

칼라 영상에서의 객체 영역 추출에 관한 연구

김성영*, 권규복***, 박창민***, 김민환**

* 창원전문대학 멀티미디어과

** 부산대학교 컴퓨터공학과

*** 성심외국어대학 경영정보시스템전공

A Study on Extraction of Object in Color Image

Sung-Young Kim**, Kyu-Bok Kwen*, Chang-Min Park***, Min-Hwan Kim**

* Dept. of Multimedia Changwon College

** Dept. of Computer Engineering Pusan National University

*** Dept. of Management Information System Sung Sim Foreign College

Email : kyubok@hyowon.pusan.ac.kr, sykim@changwon-c.ac.kr,

cmpark@sungsim.ac.kr, mhkim@hyowon.pusan.ac.kr

요 약

본 논문에서는 칼라영상에 포함된 객체를 추출할 수 있는 방법을 제시한다. 객체는 상이한 칼라와 텍스처로 구성되는데 이들을 포괄하여 사람이 인식하는 객체와 유사한 형태로 추출할 수 있도록 하였다. 이를 위해 해상도를 줄여가며 영역 분할한 후 그 결과를 조합함으로써 사소한 변화는 흡수하고 중요한 변화가 발생하는 영역 단위의 분할이 가능하도록 하였다. 분할된 영역들의 공간적 정보와 영역 크기 정보를 이용하여 분할된 영역 중에서 핵심 객체영역과 핵심 배경영역을 선택하고 객체 후보 영역에 대해 이들과 칼라 분포 유사도 조사를 하여 핵심객체영역과 유사한 유사도를 갖는 영역들을 최종 객체 영역으로 추출하였다. 본 논문에서 제안된 방법은 다양한 실험을 통해 단순한 배경을 가진 영상뿐만 아니라 복잡한 배경을 갖는 영상에 대해서도 어느 정도 만족할 만한 결과를 얻을 수 있음을 확인하였다. 본 논문에서 제안된 방법은 내용기반검색, 영상 DB의 인덱싱 등 다양한 분야에 활용될 수 있을 것으로 예상된다.

1. 서론

최근 컴퓨터 상에서 영상의 사용이 증대됨에 따라 영상을 다루는 기술들이 급속하게 발전하고 있다. 그런데 이들 방법들은 대부분 영상의 전체적인 정보를 이용할 뿐 영상에 포함된 객체 정보는 제대로 활용하지 못하고 있다. 영상에 포함된 객체를 추출할 수 있다면 영상 데이터베이스의 인덱싱, 내용기반검

색 등의 다양한 분야에서 유용하게 활용할 수 있을 것이다. 이미 MPEG 4[1]에서는 영상에 포함된 객체들을 분리해야 되는 필요성이 대두되었으며 다른 많은 응용에서도 객체 추출에 대한 요구가 발생하고 있다.

그런데 영상에 포함된 객체를 추출하는 것은 매우 어려운 작업이다. 영상에 포함된 객체는 유사한 색이나 텍스처로 구성되는 경우(그림1(a))도 있지만

상이한 몇 가지 색과 텍스처로 구성되는 경우도 많다(그림1(b)). 또한 복잡한 배경에 객체가 포함되어 객체 영역 판단에 어려움을 주는 경우도 상당수 존재한다. 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 유사한 작업을 수행하도록 하기 위해서는 많은 어려움이 따른다.

[2]에서는 색상이나 패턴의 변화가 거의 없는 평탄한 배경을 가진 영상에서 전경을 분리하고 있다. 이들이 제안한 방법은 복잡한 배경을 가지는 영상에서는 배경과 전경을 분리하기 힘들다. 내용기반검색에서 객체의 정보를 활용하기 위해 시도한 Netra[3], Blobworld[4]에서는, 객체 단위의 영역 추출이 힘들기 때문에 영상 검색 질의를 수행할 때 객체의 중요 특징을 포함한 영역을 사용자가 직접 선택하게 함으로써 많은 사용자 개입을 요구하고 있다.

