

칼라-공간 히스토그램의 통계 정보를 이용한 자연 영상의 영역 분할 및 레이블링 기법

신수연, 김우생
광운대학교 컴퓨터공학부

Natural Image Segmentation and Labeling Technique by Color-Spatial Histogram and Statistics

Suyoun Shin, Woosaeng Kim
Dept of Computer Engineering, Kwangwoon University
E-mail: comma337@hotmail.com, kws@cs.kwangwoon.ac.kr

요약

영역 분할과 영역 레이블링은 내용에 기반한 영상 검색이나 영상 이해를 위해 선행되어야 하는 중요한 작업 중의 하나이다. 본 논문에서는 칼라-공간 히스토그램의 통계정보를 통해 자연 영상내의 영역을 효율적으로 분할하고 또한 이러한 데이터를 생성규칙으로 만들어 레이블링 하는 새로운 방법을 제안한다. 제안하는 방법은 자연영상 처럼 많지 않은 영역으로 이루어진 경우 매우 효율적임을 보였다.

1. 서론

웹의 발전과 멀티미디어 데이터에 대한 필요에 의해 영상에 대한 자동 검색, 분류 등에 대한 기술이 점점 더 요구되고 있는 상황이다. 또한 공장 자동화, 로버틱스 등의 컴퓨터 비전의 응용에서는 영상을 자동으로 인식하고 이해하는 기술들이 요구되고 있다. 이러한 영상의 검색, 분류, 인식, 이해 등의 작업을 하기 위한 하나의 선행 작업으로써 영상 내의 각 영역을 분할하고 레이블링 하는 기술이 요구된다.

기존의 영역 분할 방법으로는 군집화, 히스토그램, 영역에 기반한 기법, 에지에 기반한 기법 등 여러가지 기법들이 제안되고 있다[1, 2, 3, 4]. 또한 분할된 영역에 레이블링을 하는 기법으로는 이산완화(discrete relaxation), 확률완화, 탐색트리, 유전자 영상해석 기법 등 다양한 방법들이 연구되고 있다[5, 6, 7, 8].

본 연구에서는 영상의 칼라 분포 뿐 아니라 공간 정보도 함께 나타낼 수 있는 칼라-공간 히스토그램을 사용 하여 영상의 영역들을 자동으로 분할하는 방법

* 본 연구는 정보통신부 대학기초연구지원사업
(과제번호:2001-042-3)에 의해 수행되었음.

을 제안한다. 또한 영상의 사전 정보를 통한 생성 규칙들을 사용하여 분할된 영역들에 레이블링을 하는 방법도 제안한다. 실험을 통하여 자연영상 같이 객체가 많지 않은 경우는 제안한 방법이 매우 효율적임을 보였다.

본 논문의 구성은 다음과 같다. 2장에서는 칼라-공간 히스토그램과 빈 영상, 3장에서는 빈 영상을 통한 영역 분할과 레이블링 방법, 그리고 4장에서는 실험을 통하여 제안하는 방식의 성능을 보였다. 마지막으로 5장에서는 결론과 미래의 연구에 관하여 언급하였다.

2. 칼라-공간 히스토그램 및 빈영상

칼라 히스토그램은 영상내의 색의 분포를 표현하는 기본적인 방법이다. 그러나 영상내의 영역을 분할하고 레이블링 하기 위해서는 영상의 색 분포 뿐 아니라 공간 정보도 필요하다. 따라서 우리는 영상의 칼라 분포 뿐 아니라 공간적인 정보도 표현 할 수 있는 칼라-공간 히스토그램을 사용한다. 칼라-공간 히스토그램은 각 빈에 속한 모든 화소들의 칼라, 숫자, 중심점과 분산 값을 포함하는 벡터로서 표현이 된다 [9].

