

색 변화 특징을 이용한 자연이미지에서의 장면 텍스트 추출

송영자* · 최영우

숙명여자대학교 컴퓨터과학과

Scene Text Extraction in Natural Images Using Color Variance Feature

Youngja Song* · Yeongwoo Choi

Dept. of Computer Science, Sookmyung Women's University

E-mail: soyo303@cs.sookmyung.ac.kr

Abstract

Texts in natural images contain significant and detailed informations about the images. Thus, to extract those texts correctly, we suggest a text extraction method using color variance feature. Generally, the texts in images have color variations with the backgrounds. Thus, if we express those variations in 3 dimensional RGB color space, we can emphasize the text regions that can be hard to be captured with a method using intensity variations in the gray-level images. We can even make robust extraction results with the images contaminated by light variations. The color variations are measured by color variance in this paper. First, horizontal and vertical variance images are obtained independently, and we can find that the text regions have high values of the variances in both directions. Then, the two images are logically ANDed to remove the non-text components with only one directional high variance. We have applied the proposed method to the multiple kinds of the natural images, and we confirmed that the proposed feature can help to find the text regions that can be missed with the following features - intensity variations in the gray-level images and/or color continuity in the color images.

요 약

이미지에 포함되어 있는 텍스트들은 이미지의 내용

을 함축적이며 구체적으로 표현하는 정보를 갖고 있다. 본 논문에서는 이러한 정보를 정확히 추출하기 위해서 색 변화 특징을 이용한 텍스트 영역 추출 방법을 제안한다. 관찰에 의하면 이미지 내의 텍스트들은 주변 배경과의 색 변화가 존재하며, 이러한 색 변화를 3차원 RGB공간에서 표현한다면, 명도이미지에서의 밝기 변화에서 표현하기 어려운 영역들을 강조시킬 수 있으며, 조명 변화에도 민감하지 않은 결과를 만들어 낼 수 있다. 색 변화 정도는 3차원 RGB 공간에서의 색 분산(Variance)으로 측정한다. 처리 과정으로서 우선 수평 및 수직 방향의 분산 이미지를 구하는데, 텍스트 영역은 두 방향의 분산 값이 모두 높은 특징이 있다. 다음으로 두 결과의 논리적 AND 연산을 수행하여 불필요한 잡영들을 제거한 후 연결요소를 분석, 검증하여 영역을 최종 확정한다. 다양한 종류의 자연이미지로 제안한 방법을 검증한 결과 밝기 변화 또는 색 연속성 특징들을 이용한 방법에서 찾기 어려운 텍스트 영역들을 찾을 수 있는 것을 확인할 수 있었다.

I. 서론

이미지 및 동영상에 인위적으로 또는 자연적으로 포함된 텍스트는 이미지의 내용을 함축적이고 구체적으로 표현하는 중요한 정보들이다. 이러한 정보들을 실시간에 추출하여 인식할 수 있다면 무인 자동차 운행, 시각 장애인 보행 안내 등과 같이 유용하게 사용될 수 있을 것이다[1]. 특히 자연이미지에 포함되어 있는 텍스트 영역의 추출은 자막과는 달리 배경에 다양한 종류의 잡영을 포함하고 있으며, 텍스트의 위치, 크기, 색상 등이 자유롭게 변하여 추출에 어려움이 있다. 대부분의 관련 연구들은 문서이미지에서의 텍스트 영역 추출과 제한된 환경에서 얻어진 이미지에서의 장면 텍스트 추출 연구로 생각할 수 있다. 최근에는 복잡한 배경과 다

양한 종류의 잡영이 포함되어 있는 이미지에서의 장면 텍스트 추출에 관한 연구도 활발하게 진행되고 있다 [2,3].