본 논문에서는 영상에 포함된 저수준의 특징을 추출하고 조합하여 사람이 인식하는 객체와 유사한 형태로 객체를 추출할 수 있는 방법을 제안한다. 영상에 포함된 객체에는 그림자, 하이라이트 영역 등이 존재하여 객체 단위의 영역 추출을 어렵게 하고 있다. 이를 해결하기 위해, 영상의 칼라화 정도를 파악하여 정보로써 활용할 수 있도록 한다. 또한 객체내의 사소한 변화는 무시하며 포괄적인 객체 정보를 파악하기 위해, 해상도를 줄여가며 계층적인 영상 집합을 생성한 후 각 영상에 대해 영상 분할하고 그 결과를 조합하여 객체에 해당되는 핵심 객체 영역 및 핵심 배경 영역을 선택한다. 최종 객체 영역은 관심 영역 창(attention window, AW)을 이용하여 창에 포함되는 객체 후보 영역들을 핵심 객체 영역 및 핵심 배경 영역과 유사도 비교를 하여 최종 결정할 수 있도록 한다.

영상 내에는 보는 관점에 따라 많은 수의 객체가 존재할 수 있다. 사람은 포함된 객체들의 중요 정도를 판단하여 의미있는 객체를 선택할 수 있지만 영상 처리를 통해서 의미있는 객체에 대한 판단 역시 쉽지 않다. 본 논문에서는 촬영의 중심이 되어

대체적으로 영상의 가운데 위치에 존재하며 크기가 비교적 큰 객체를 추출의 대상이 되는 의미있는 객체로 간주하였다. 의미있는 객체가 포함되지 않은 영상은 접근 방법을 달리해야 되므로 본 논문의 처리 대상에서는 제외하였다(그림1(c)).



(a) 단순배경 (b) 복잡한 배경 (c) 원경

그림 1. 다양한 영상의 예

2장에서는 해상도 축소에 따른 계층적 이미지에 대한 영상 분할 방법을 소개하고 3장에서는 분할된 결과로부터 핵심 객체영역 및 핵심 배경영역을 선택하고 영역의 통계적, 공간적 특징을 이용하여 최종 객체 영역을 추출하는 방법에 대해 다룬다. 이어서 4장에서는 실험 결과를 보이며 5장에서는 결론 및 향후 연구 방향에 대해 설명한다.

2. 객체 영역 추출을 위한 영역 분할

본 논문에서는 영상 분할을 위한 방법으로 Netra에서 제안한 방법[5]을 이용하였다. 이 방법은 영역 내부의 색상과 텍스처 분포 유사도를 반영하는 J값에 기반한다. J값이 큰 영역일수록 영역 경계일 가능성이 높고, 작을수록 영역 내부일 가능성이 높도록 정의되었다. 단계적으로 마스크의 크기를 바꾸며 마스크 내부에서 J값을 계산한다. J값이 가장 작은 영역부터 시작하여 인접 픽셀을 병합하면서 주위 영역으로 확장하며 최종 영역 분할 결과를 얻게 된다.

Netra에서 제안된 방법은 텍스처에 비교적 무관하게 영상을 분할하지만 객체를 추출하기 위해서는 너무 세부적으로 영역이 분할되는 문제가 있다. 또한 그림자, 하이라이트 등으로 인해 객체가 세분화되는 문제는 피할 수 없다.

본 논문에서는 입력 영상에 대해 해상도를 달리하며 계층적인 영상 집합을 생성하여 영상 분할하고

그 결과를 조합함으로써 영상 내부의 사소한 변화는 무시하며 영역을 추출할 수 있도록 하였다.

그림 2(a)는 입력 영상의 가로, 세로를 절반씩 줄여가며 만든 계층적인 영상 집합이고, 그림 2(b)는 각 해상도에 영상 분할을 수행한 후 동일한 크기로 표현한 것이다. 저해상도에서 영상 분할된 결과에는 세부적인 변화 및 그림자 등이 어느 정도 무시되며 포괄적인 객체의 형태를 나타내며 영역 분할된 것을 볼 수 있다.

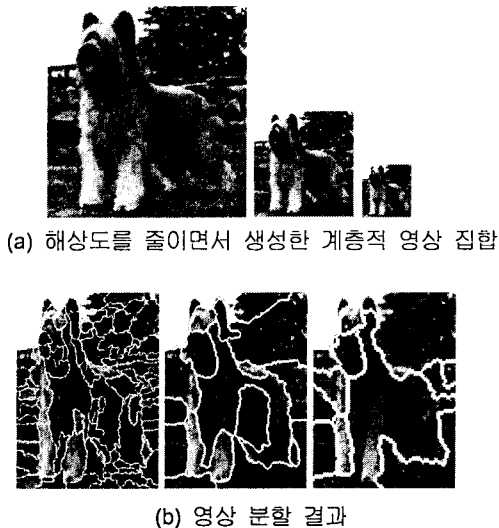


그림 2. 해상도에 따른 영상 분할 결과

저 해상도에서의 분할 결과는 너무 개략적이므로 각 해상도에서 분할된 결과는 다음과 같이 조합하여 최종 분할된 결과를 얻을 수 있도록 하였다. 각 해상도에서 분할된 영역에 대해 영상 경계와 접하는 영역은 경계 영역(BR), 접하지 않는 영역은 비경계 영역(NBR)으로 정의하였다. BR은 배경에 해당될 가능성이 높은 영역이고, NBR은 객체에 해당될 가능성이 높은 영역이다.

