

# 칼라-공간 히스토그램과 생성 규칙을 이용한 자연 영상 레이블링 및 분류 기법

김준영, 신수연, 김우생  
광운대학교 컴퓨터공학부  
전화: 02-940-5217

## Natural Image Labeling and Classification Technique by Color-Spatial Histogram and Production Rules

Junyoung Kim, Suyoun Shin, Woosaeng Kim  
Dept of Computer Engineering, Kwangwoon University  
E-mail: [joon0@dreamwiz.com](mailto:joon0@dreamwiz.com), [syshin@cs.kwangwoon.ac.kr](mailto:syshin@cs.kwangwoon.ac.kr), [kws@cs.kwangwoon.ac.kr](mailto:kws@cs.kwangwoon.ac.kr)

### Abstract

The image labeling and classification is one of the important tasks for a content-based image retrieval and an image understanding. This paper propose a new technique to label and classify natural images with a color-spatial histogram and production rules. We show that our proposed method is very efficient for a natural image composed of a few regions.

본 연구에서는 영상의 칼라 분포 뿐 아니라 공간 정보도 함께 나타낼 수 있는 칼라-공간 히스토그램 (또는 빈영상)을 사용 하여 자연 영상의 분할된 영역들을 레이블링하고 분류하는 방법을 제안한다. 영역들에 레이블링하고 분류 하기 위해 빈영상의 대표 빈화소들에 표현되는 영상의 사진 정보를 생성 규칙들로 만들어 사용한다. 실험을 통하여 자연영상 같이 주요한 몇 개의 영역이 존재하고 영역간에 상대적인 관계가 존재할 때 제안한 방법이 효율적임을 보였다.

### I. 서론

영상을 레이블링하고 분류하는 작업은 내용기반 영상 검색이나 영상 이해등을 위한 중요한 작업이다. 기존의 영역에 레이블링을 하는 기법으로는 이산완화(discrete relaxation), 확률완화, 탐색 트리, 유전자 영상해석 기법 등 다양한 방법들이 연구되고 있다[1,2,3,4]. 또한 영상 분류를 하기위한 방법들은, 크게 영상의 하위레벨의 특징들을 이용하는 방식[5], 하위레벨과 상위레벨의 특징들을 동시에 이용하는 방식[6], 그리고 영상내의 영역들간의 색과 공간관계[7,8] 등을 고려한 방식등이 있다.

### II. 빈영상의 대표 빈화소를 통한 영역분할

칼라-공간 히스토그램은 각 빈에 속한 모든 화소들의 칼라, 숫자, 중심점과 분산 값을 포함하는 벡터로서 표현이 된다[9]. 칼라-공간 히스토그램은 다양한 통계정보를 포함 하기 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) 원영상의 일반적인 칼라 히스토그램이다. 그림 1(c)는 빈영상으로 그림 1(b)의 칼라 히스토그램의 각 빈에 속한 모든 화소들의 중심점과 색에 대응하는 빈화소의 위치와 색을 보여주고 있다. 이때 원영상을 해석하는데 별로 필요 없는 빈화

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

소들을 제거하는 것이 도움이 된다. 따라서 빈화소를 구성하는 화소 수가 작은 경우 잡음으로 간주하고 빈영상에서 이들을 제거한다. 이 과정 후에, 빈영상에는 원영상의 영역과 관련된 빈화소들만 남게 된다. 그러나 원영상에서 유사한 색의 이웃 하는 영역들은 하나의 객체를 표현할 가능성이 많기 때문에 이들을 하나의 객체로 병합하는 것이 바람직하다. 따라서 빈영상에서 서로 이웃하며 색이 유사한 빈화소들을 합병해 하나의 대표 빈화소로 만들어 준다. 따라서 이러한 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[9].

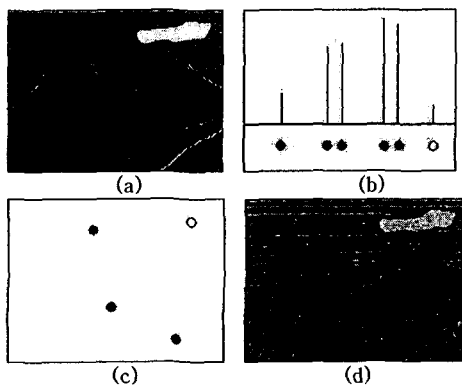


그림 1. 대표 빈화소를 통한 영역 분할

### III. 생성규칙을 통한 영역 레이블링 및 분류

본 연구에서 대상으로 하는 자연 영상들은 산, 들, 강 의 3가지 클래스 영상들로 가정을 하였다. 본 연구에서는 각 클래스 영상을 특징 지을 수 있는 다음의 중요한 영역들만을 레이블 대상으로 선정하였다. 단, 하늘 영역 내에는 구름이 존재할 수 있다.