밝기 변화를 이용한 텍스트 영역 추출 방법으로서 Johnson이[4] 제안한 영역화 알고리즘의 RLSA(Run Length Smearing Algorithm)가 있다. 이 방법은 서로 인접한 검은 픽셀을 연결한 후 수직 및 수평 방향으로 알고리즘을 적용한 결과를 결합함으로써 영역을 찾는다. Ohya는[5] 지역적 이진화 방법과 인접 영역의 명도차를 고려한 텍스트 영역 추출 방법을 제안하였으나, 이미지 전체에 나타나는 불필요한 에지 값들 때문에 다양한 종류의 재처리 작업을 필요로 하는 단점이 있다. Zhong은[6] 텍스트 영역은 명도 값의 공간적 변화가 크다는 특징을 이용하기 위하여 수평 방향의 분산 이미지를 구하는 방법을 제안하였다. 이 방법은 세로 방향의 문자열은 찾지 못하는 경우가 많았고, 명도이미지의 1차원적인 Gray 값을 사용하였기 때문에 3차원 RGB 공간에서의 변화를 적절하게 반영하지 못하는 단점이 있다. Gu는[7] Differential top-hats 연산을 이용한 텍스트 추출 방법을 제안하였다. 문자 획을 크기에 따라 8단계로 나누고, 흰 문자와 검은 문자로 구분하여 연산자를 적용하였다. 이 방법은 문자열을 비교적 정확히 찾지만 많은 수행 시간을 필요로 하는 단점이 있다.

본 논문은 RGB 3차원 공간의 색 변화 값을 특징으로 이용한 추출 방법을 제안하여 배경과 텍스트의 명도 값이 비슷하여 밝기 변화가 잘 나타나지 않는 영역도 찾을 수 있으며, 수평 및 수직 방향의 색 분산 결과에 논리적 AND 연산을 수행하여 불필요한 잡영을 쉽게 제거하는 방법도 함께 제안한다.

II. 색 변화 특징

동영상에 포함된 자막이나 문서 이미지의 텍스트대부분은 텍스트 색과 주변 색을 보색으로 이용하거나 텍스트 주변을 단순하게 하여 텍스트와 배경을 비교적 뚜렷하게 구분하였다. 그러나 거리에서 취득한 간판이나 표지판 이미지와 같은 자연이미지는 빛과 조명의 영향으로 또는 인위적인 의도에 의해서 텍스트와 배경의 구분이 명확하지 구분되지 않는 경우도 빈번하게 발생한다. 이러한 경우 색 변화를 특징으로 이용한다면 보다 정확한 영역 추출을 기대할 수 있다. 이 방법은 밝기 변화를 이용한 텍스트 영역 추출 방법이 조명의 변화에 따라 에지 추출의 오류가 쉽게 발생하여 생기는 오검출의 빈도수를 줄일 수 있다. 그림 1(a)는 밝기 변화를 이용한 텍스트 영역 추출 결과이며, (b)는 색 변화를 분산 값을 이용한 추출한 결과로서 (a)보다 개선된 결

과를 보여준다.

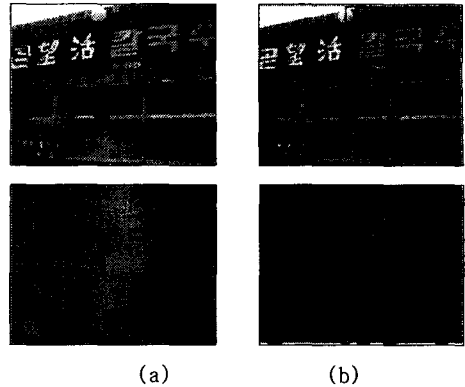


그림 1. 영역 추출 결과 비교: (a) 밝기 변화를 이용한 추출, (b) 색 분산을 이용한 추출

밝기 또는 색 변화를 분산으로 표현할 수 있다. 명도이미지의 Gray 값을 이용하여 분산을 구한다면 문서 이미지처럼 글자와 배경의 구분이 뚜렷한 경우 분산 값이 크게 나타나지만, 명도 값의 변화가 작은 경우 분산 값이 작게 나타난다. 그러나 RGB로 표현된 색 정보는 3차원 RGB 공간에서 픽셀 값이 표현되어 분산 값이 크게 구해지는 장점이 있다. 1차원 명도이미지에서 분산 값이 작은 이유는 3차원 공간에서 멀리 나타날 수 있는 픽셀들의 위치가 동일 선상에 선형적으로 투영되어 색 변화 값이 축소되기 때문이다. 따라서 색 이미지에서의 분산 값은 명도이미지에서의 분산 값보다 크고 변화를 정확하게 측정할 수 있다. 그림 2(b)는 주어진 이미지(a)에서 명도 값을 이용하여 분산을 계산하여 영역을 추출한 결과이며, (c)는 RGB 색 값으로 분산을 계산하여 영역을 추출한 결과로서 글자와 배경의 변화가 보다 명확하게 표현된 것을 확인할 수 있다.