칼라-공간 히스토그램은 다양한 통계정보를 포함하기 때문에 이것을 비주얼 하게 표현하기 위해 빈영상을 사용한다. 빈영상에 존재하는 화소(빈화소)들의 위치와 색은 칼라-공간 히스토그램의 각 빈에 속한 모든 화소들의 중심점과 색에 대응한다. 그림 1(b)는 그림 1(a) 원 영상의 일반적인 칼라 히스토그램이다. 그림 1(c)의 2차원 빈영상은 그림 1(b)의 칼라 히스토그램의 각 빈에 속한 모든 화소들의 중심점과 색에 대응하는 빈화소의 위치와 색을 보여주고 있다. 반면에, 그림 1(d)의 3차원 빈영상은 빈화소들의 위치와 색뿐 아니라 대응되는 칼라 히스토그램의 각 빈에 속한 모든 화소들의 숫자와 그들의 x 축, y 축 분산 값들도 함께 보여주고 있다. 칼라-공간 히스토그램의 이러한 통계적 데이터들로부터 우리는 원 영상내의 영역들의 형태 및 특징들을 예상할 수 있다. 예

를 들어 만약 특정 빈화소의 위치가 상단에 있고 화소 수에 비해 분산 값이 크지 않다면, 원 영상의 대응되는 영역은 상단에 모여있는 형태일 것이다.

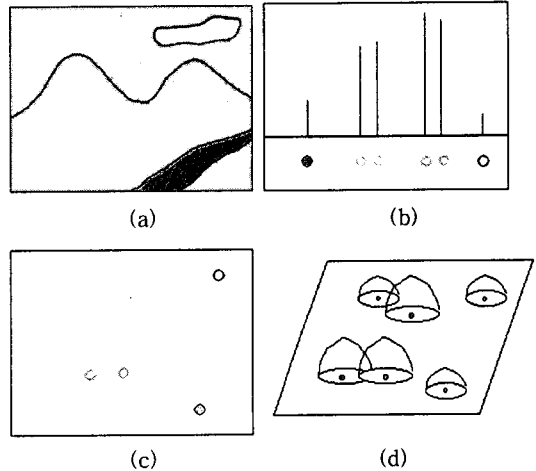


그림 1. 빈 영상의 2차원과 3차원 뷰

3. 빈영상을 통한 영역 분할 및 레이블링 기법

전체적인 과정은 시스템은 우선 원 영상을 빈 영상으로 바꾼다. (실제적으로 빈영상은 단지 칼라-공간 히스토그램을 이해 시키는 도구로 사용될 뿐이지 시스템은 칼라-공간 히스토그램을 통하여 작업을 한다.) 그런후 시스템은 빈영상에서 원영상의 영역들에 대응하는 대표 빈화소들을 구한다. 대표 빈화소들은 대응하는 영역에 속한 화소들을 갖고 있기 때문에 이것을 통하여 영역을 분할하고 대표 빈화소들의 색과 위치 정보를 통하여 영역들의 레이블링을 수행한다.

3.1 영상을 빈 영상으로 변환

시스템은 우선 원 영상을 빈 영상으로 변환한다. 본 연구에서 영역 분할을 시도하는 자연 영상들의 경우 단지 몇개의 객체들을 포함하고 있기 때문에 적은 수의 색으로도 객체들을 구분할 수가 있어 HSV를 36 단계로 양자화 시키는 방식을 사용하였다[10]. 이 방식은 여러 칼라 모델 중 인간이 색을 인지하는

방식과 유사하며 RGB등 다른 칼라 모델과는 다르게 유사한 칼라의 빈들은 서로 가까이 위치하는 장점이 있다. 즉 특정 빈들의 인덱스가 가까울 경우 해당 영역들은 비슷한 색들의 영역들로 쉽게 간주 할 수가 있다. 또한 이 방식은 36개의 적은 수의 양자화를 사용하기 때문에 영역 분할과 레이블링의 계산 복잡도를 상당히 줄여주는 장점이 있다.

3.2 대표 빈화소와 영역 분할

원 영상을 해석하는데 별로 필요없는 빈화소들을 제거하는 것이 도움이 된다. 따라서 우리는 빈화소를 구성하는 화소 수가 원 영상 전체 화소 수의 0.5%보다 작은 경우 잡음으로 간주하고 빈 영상에서 이들을 제거하였다. 이 과정 후에, 빈영상에는 그림 2(b)와 같이 원 영상의 영역과 관련된 빈화소들만 남게 된다. 그러나 원 영상에서 유사한 색의 이웃 하는 영역들은 하나의 객체를 표현할 가능성이 많기 때문에 이들을 하나로 병합하는 것이 바람직 하다. 따라서 빈 영상에서 서로 이웃하며 색이 유사한 빈화소들을 합병해 하나의 대표 빈화소로 만들어 준다. 그림 2(c)는 대표 빈화소들로 구성된 빈영상을 보여준다. 빈화소들을 합병하여 대표 빈화소로 만드는 방식은 다음과 같다.