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)

예를 들어 산영상 (하늘/산)의 표기는, 일부 산영상들은 두개의 영역으로 분할 되어 있으며 위는 하늘영역, 아래는 산 영역으로 구성된다는 의미이다. 사전에 구할 수 있는 이러한 영상의 특징과 영역들간의 상대적인 관계를 다음과 같은 생성 규칙으로 만들어 사용한다.

If condition X holds THEN action Y is appropriate

여기서 대표 빈화소 R의 색, 위치, 대응하는 빈 인덱스 번호, 대응하는 영역의 이름은 각각  $R_c$ ,  $R_p$ ,  $R_i$ ,  $R_N$ 로 표현하고, 변수  $x$ ,  $y$ 는 대표 빈화소에 대응하는 원 영상의 영역을 의미한다. 우선 대표 빈화소에 대응하는 영역의 색을 빈 인덱스 범위 값으로 결정하는 규칙들 (규칙1)은 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 산, 들, 바다의 경우는 어두운 색의 숲이나 바다로 존재할 수 있기 때문에 빈 인덱스 영역이 하나 이상으로 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  $x$ 가 위의 어떤 색에도 포함되지 않을 경우는 빈 영상에서  $x$ 와 가장 가까운 위치에 있는 대표 빈화소  $y$ 의 색을  $x$ 의 색으로 결정한다.

대표 빈화소의 색과 위치를 통해 대응하는 영역의 종류를 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 같이 표현하였다. 여기서 대표 빈화소 중심점의  $y$  축의 위치에 따라  $R_p$ 를 상단, 중단, 하단으로 구분하였다. 표1에서 대표 빈화소의 색이 하늘색이며 상단에 위치할 때 대응하는 영역이 하늘 영역 일 확률을 생성규칙 (규칙2)으로 표현하면 다음과 같다.

IF  $R_c(x) = \text{하늘색} \ \& \ R_p(x) = \text{상단}$  THEN  $R_N(x)$  is 하늘 영역 with 90%

표 1. 대표 빈화소의 색과 위치에 따른 영역의 이름

$R_c \backslash R_p$	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(0.5)	들(0.5)
강색	강(0.0)	강(0.2)	강(0.8)

이때 특정 영역에 대하여 한 개 이상의 이름이 가능하지만 서로 다른 신뢰도로서 표현이 되기 때문에

신뢰도가 큰 것으로 레이블링을 하는 것이 바람직하다. 그러나 잘못된 영역 분할이나 잡음 등으로 인해 이러한 방식의 레이블링 결과가 틀릴 수도 있다. 예를 들어 어떤 영상이 2개의 영역으로 분할이 되었는데 규칙 1, 2에 의해 위 영역은 산(70%), 강(20%) 또는 하늘(10%) 영역으로, 아래 영역은 들(60%) 또는 산(20%) 영역으로 이름이 선정 되었다면, 가장 높은 신뢰도인 (산/들)로 레이블 될 것이다. 그러나 이러한 영상은 실제로 존재하지 않는 영상이므로 차라리 (하늘/산) 혹은 (하늘/들)의 조합을 선정하고 이들 중 신뢰도가 가장 큰 (하늘/들)로 레이블 하는 것이 더 바람직 하다.

따라서 본 연구에서는 규칙1과 규칙2를 통해서 레이블을 결정 짓지 않고, 대신 각 분할된 영역에 모든 신뢰도의 영역 이름들을 포함 시킨다. 이러한 서로 다른 신뢰도를 갖는 영역 이름들은 다음 단계에서 실제 영상들과 비교를 통해 최종 레이블로 결정된다. 예를 들어 어떤 영상 Q가 두개의 위/아래 영역으로 분할되고 규칙1과 규칙2를 통해 위 영역에  $A_1..A_n$ 의 영역이름이, 아래영역에는  $B_1..B_m$ 의 영역이름이 설정 되었다고 가정하자. 반면 두개의 영역들을 포함하는 실제 영상 클래스들은  $(U_1/V_1, U_2/V_2, U_3/V_3)$ 의 영역 이름 조합들이라고 가정하자. 시스템은 먼저 영상 Q를 통해 실제 영상 클래스인  $U_i/V_i$ 로 레이블이 가능한가를 조사한다. 만약 영상 Q의  $A_2$ 가  $U_1$ 과 이름이 같고,  $B_3$ 가  $V_1$ 와 이름이 같다면, 영상 Q가  $(U_1/V_1)$ 로 레이블링 될 신뢰도는  $P(A_2|위치) + P(B_3|위치)$ 로 계산된다. 나머지 실제 영상 클래스에 대해서도 레이블링 될 경우의 신뢰도를 구해 가장 신뢰도가 높은 영역 이름들의 조합을 Q의 최종 레이블로 결정 한다. 영상이 3개로 분할되는 경우도 3개의 영역으로 구성되는 실제 영상 클래스들과 비교를 통해 최종 레이블을 만든다.