1. 원 영상의 분할된 영역에 대해 인접성을 조사하여 영역 인접 그래프(Region Adjacent Graph, RAG)를 생성한다.
2. 인접 영역에 대해 하위 해상도에서 동일 영역 포함 여부를 조사하여 동일 영역에

포함되면 병합한다. 이때 BR과 NBR이 서로 병합되지 못하도록 하였다.

3. 병합된 결과에 대해 RAG를 생성한 후 최하위 해상도 영상이 될 때까지 2번 과정을 반복한다.

그림 3(a)는 Netra에서 제안된 방법으로 분할된 결과이고 그림 3(b)는 제안된 방법으로 추출된 결과이다. 제안된 방법에서 객체 추출에 용이하게 영역이 추출된 것을 볼 수 있다.

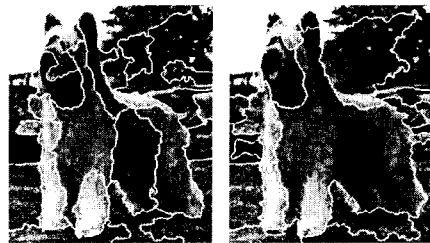
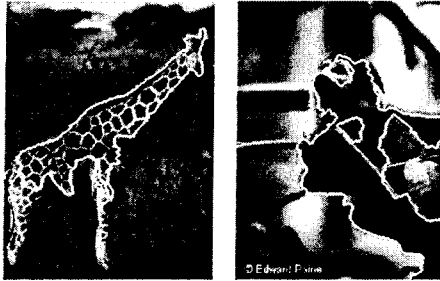


그림 3. 영상 분할 결과

3. 객체 영역 추출

앞에서 분할된 결과를 이용하여 객체 영역을 추출하기 위해 우선 객체의 대표적인 특성을 지니고 있는 핵심 객체 영역을 선택한다. 핵심 객체 영역은 영상 내에서 객체일 가능성이 가장 높은 영역이다. 핵심 객체 영역 선택을 위해 관심 영역 창(AW)을 설정한다. AW는 객체가 포함될 가능성이 높은 영역에 설정되는데 영상내에 의미있는 객체가 포함된 경우 주로 영상의 가운데에 나타나므로 기본적으로 영상의 중심을 기준으로 사각형 형태로 설정하였다.

핵심 객체 영역은 AW 내부에 포함된 NBR중 가장 넓은 면적을 가지는 것을 선택한다(그림 4(a)). 그런데 경우에 따라 객체가 근접 촬영되어 객체가 영상의 경계와 접하는 경우도 발생하므로 BR중에서 AW 내부에서 대부분의 면적을 차지하는 것이 존재하면 NBR 대신에 이 BR을 핵심 객체 영역으로 선택한다(그림 4(b)).



(a) NBR 선택결과 (b) BR 선택결과

그림 4. 핵심객체영역

복잡한 영상의 배경은 일반적으로 하나의 특징만으로는 표현되어질 수 없다. 예를 들어, 동물 사진을 촬영할 때 산과 하늘이 배경으로 함께 촬영되는 경우가 흔하다. 따라서 영상 전체의 배경을 표현하기 위해서는 각 배경을 대표하는 특징이 필요하다. 각 배경을 대표하는 핵심배경영역을 선택하기 위해 영상의 네 코너를 이용하여, 코너에 접하는 BR을 사용하였다(그림 5). 선택된 핵심배경영역 중 핵심객체영역과 칼라 분포 유사도가 높은 영역은 핵심배경영역에서 배제하였다.



그림 5. 핵심 배경영역

최종 객체 영역을 추출하기 위해서는 모든 객체 영역이 포함된 객체 후보 영역으로부터 확실한 객체 영역을 결정해 간다. 객체 후보 영역은 AW에 속하거나 울리는 모든 NBR로 설정하였다(그림 6).