1. 빈영상에 빈화소가 존재하면 (2)를 수행하고 더 이상 존재 하지 않으면 종료한다.
2. 빈영상의 빈화소들 중 포함하는 화소 수가 가장 많은 빈화소를 선택한다.
3. 선택된 빈화소와 이웃 하는 빈화소들 중에서 유사한 색을 갖는 빈화소들을 대표 빈화소로 합병한다.
4. 합병에 참가한 빈화소들을 제거하고 대표 빈화소의 통계적인 데이터들을 수정한 후 (1)로 간다.

본 연구에서 사용하는 칼라-공간 히스토그램은 기존의 칼라 히스토그램을 확장한 것이기 때문에 다중 문턱 값(threshold value)을 통한 영역 분할 기법을 적용 할 수 있다. 기존 히스토그램을 통한 영역 분할 기법의 문제점은 대상 영역을 효과적으로 분리시키는 적절한 여러 개의 문턱치를 설정하고 찾는 것이

쉽지 않다는 것이다. 반면 본 연구에서 구한 대표 빈화소들의 경우 대표 빈화소에 속한 화소들은 이미 원영상의 구분된 영역을 구성하는 화소들이기 때문에 그림 2(d)와 같이 각 대표 빈화소에 속한 화소들을 하나의 영역으로 분할을 해주면 된다. 또한 대표 빈화소의 경우 색과 위치도 함께 고려해 문턱 값을 결정하는 효과를 주기 때문에 비슷한 색의 서로 다른 두 영역이 떨어져 있는 경우 서로 다른 영역으로 분할 시키는 효과가 있다.

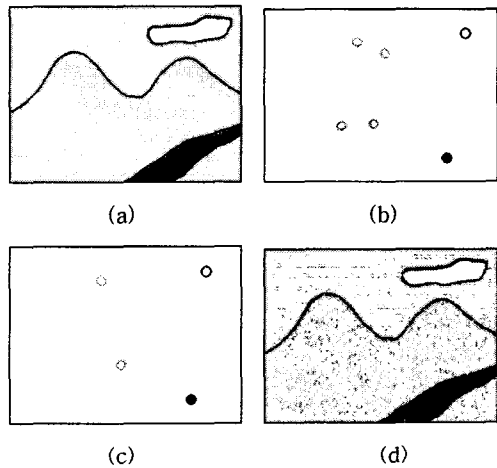


그림 2. 대표 빈화소를 통한 영역 분할

3.3 영역 레이블링을 위한 생성 규칙들

본 연구에서 대상으로 하는 자연 영상들은 산, 들, 강의 3가지 클래스 영상들로 가정을 하였다. 사실 영상에는 중요한 영역 외에도 여러 가지 작은 영역들이 포함 될 수 있다. 이러한 모든 작은 영역들은 사람의 눈으로도 구분하기 어려울 때가 있기 때문에, 본 연구에서는 각 클래스 영상을 특징 지을 수 있는 다음의 중요한 영역들만을 레이블 대상으로 선정하였다. 단, 하늘 영역 내에는 구름이 존재할 수 있기 때문에 레이블 대상이 되는 영역은 하늘, 구름, 산, 들, 강의 5개이다.

1. 산영상: (하늘, 산) 또는 (하늘, 산, 들)
2. 들영상: (하늘, 들) 또는 (들)
3. 강영상: (하늘, 산, 강) 또는 (하늘, 들, 강) 또는

(들, 강)

사전에 구할 수 있는 이러한 영역의 특징들과 영역들간의 상대적인 관계를 다음과 같은 생성 규칙으로 만들어 사용한다.

If condition X holds THEN action Y is appropriate

생성 규칙에서 대표 빈화소 R의 색, 위치, 대응하는 빈 인덱스 번호, 대응하는 영역의 이름은 각각 R_C , R_L , R_I , R_N 로 표현하고, 변수 x, y는 대표 빈화소에 대응하는 원 영상의 영역을 의미한다. 우선 대표 빈화소에 대응하는 영역의 색을 빈 인덱스를 통해 결정하는 규칙들은 다음과 같다.