실제의 영상 클래스가 두개의 영역( $U_i/V_j$ )로 구성 될 때 ( $1 \leq i, j \leq 4$ ), 영역 이름 U, V에는 하늘, 산, 들 또는 강이 될 수 있다. 또한 주어진 영상도 크게 두개의 영역 ( $A_i/B_j$ )으로 분할 되었다고 할 때 ( $1 \leq i, j \leq 5$ ), 영역이름 A, B에는 4개의 영역 이름 외에 공(empty) 영역 이름을 추가한다. 이것은 주어진 영상의 특정 영역에 설정된 이름이 실제 영상 클래스의 해당 영역에 있는 어떠한 이름과도 같지 않을 경우에 설정하는 이름으로 이때의 신뢰도  $P(공영역|위치)$ 는 0이 된다. 주어진 영상 Q가 두개의 영역으로 분할 되었을 때 다음의 규칙 (규칙3)을 적용하여 가장 신뢰도가 큰 조합으로 레이블 한다.

IF  $A_i = U_r$  &  $B_j = V_s$  THEN Label $_Q(A/B)$  is  $(A_i/B_j)$  with  $P(A_i|위치) + P(B_j|위치)$

실제의 영상 클래스가 세 개의 영역( $U_i/V_j/W_k$ )으로 구성되었으며 또한 주어진 영상 Q도 세 개의 영역 ( $A_i/B_j/C_k$ )로 분할 되었을 때 다음의 규칙 (규칙3)을

적용하여 가장 신뢰도가 큰 조합으로 레이블 한다.

IF  $A_i = U_r$  &  $B_j = V_s$  &  $C_k = W_t$  THEN Label $_Q(A/B/C)$  is  $(A_i/B_j/C_k)$  with  $P(A_i|위치) + P(B_j|위치) + P(C_k|위치)$

일단 영상의 주요 영역들이 레이블링이 되면 이를 통해 영상을 분류할 수 있다. 주어진 자연 영상 Q가 두개의 영역으로 구성 되었을 때 영상을 분류하기 위해 다음과 같은 규칙 (규칙4)을 사용한다.

IF Label $_Q(A/B) = (하늘/산)$  THEN Q is 산영상  
 IF Label $_Q(A/B) = (하늘/들)$  THEN Q is 들영상  
 IF Label $_Q(A/B) = (들/강)$  THEN Q is 강영상

마찬가지로 주어진 자연 영상 Q가 세 개의 영역으로 구성 되었을 때 영상을 분류하기 위해 다음의 규칙 (규칙4)를 적용한다.

IF Label $_Q(A/B/C) = (하늘/산/들)$  THEN Q is 산영상  
 IF Label $_Q(A/B/C) = (하늘/산/강)$  THEN Q is 강영상  
 IF Label $_Q(A/B/C) = (하늘/들/강)$  THEN Q is 강영상

사전 지식들을 이러한 생성 규칙들로 표현하여 새로운 지식들을 추론할 수 있게 된다. 따라서 영역의 색과 상관관계 등을 이러한 규칙에 적용하여 주어진 영상을 레이블링 하고 분류하게 된다.

#### IV. 실험 및 결과

칼라-공간 히스토그램에 기반 한 영역분할과 영역 레이블링을 평가하기 위해 Corel Stock Photo에 있는 영상들을 사용하였다. 각 영상은 192\*128 크기로 정규화를 하였으며, 각 영상 클래스 당 25개의 영상을 사용하여 실험을 하였다.

##### 4.1 영역 레이블링 결과

그림 2는 대표 빈화소의 색과 위치에 따른 정보를 통하여 영역들을 레이블링 한 결과를 보여준다. 그림 2(a), 2(b)는 각각 산영상, 강영상의 영역들을 분할하여 레이블링 한 결과이다. 각 영역은 레이블링 가능한 영역 종류를 가지고 있으며, 이 영역 종류의 조합 중에서 신뢰도 합이 가장 큰 것으로 영역이 레이블링 된다.

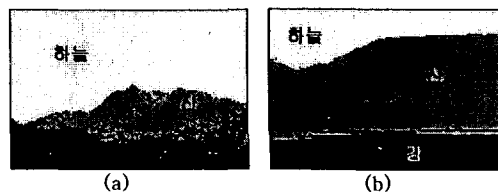


그림 2. 영역 레이블링 결과

그림 3(a)는 규칙 2까지 적용한 결과이며 그림 3(b)는 규칙 3까지 적용한 결과이다. 그림에서 보듯이 규칙 2만 적용했을 때 신뢰도가 가장 큰 바다로 잘못 레이블링이 된 영역이, 실제 영상의 사전지식을 이용한 규칙 3에 의해서 산으로 제대로 레이블링이 되는 것을 볼 수 있다.