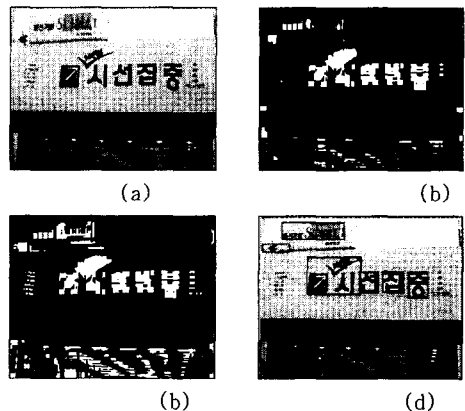


그림 2. 명도 분산과 색 분산 비교: (a) 원이미지, (b) 명도 분산 이미지, (c) 색 분산 이미지, (d) 색 분산으로 찾은 글자 영역

III. 색 분산을 이용한 텍스트 영역 추출

색 분산을 이용한 텍스트 영역 추출 방법은 수평 및 수직 방향의 색 분산 이미지를 먼저 구한 후, 두 이미지의 논리적 AND 연산을 수행한다. 다음으로 모폴로지를 이용하여 후보영역을 강조하고, 연결요소의 분포 및 크기 등을 고려하여 추출된 후보 영역을 검증한다.

3.1 수평 및 수직 분산 이미지 생성

320x240 크기의 입력된 색 이미지에 대해서 1*21 크기의 수평 방향 원도우와 17*1 크기의 수직 방향 원도우를 각각 설정하여 분산 값을 구한다. 이러한 원도우의 크기는 실험에 의해서 결정하였다. 구해진 분산 값은 식(1)을 이용하여 값의 범위를 0과 255 사이로 정규화시키며 여기서 max_1 과 min_1 은 분산값 v_{new} 의 최대, 최소 값, max_2 와 min_2 는 새로운 범위의 밝기값을 의미한다. 그림 3(b)와 (c)는 정규화된 수평 및 수직 방향 분산이미지의 결과를 보여준다. 원 이미지 (a)에서 색 변화가 큰 부분은 밝게 표현되고, 색 변화가 작은 부분은 어둡게 나타난다. 텍스트 영역은 수평 및 수직 분산 값이 모두 크게 나타나는 특징을 찾을 수 있다.

$$v_{normalized} = \left(\frac{v_{new} - min_1}{max_1 - min_1} \right) (max_2 - min_2) + min_2 \quad (1)$$

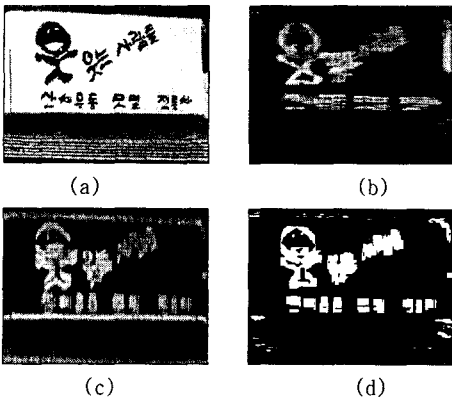


그림 3. 수평 및 수직 분산 이미지: (a) 원 이미지, (b) 수평 분산 이미지, (c) 수직 분산 이미지, (d) 논리적 AND 연산 결과

3.2 AND 연산을 이용한 잡영 제거

텍스트 영역은 수평 및 수직 방향 분산이미지 모두에서 분산 값이 크게 나타나지만, 단순한 선들과 같은 불필요한 요소들은 수평 또는 수직 한 이미지에서만 분산 값이 크게 나타나는 특징이 있다. 따라서 두 분산 이미지를 대상으로 논리적 AND 연산을 수행한다면 선과

같은 불필요한 정보들을 쉽게 제거할 수 있다. 논리적 AND 연산을 수행하기 위해서 먼저 이진화를 수행한다. 이진화는 정규화된 분산 이미지의 평균 값을 이용하여 임계값을 결정하는 방법을 사용하였다. 그림 3(d)는 두 이미지의 논리적 AND 결과로서 명도이미지의 에지로 표현한다면 포함될 수 있는 많은 선들이 제거된 결과를 볼 수 있다.