IF $28 \leq R_I(x) \leq 30$ THEN $R_C(x)$ is 하늘색

IF $5 \leq R_I(x) \leq 7$ THEN $R_C(x)$ is 구름색

IF $26 \leq R_I(x) \leq 30$ THEN $R_C(x)$ is 강색

IF $0 \leq R_I(x) \leq 6$ or $28 \leq R_I(x) \leq 32$ THEN $R_C(x)$ is 산색

IF $0 \leq R_I(x) \leq 4$ or $12 \leq R_I(x) \leq 24$ THEN $R_C(x)$ is 들색

위의 규칙에서 산이나 들의 경우는 어두운 숲들로 존재할 수 있기 때문에 빈 인덱스 영역이 2군데로 설정이 되었다. 또한 규칙에서 보듯이 서로 다른 색이라도 빈 인덱스가 겹치는 경우가 존재한다. 따라서 어떤 대표 빈화소의 경우는 여러 개의 색으로 간주될 수 있다. 잡음 등으로 인해 대표 빈화소 x가 위의 어떤 색에도 포함되지 않을 경우는 빈 영상에서 x와 가장 가까운 위치에 있는 대표 빈화소 y의 색을 x의 색으로 결정한다.

대표 빈화소의 색과 위치를 통해 대응하는 영역을 결정할 수 있다. 하지만 영역의 위치가 항상 절대적인 것은 아니기 때문에 특정 대표 빈화소의 색과 위치가 주어졌을 때의 대응하는 영역의 신뢰도를 실험을 통하여 표 1과 같은 확률로써 표현하였다. 여기서 대표 빈화소 중심점의 y축의 위치에 따라 R_L 을 상단,

중단, 하단으로 구분하였다. 표1의 첫번째 행 (대표 빈화소의 색이 하늘색 일 때 위치에 따라 하늘 영역 일 확률)을 생성규칙으로 표현하면 다음과 같다.

IF $R_C(x)$ is 하늘색 & $R_L(x) =$ 상단 THEN $R_N(x)$ is 하늘영역 with 90%

IF $R_C(x)$ is 하늘색 & $R_L(x) =$ 중단 THEN $R_N(x)$ is 하늘영역 with 10%

IF $R_C(x)$ is 하늘색 & $R_L(x) =$ 하단 THEN $R_N(x)$ is 하늘영역 with 0%

표 1. 대표 빈화소의 색과 위치에 따른 영역의 이름

$R_C \backslash R_L$	상단	중단	하단
하늘색	하늘(0.9)	하늘(0.1)	하늘(0.0)
구름색	구름(0.8)	구름(0.2)	구름(0.0)
산색	산(0.1)	산(0.7)	산(0.2)
들색	들(0.0)	들(0.4)	들(0.6)
강색	강(0.0)	강(0.2)	강(0.8)

사전 지식들을 이러한 생성 규칙들로 표현하여 새로운 지식들을 추론할 수 있게 된다. 따라서 대표 빈화소의 통계치 데이터들을 이러한 규칙에 적용하여, 대응하는 영역이 어떤 영역인지 추론하여 레이블링하게 된다. 이때 특정 영역에 대하여 한 개 이상의 레이블링이 가능하지만 서로 다른 신뢰도로서 표현이 되기 때문에 신뢰도가 큰 것으로 결정을 한다.

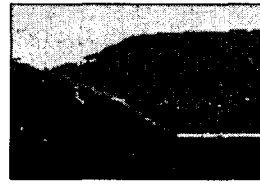
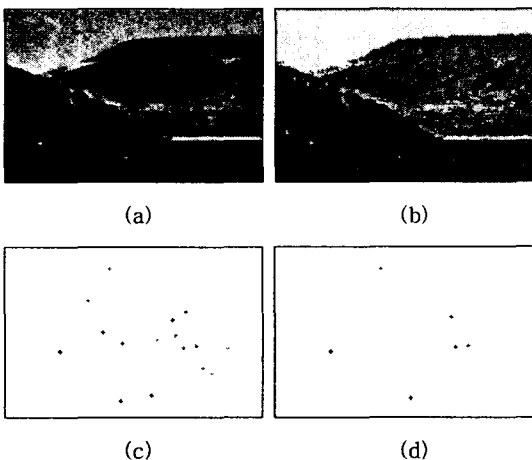
4. 실험 및 결과

