

그림 9는 그림 8의 추출 정확도가 0인 세 개의 영상(A, B, C)을 표시한 그림이다. A, B 영상은 객체의 일부 영역을 포함하여 추출된 형태이지만 객체의 크기에 비해 배경 영역이 너무 많이 포함됨으로써 추출 정확도가 0으로 계산된 경우이다. C 영상은 객체 영역이 추출되지 못하고 배경 영역만이 추출된 결과이다. 이런 현상은 핵심객체영역이 객체 내부의 영역이 선택되지 못하고 배경 영역 중에서 선택됨으로써 발생된다. 핵심객체영역은 중심 객체의 정의에 의해 결정된다. 따라서 중심 객체가 영상의 중심에서 어느 정도 벗어나 있으며 또한 너무 작은 영역들로 세분화되어 영상 분할된 경우에는 영상의 중심 부근에 있으며 비교적 크기가 큰 배경 영역이 오히려 영상의 핵심객체영역으로 선택되어 현재의 방법으로는 잘못된 결과를 초래하게 된다. 이런 현상을 줄이기 위해서는 앞으로 보완이 필요할 것이다.

그림 10은 배경의 복잡도에 따라 객체 영역 추출 성공률을 그래프로 표시한 결과이다. 배경의 복잡도는 사람이 수동으로 추출한 객체 영역을 제외한



영상 A 영상 B 영상 C

그림 9. 추출 정확도가 0인 결과 영상

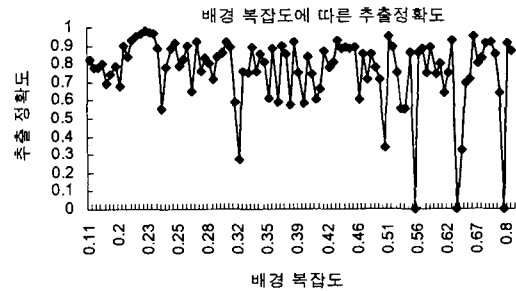


그림 10. 배경의 복잡도에 따른 추출 성공률

나머지 영역의 채널별 분산의 평균값과 밝기 변화에 대한 경계 강도의 합으로 계산된다. 그림 10에서 알 수 있듯이 배경 복잡도가 낮은 경우에는 비교적 높은 추출 성공률을 나타내었으며 배경이 복잡해질수록 추출 정확도는 다양한 값을 나타내었지만 일정수준 이상의 추출 성공률을 얻을 수는 있었다.

그림 11은 명도 성분을 배제한 경우와 칼라 성분 및 명도 성분을 모두 사용한 경우의 객체 추출 결과를 비교한 결과이다. 사슴의 몸통에 있는 털의 색상이 명도 차이로 인해 앞쪽과 뒤쪽이 다르게 나타나 있는데 칼라 성분과 명도 성분을 모두 사용한 경우에는 사슴 뒤쪽 몸통이 객체 영역에서 배제되어 객체 영역이 추출되지만 명도 성분을 제외시킴으로써 좀더 나은 객체 형태의 추출이 가능하였다. 특히 일찍 영상은 양자화됨으로써 사람에게서는 유사하게 보였던 색상이 처리 단계에서는 어느 정도 차이를 가지고 나타나게 된다.

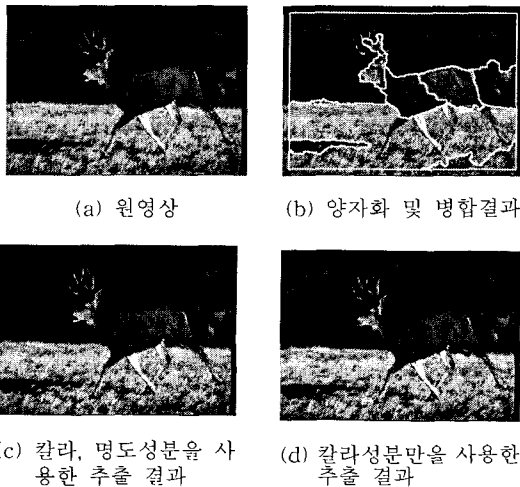


그림 11. 명도 성분 사용 여부에 따른 추출 결과 비교

5. 결론 및 향후 연구

본 논문에서는 객체가 포함된 칼라 영상에서 영상 내용에 대한 중요한 정보를 담고 있는 객체 영역을 추출하는 방법을 제안하였다. 이를 위해 해상도를 줄여가며 만든 영상 집합에 대한 영상분할과 계층적인 영역 병합을 수행하여 사소한 차이는 흡수하고 비교적 큰 차이를 나타내는 영역들로 영상을 분할할 수 있도록 하였다. 그리고 영역의 공간적 정보 및 크기

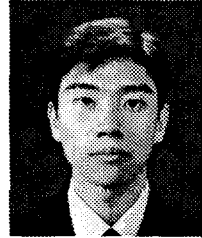
정보를 활용하여 핵심객체영역 및 핵심배경영역을 선택하고 객체 후보 영역에 대해 칼라 분포 유사도를 계산하여 최종 객체 영역을 추출할 수 있도록 하였다. 제안된 방법은 단순한 배경뿐만 아니라 복잡한 배경을 가지는 영상에 대해서도 어느 정도 만족할 만한 결과를 얻을 수 있음을 확인하였다. 본 논문에서 제시한 객체 추출 방법은 내용기반 영상 검색에서 영상의 전역적 특징을 이용한 검색이 아닌 객체단위로 검색할 수 있는 기반을 제공함으로써 검색 효율(속도, 정확성)을 보다 높일 수 있을 것으로 예상된다. 또한 영상 데이터베이스 검색 및 색인화에도 유용하게 사용될 수 있을 것이다.