3.3 영역 추출 및 검증

논리적 AND 연산을 수행한 결과 이미지에는 분산 값이 높은 색 변화가 큰 영역들이 남게 되지만 불필요한 잡영도 부분적으로 포함된다. 따라서 텍스트 영역의 처리를 쉽게 하고 간단한 잡영들을 제거시키기 위해서 모폴로지 연산을 수행한다. 모폴로지 연산은 수평 방향의 텍스트를 주로 강조하기 위해서 2*5 크기의 Dilation 연산자와 3*2 크기의 Erosion 연산자를 구성하여, 식(2)와 같이 열림(Opening)과 닫힘(Closing)연산을 반복적으로 적용한다.

$$I = Closing + Opening + Opening \quad (2)$$

그림 4(a)는 그림 3(d)의 AND 연산 결과에 제안한 모폴로지 연산자를 적용한 결과로서, 적용 전의 이미지보다 잡영은 제거되고 텍스트 영역은 강조되며 연결요소의 수가 줄어든 것을 확인할 수 있다. (b)는 모폴로지 결과에서 연결요소와 연결 요소의 외곽사각형을 구한 결과이며, 외곽사각형의 위치, 크기, 종횡비(Aspect ratio)와 연결요소의 크기 등이 저장된다.

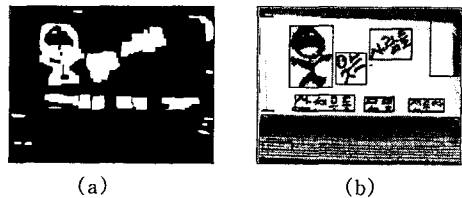


그림 4. (a) 모폴로지 적용, (b) 연결요소의 외곽사각형 추출

그림 4(b)에서 텍스트 영역이 아니라고 판단되는 후보 요소들을 경험적인 규칙에 의해서 먼저 제거한다. 배경과 같이 외각사각형의 크기가 임계값 이상으로 큰 것을 제거하고, 영역의 크기가 너무 작아서 식별할 수 없는 후보영역이 생기지 않도록 임계값 이하의 크기를 갖는 외각사각형도 제거한다. 또한, 외각사각형의 가로에 대한 세로의 비, 외각사각형을 차지하는 연결요소의 비율에 제한을 두어 임계값 이하로 작거나 크면 텍스트 후보영역에서 제외한다.

다음으로 문자 단위의 의미 있는 외곽사각형을 구성하기 위해서 외곽사각형을 결합한다. 결합은 외곽사각형들의 포함관계, 외곽사각형 사이의 거리 및 X, Y축 방향으로의 투영의 겹침 정도를 이용한다. 추출된 텍스트 후보 영역들의 일부는 문자열이 하나의 외곽사각형으로 존재하지만 글자 단위로 존재하는 것도 있기 때문이다. 그림 5는 외곽사각형이 결합되어 최종적으로 얻어진 텍스트 영역을 보여준다.

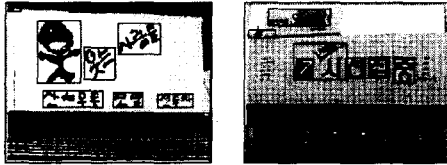


그림 5. 추출된 텍스트 영역

IV. 실험 및 결과

학교, 병원, 지하철, 간판 등의 실내외에서 취득한 110개의 자연이미지를 대상으로 Visual C++ 6.0을 이용하였다. 320× 240 크기의 대상 이미지에는 다양한 색, 크기, 폰트, 방향 등의 문자열이 포함되어 있다. 평가를 위해서 수평 방향으로 글자 간격이 너무 큰 것과 세로 방향의 텍스트는 문자 단위로 구분하여 텍스트 영역으로 설정하였다.