칼라-공간 히스토그램에 기반한 영역분할과 영역 레이블링을 평가하기 위해 Corel Stock Photo에 있는 영상들을 사용하였다. 각 영상은 192*128 크기로 정규화를 하였으며, 실험을 통하여 다음과 같은 매개변수 값들을 사용하였다. HSV 색을 36개의 색으로 양자화하였을 경우 색은 크게 명암도(gray) 색의 군(빈0-빈7)과 휴(hue) 색의 군으로(빈8-빈35) 구분된다. 그런데 휴 색의 군은 다시 4개씩의 비슷한 색의 서브 군으로 나눌 수 있다. 따라서 본 연구에서는 주어진 대표 빈화소의

색이 명암도 색일 경우는 앞뒤의 빈 번호에 해당하는 명암도 색을 유사한 색으로, 휴 색일 경우는 해당 서브 군에 해당하는 휴 색들을 유사한 색으로 결정하였다. 또한 대표 빈화소들은 그들의 거리가 빈 영상 대각선 길이의 10% 안에 위치할 때 이웃 하는 대표 빈화소로 간주하였다.

4.1 영역 분할 결과

다음의 그림들은 원 영상을 빈 영상으로 바꾼 후 대표 빈화소들을 통해 영역을 분할한 결과를 보여준다. 시스템은 먼저 강영상 그림 3(a)를 36 색으로 그림 3(b) 처럼 양자화 시킨다. 그림 3(c)는 양자화된 영상의 칼라-공간 히스토그램의 통계 데이터를 2차원의 빈 영상으로 나타낸 것이다. 그림 3(d)는 빈 영상의 이웃하며 유사한 빈화소들을 군집화 한 대표 빈화소들을 보여준다. 마지막으로 그림 3(e)는 대표 빈화소들에 속한 화소들로 영역을 분할 한 결과를 보여준다. 이 그림에서 보듯이 영역을 분할 한 결과는 원 영상을 36 색으로 양자화한 결과와 비슷하지만 이웃의 비슷한 색의 작은 영역들은 큰 영역에 포함되어서 나오게 된다. 이는 대표 빈화소가 이웃하는 유사한 색의 빈화소들을 군집화 시키는 역할을 하기 때문에 영역을 객체 단위로 분할하기 때문이라고 볼 수 있다.



(e)

그림 3. 대표 빈화소를 통한 영역분할 결과

4.2 영역 레이블링 결과

그림 4는 대표 빈화소의 색과 위치에 따른 정보를 통하여 영역들을 레이블링한 결과를 보여준다. 그림 4(a), 4(b), 4(c)는 각각 산영상, 들영상, 강영상의 영역들을 분할 하여 레이블링을 한 결과를 보여준다. 이때 한 개 이상의 레이블링 되는 영역의 경우는 가장 확률이 높은 값으로 결정해준다. 예를들어 4(c)에 강영역 같은 경우 대표 빈화소의 인덱스 값은 28이고 위치는 하단에 해당된다. 따라서 색으로는 하늘, 산, 강역이 다 가능하나, 위치 정보의 경우 강의 신뢰도가 가장 높기 때문에 강영역으로 최종 레이블링 된다.