영상에 포함된 객체를 추출하는 작업은 사람이 가지고 있는 개념을 영상으로부터 추출된 저수준의 특징들을 조합하여 표현하는 과정이므로 상당히 어려운 작업이다. 따라서 향후에도 계속적인 연구가 필요하다. 특히 복잡한 배경이 포함된 영상에 대해서는 좀더 보완이 필요하다. 이를 위해 칼라 유사도 뿐만 아니라 질감 정보 및 영역간의 위치 정보 등 좀더 다양한 특징들을 조합할 수 있는 방법이 요구된다. 조명의 변화에 의한 그림자 혹은 하이라이트 등도 객체 추출에 어려움을 주는 요소이다. 계층적 영역 병합, 명도 성분을 제외한 칼라 유사도 비교를 사용함으로써 이런 현상을 어느 정도 극복하였지만 이에 대한 보완 연구도 필요하다.

참 고 문 헌

- [1] MPEG SPEC. ISO/IEC JTC1/SC29/WG11, Vol 1-4, 1995.
- [2] Qian Huang, Byron Dom, David Steele, Jon Ashley, and Wayne Niblack. "Foreground/Background segmentation of color images by integration of multiple cues", In Proceedings of 1995 IEEE Conference on Image Processing, pages I246-I249, Washington D.C, U.S.A., 1995.
- [3] Qian Huang, Byron Dom, David Steele, Jon Ashley, and Wayne Niblack, "Segmentation and Representing Background in Color Images", Proceedings of ICPR '96, 1996.
- [4] Y. Lu and H. Guo, "Background Removal in

- Image indexing and Retrieval”, International Conference on Image Analysis and Processing, pp. 933-938, 1999.
- [5] Jian Cheng, Siegbert Drué, Georg Hartmann, joerg Thiem, Pohlweg 47-49, D33098 Paderborn, Germany, “Efficient Detection and Extraction of Color Objects from Complex Scenes”, 2000 IEEE Conference on Image Processing, 7.
- [6] R. Schettini, “Multicolored object recognition and location”, Pattern Recognition Lett, 15, 1089-1097, 1994.
- [7] F. Ennesser and G. Medioni, “Finding Waldo, or focus of attention using local color information”, IEEE Trans. Pattern Analysis Mach. Intell. 17 805-809, 1995.
- [8] V. Vinod and H. Murase, “Focused color intersection with efficient searching for object extraction”, Pattern Recognition 30(10), 1787-1797, 1997.
- [9] Yu Zhong, Anil K. Jain, “Object localization using color, texture and shape”, Pattern Recognition 33, 671-684, 2000.
- [10] A.H.Kam, T.T.Ng, N.G.Kingsbury and W.J.Fitzgerald, “Content Based Image Retrieval through Object Extraction and Querying”, IEEE Workshop on Content-based Access of Image and Video Libraries, pp. 91-95, 2000.
- [11] S. Belongie, C. Carson, H. Greenspan, and J. Malik, “Color and texture-based image segmentation using EM and its application to content-based image retrieval”, In IEEE Conf. on Computer Vision and Pattern Recognition, pp 675-682, 1998.
- [12] Y. Deng, B. S. Manjunath, and H. Shin, “Color Image Segmentation”, IEEE Conference on Computer Vision and Pattern Recognition, pp. 446-451, 1999.
- [13] M.J Swain and D.H Ballard, “Color Indexing”. Int. J. Comput. Vision 7(1), 11-32, 1991.



김성영

1994년 부산대학교 컴퓨터공학과 공학사
 1996년 부산대학교 대학원 컴퓨터공학과 공학석사
 1998년 부산대학교 대학원 컴퓨터공학과 박사과정 수료
 1998년 3월~현재 창원전문대학

멀티미디어과 전임강사

관심분야 : 칼라영상분할, 내용기반영상검색



박창민

1991년 독일 도르트문트대학교 컴퓨터공학과 공학사
 1995년 독일 도르트문트대학교 대학원 컴퓨터공학과 공학석사
 2000년 부산대학교 컴퓨터공학과 박사과정 수료

1997년 9월~현재 성심외국어대학 정보통신학부 조교수

관심분야 : 칼라영상처리, 내용기반영상검색



권규복

2000년 밀양대학교 컴퓨터공학과 공학사
 2002년 부산대학교 대학원 컴퓨터공학과 공학석사
 현재 한국 정보통신 교육원 강사

관심분야 : 칼라영상처리, 컴퓨터비전, 내용기반영상검색



김민환

1980년 서울대학교 전기공학과 학사
 1983년 서울대학교 대학원 컴퓨터공학과 석사
 1988년 서울대학교 대학원 컴퓨터공학과 박사

1991년~1992년 Univ. of Washington 객원연구원

1986년~현재 부산대학교 컴퓨터공학과 교수
 관심분야 : 화상처리 및 이해, 칼라공학