그림 6. 객체 후보 영역

객체 후보 영역들은 이미 결정된 핵심객체영역 및 핵심배경영역과 칼라 분포 유사도를 비교하여 핵심 객체 영역과 보다 유사한 칼라 분포를 갖는 영역들을 최종 객체 영역으로 추출한다. 이때 칼라 유사도는 히스토그램 인터섹션[6]을 사용하였다.

객체에 포함된 그림자, 하이라이트 등은 객체 추출에 많은 어려움을 초래한다. 계층적 영역 분할을 통해 그림자, 하이라이트 영역으로 발생하는 어려움을 어느 정도 극복하였지만 여전히 어려움은 존재한다. 본 논문에서는 객체 후보 영역에 대한 칼라 유사도를 계산할 때 영역의 칼라화 정도를 계산하여 칼라화 정도가 높은 영역은 칼라와 명도를 함께 사용하도록 하였고 칼라화 정도가 낮아 회색톤이 강한 영역에서는 명도를 제외한 칼라 성분만을 사용할 수 있도록 하였다.

4. 실험 및 토의

본 논문에서 사용된 테스트 영상은 인터넷을 통하여 객체가 포함된 것을 무작위로 획득하였다. 객체가 단순한 배경에 포함되어 있는 것뿐만 아니라 복잡한 배경에 포함되어 있는 것도 함께 사용하였다. 영상으로부터 최종 추출된 객체 영역은 그림 7과 같다. 복잡한 배경에 상이한 칼라 및 텍스처로 구성된 객체라도 사람이 인식하는 것과 유사하게 추출된 것을 볼 수 있다. 150개 이상의 영상을 대상으로 실험한 결과 배경이 단순한 영상에서는 90% 정도의 좋은 결과를 나타내었고 복잡한 배경을 가지는 영상에서는 70%의 성능을 나타내었다.



그림 7. 최종 결과

그림 8은 원하는 수준까지의 객체 추출에 실패한 경우이다. 객체 추출에 실패한 영상들은 대부분은 배경이 객체와 비슷한 칼라 분포를 갖는 경우로써 완전히 잘못된 영역을 객체로 선택한 것은 아니고 객체를 전체적으로 추출하지 못하고 단지 일부분만을 추출한 것이 대부분이다. 또한 객체가 현저하게 차이나는 두 가지 이상의 칼라 분포를 갖는 경우에도 실패한 결과가 비교적 많이 발생하였다. 이에 대한 보완은 필요하다.



그림 8. 객체 추출에 실패한 결과

본 논문에서 제시한 객체 추출 방법은 내용기반 영상 검색에서 영상 전체의 특징을 이용한 검색이 아닌 객체단위로 검색할 수 있는 기반을 제공함으로써 검색 효율(속도, 정확성)을 보다 높일 수 있을 것으로 예상된다.

5. 결론 및 향후 연구

본 논문에서는 객체가 포함된 칼라 영상에서 의미 있는 객체 영역을 추출하는 방법을 제안하였다. 이를 위해 해상도를 줄여가며 만든 영상 집합에 대한 계층적인 영상분할을 수행하여 사소한 차이는 흡수하

고 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 어느정도 만족할 만한 결과를 얻을 수 있음을 확인하였다. 하지만 복잡한 배경이 포함된 영상에 대해서는 아직 보완이 필요하다. 이를 위해 칼라 유사도 뿐만 아니라 텍스처 정보 및 영역간의 위치 정보를 조합할 수 있는 방법이 요구된다.

[참고문헌]

- [1]. MPEG SPEC. ISO/IEC JTC1/SC29/WG11
Vol 1-4
- [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 1246-1249, Washington D.C, U.S.A., 1995
- [3]. W.Y. Ma, "NETRA: A Toolbox for Navigating Large Image Databases," Ph.D. Dissertation, Dept. of Electrical and Computer Engineering, University of California at Santa Barbara, 1997
- [4]. C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik, "Blobworld: A System for region-based image indexing and retrieval," the Third Int. Conf. On Visual Information Systems, 1999
- [5]. Y. Deng, B. S. Manjunath, and H. Shin, "Color Image Segmentation", IEEE Conference on Computer Vision and Pattern Recognition, pp. 446-451, 1999
- [6]. M.J Swain and D.H Ballard, Color Indexing. Int. J. Comput. Vision 7(1), 11-32(November 1991).