실험 결과는 표 1과 같으며, 표 1에서의 Total은 실험 대상 이미지에서의 전체 텍스트 수이며, Correct는 정확히 찾은 텍스트 개수, Partial은 텍스트의 일부분만을 찾은 개수, Missing은 텍스트를 찾지 못한 개수이며, False는 텍스트 영역이 아닌 영역을 텍스트로 잘못 찾은 개수이다. Precision = (Total number of correct texts)/(Number of detected texts)이며, Recall = (Total number of correct texts)/(Total number of texts)로 계산하였다.

표 1 영역 추출 실험 결과

		Total	Correct	Partial	Missing	False
텍스트 이미지	영역 개수	677	591	15	47	394
		Precision = 59.1%		Recall = 87.3%		

표 1과 같이 다양한 종류의 실외 이미지에 대해 실험한 결과 조명의 영향이 있는 이미지에서도 색 변화 특징을 이용하여 비교적 정확히 텍스트 영역을 추출한 것을 알 수 있다. 그림 6은 추출 결과 이미지를 보여주고 있다. 결과에서 비텍스트 영역을 텍스트 영역으로 잘못 찾은 오류는 또 다른 검증으로 해결할 수 있을 것이다.

V. 결론

본 논문에서는 색 변화 특징을 이용한 장면 텍스트 추출 방법을 제안하였다. 제안한 알고리즘은 조명에 영향을 받은 이미지에 대해서도 비교적 정확하게 텍스트 영역을 추출할 수 있었는데, 이는 RGB 3차원 공간에서의 분산 값을 이용함으로써 밝기 변화로 표현하기 어려운 영역들을 추출할 수 있었기 때문이다. 향후 연구에서는 밝기 변화, 색 연속성을 이용한 추출 방법과의 결합 및 검증 방법을 개발하여 추출 성능을 높이는 것이다.

* 본 연구는 과학기술부 뇌신경정보학사업의 지원을 받아 수행되었습니다.(M1-0107-00-0009)

VI. 참고문헌

- [1] H.R Byun, M.C. Roh, K.C.Kim, Y.W.Choi, S.W.Lee, "Scene Text Extraction in Complex Images," Lecture Note and Computer Science, Vol. 2423, pp. 329-340, 2002.
- [2] R. Lienhart, F. Stuber, "Automatic Text Recognition in Digital Videos," In Image and Video Processing IV 1996, Proc. SPIE 2666-20 (1996).
- [3] Y. Zhong, H. Zhang and A. K. Jain, "Automatic Caption Localization in Compressed Video," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 4, pp. 385-392, 2000.
- [4] E.G.Johnston, Printed text discrimination, Comp. Graphics Image Processing 3(1) (1974)83-89.
- [5] J. Ohya, A. Shio, S. Akamatsu, "Recognizing Characters in Scene Images," IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-16(2), pp. 67-82, 1994.
- [6] Yu Zhong, Kalle Karu, Anil K. Jain, "Locating Text in Complex Color Images," Pattern Recognition, Vol. 28, No. 10, pp. 1523-1535, 1995.
- [7] Lixu Gu, Toyahisa Kaneko, "Robust extraction of Characters from color scene Image Using Mathematical morphology", Proc. of 7th Int. Conf. on Pattern Recognition, Vol. 2, pp. 1002-1004.

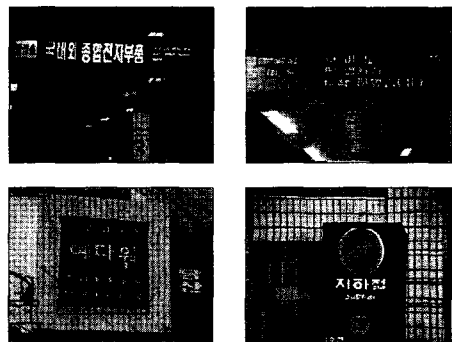


그림 6. 추출 결과 예