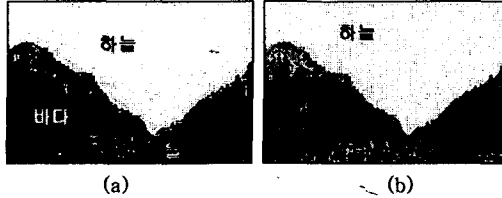


그림 3. 규칙 2와 규칙3을 적용한 레이블링 결과

표 2는 실험에 사용한 3가지 자연 영상에서 각 영역들이 얼마나 제대로 레이블링이 되는지를 보여준다. 다른 영역에 비해 산영역의 경우 낮은 인식률을 보이는 것은 산 영역과 들 영역의 색이 많은 부분 겹치는데, (하늘/산) 영상에서 산 영역이 중단/하단에 모두 위치하는 경우 대표 빈화소가 하단에 생성되면 더 높은 신뢰도를 갖는 들 영역으로 잘못 레이블링 될 수 있기 때문이다.

표 2. 자연영상의 각 영역에 대한 레이블링 결과

	하늘	구름	산	들	강
산영상	97%	100%	74%	96%	
들영상	100%	78%		97%	
강영상	91%	100%	89%	96%	82%

표 3은 영상들의 분류결과를 보여준다. 강영상에 비해 산영상과 들영상의 영상분류 결과가 낮게 나타난다. 이는 산영상의 경우 낮게 위치하는 산 영역이 들 영역으로 잘못 레이블링 되어 들영상으로 분류되는 경우가 발생하기 때문이며, 또 들영상의 경우는 (하늘/들) 영상 중 하늘이 상/중단에 걸쳐 있고 먹구름이 중단에 나타나면 산색과 겹치는 구름의 대표 빈화소가 더 높은 신뢰도를 갖는 산으로 잘못 레이블링 하는 경우가 발생하기 때문이다.

표 3. 영상분류 결과

영상종류	산영상	들영상	강영상
분류도	87%	88%	96%

## V. 결론 및 향후 연구 과제

본 논문은 칼라-공간 히스토그램과 생성 규칙을 이용해 자연 영상을 레이블링 하고 분류하는 방법을 제시하였다. 원 영상의 영역들을 빈 영상의 대표 빈화소로 표현된

다. 따라서 빈 영상의 대표 빈화소의 통계적인 데이터로 원 영상의 대응하는 영역들의 정보를 추출할 수 있기 때문에 이들을 생성 규칙에 적용해서 분할 된 영역들에 레이블링하고 분류하는 것이 가능해진다. 실제의 자연 영상들을 대상으로 하여 제안하는 방법이 좋은 결과를 보여주는 것을 알 수 있었다. 추후 연구과제로는 자연 영상 클래스 내에 존재하는 색과 공간관계 뿐 아니라 다른 특징 정보들을 추출하여 이들을 생성 규칙에 사용하는 것이 필요하다.

## 참고문헌

- [1] E. R. Hancock and J. Kittler, "Discrete relaxation", Pattern Recognition, Vol. 23, No. 7, pp. 711-733, 1990.
- [2] 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.
- [3] M. Kamada, K. Toraiichi, R. Mori, K. Yamamoto, and H. Yamada, "Parallel architecture for relaxation operations", Pattern Recognition, Vol. 21, No. 2, pp. 175-181, 1988.
- [4] 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.
- [5] Szummer, M. and Picard, R.W., "Indoor-outdoor image classification, Content-Based Access of Image and Video Database," 1998. Proceedings., 1998 IEEE International Workshop on, pp. 42 -51
- [6] Ratan, A.L., Maron, O., Grimson, W.E.L. and Lozano-Perez, T., "A Framework for Learning Query Concepts in Image Classification," Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on. , 1999 -429 Vol. 1
- [7] Smith, J.R. and Li, C.S., "Decoding Image Semantics Using Composite Region Templates," Content-Based Access of Image and Video Libraries, 1998. Proceedings. IEEE Workshop on, 1998, pp. 9 -13
- [8] Lipson, P., Grimson, E. and Sinha, P., "Configuration Based Scene Classification and Image Indexing" 1997. Proceedings. 1997 IEEE Computer Society Conference on, pp. 1007 -1013
- [9] 신수연, 김우생, "칼라-공간 히스토그램의 통계정보를 이용한 자연영상의 영역분할 및 레이블링 기법, 한국멀티미디어 춘계 학술발표 논문집, 2002.
- [10] Lei, Z., Fuzong, L. and Bo, Z., "A CBIR Method Based on Color-Spatial Feature," 1999 IEEE TENCON, pp. 166-169, 1999.