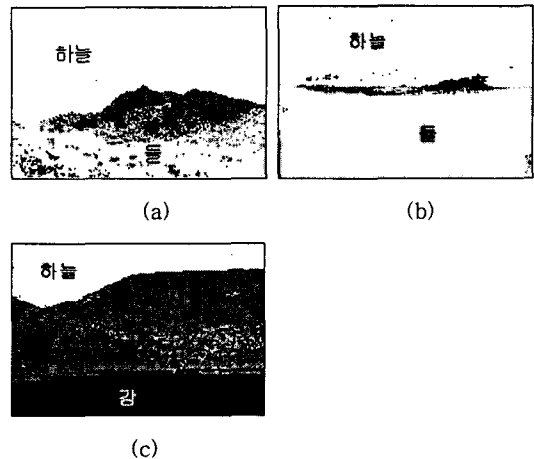


그림 4. 대표 빈화소의 통계 데이터를 통한 영역 레이블링 결과

다음의 표 3은 3개의 자연영상 (각 클래스마다 20개의 영상)에 있어서 각 영역들이 얼마나 제대로

레이블링이 되는지를 보여준다. 표 3에서 보듯이 영상에 관계 없이 하늘, 구름, 들 영역은 90% 가까운 높은 인식률을 보이는 반면 산영역의 경우는 75%의 낮은 인식률을 보인다. 이는 산영역은 어두운 색이 많고 위치도 하단에 존재하는 경우가 많기 때문에 들영역으로 오인식 될 가능성이 많다. 반면에 들영역은 색이 대부분 밝은 초록색이기 때문에 산영역으로 오인식 되는 가능성이 상대적으로 적게 나온다.

표 3. 자연영상의 각 영역에 대한 레이블링 결과

	하늘	구름	산	들	강
산영상	83%	94%	69%	94%	
들영상	94%	85%		88%	
강영상	84%	93%	78%	87%	82%

5. 결론 및 향후 연구 과제

본 논문은 칼라-공간 히스토그램의 통계 정보를 이용해 자연 영상을 분할하고 레이블링 하는 방법을 제시하였다. 원 영상의 영역들을 빈 영상의 대표 빈 화소로 표현된다. 따라서 빈 영상의 대표 빈화소의 통계적인 데이터로 원 영상의 대응하는 영역들의 정보를 추출할 수 있기 때문에 이를 통하여 영상을 분할하고 레이블링 하는 것이 가능해진다. 실제의 자연 영상들을 대상으로 하여 제안하는 방법이 영역들을 제대로 분할하고 좋은 레이블링 결과를 보여주는 것을 알 수 있었다.

그러나 영역이 과분할 되거나 레이블링이 정확하지 못한 경우도 존재한다. 이러한 문제점들의 일부는 영상의 추가적인 사전 정보를 통하여 줄일 수 있을 것으로 기대된다. 추후 연구과제로는 영상내에 존재할 수 있는 영역들과 존재 할 수 없는 영역들의 정보들을 사전지식으로 이용하여 영역 분할과 레이블링의 결과에 적용시키면 오류를 상당 부분 줄일 수 있을 것으로 기대된다.

참고문헌

[1] Koon-Pong Wong, Dagan Feng, S. T. Meikle, and M. J. Fulham, "Segmentation of dynamic PET images using cluster analysis", IEEE Transactions on , Volume: 49, pp. 200 -207

[2] J. Canny, "A computational approach to edge detection", IEEE Transaction on PAMI, Vol. 8, no. 6, pp. 679-698, 1986

[3] F. Cheevasuvit, H. Maitre, and D. Vidal-Madjar, "A robust method for picture segmentation based on split-and-merge procedure", Computer Graph Image Process. Vol. 34, pp. 268-281, 1986

[4] Rolf Adans and Leanne Bischof, "Seed Region Growing", IEEE Pattern Analysis and Machine Intelligence, Vol. 16, No. 6, 1994

[5] E. R. Hancock and J. Kittler, "Discrete relaxation", Pattern Recognition, Vol. 23, No. 7, pp 711-733, 1990

[6] W. E. L. Grimson and T. Lozano-Perez, "Localizing overlapping parts by searching the interpretation tree", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 9, No. 4, pp. 469-482, 1987

[7] M. Kamada, K. Toraichi, R. Mori, K. Yamamoto and H. Yamada, "Parallel architecture for relaxation operations", Pattern Recognition, Vol. 21, No. 2, pp. 175-181, 1988

[8] Yuntao Qian, "Image interpretation with fuzzy-graph based genetic algorithm", Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on , Vol: 1 ,pp. 545 549 1999

[9] 김택근, 김우생, "칼라 공간 히스토그램을 이용한 영상 검색 및 분류", 한국정보과학회 논문지 제출 중

[10] Lei, Z., Fuzong, L. and Bo, Z., "A CBIR Method Based on Color-Spatial Feature," 1999 IEEE TENCON, pp. 166-169, 1999