

색 분산 특징을 이용한 텍스트 추출에서의 손실된 분산 복원

최영우*, 조은숙**

Variance Recovery in Text Detection using Color Variance Feature

Yeongwoo Choi *, Eunsook Cho **

요약

본 논문은 자연이미지에 포함된 텍스트 영역을 찾기 위한 방법으로서 기존에 제안한 색 분산 특징을 이용한 방법에서 분산이 제대로 추출되지 않는 문자 획들에 대한 복원 방법을 제안한다. 이전의 색 분산 특징을 이용한 추출 방법에서는 고정된 크기의 수평 및 수직 분산 추출 윈도우를 사용함으로써 문자 획이 두껍거나 긴 경우에는 색 분산이 제대로 추출되지 않는 단점이 있었다. 따라서 본 논문에서는 미 추출된 색 분산을 연결요소 외곽사각형의 기하학적인 정보와 경험적인(Heuristic) 지식을 함께 이용하여 복원하는 방법을 제안한다. 제안한 방법은 다양한 종류의 디지털 카메라와 휴대폰 카메라를 이용해서 취득한 문서 유형의 이미지와 간판, 거리 표지판 등의 자연이미지를 사용하여 테스트하였으며, 특히 큰 글자를 포함하는 자연이미지에 대해서도 텍스트 추출의 정확성이 향상된 것을 확인할 수 있었다.

Abstract

This paper proposes a variance recovery method for character strokes that can be missed in applying the previously proposed color variance approach in text detection of natural scene images. The previous method has a shortcoming of missing the color variance due to the fixed length of horizontal and vertical windows of variance detection when the character strokes are thick or long. Thus, this paper proposes a variance recovery method by using geometric information of bounding boxes of connected components and heuristic knowledge. We have tested the proposed method using various kinds of document-style and natural scene images such as billboards, signboards, etc captured by digital cameras and mobile-phone cameras. And we showed the improved text detection accuracy even in the images of containing large characters.

▶ Keyword : 이미지 처리(Image processing), 장면 텍스트 추출(Scene text detection), 색 분산 특징(Color variance feature), 획 복원(Stroke recovery)

• 제1저자 : 최영우

• 투고일 : 2009. 08. 12. 심사일 : 2009. 08. 29. 게재확정일 : 2009. 10. 12.

* 숙명여자대학교 컴퓨터과학과 교수 ** 숙명여자대학교 컴퓨터과학과 석사

※ 본 연구는 숙명여자대학교 2006년도 교내연구비 지원에 의해 수행되었음

1. 서론

최근 고도화된 정보화 사회로의 진입에 따라 PC용 카메라, 디지털 카메라, 휴대폰 카메라 및 디지털 캠코더 등을 이용한 이미지 및 동영상 취득이 아주 쉽게 이루어지고 있다. 이러한 영상 내에 포함되어 있는 문자 정보를 정확히 인식할 수 있다면 시각 장애인을 위한 보행 보조, 로봇의 자동 주행, 비디오 자동 검색, 스팸 메일 필터 등과 같은 다양한 응용에 활용할 수 있다.

지금까지 이미지에서의 문자 추출에 관련된 연구는 OCR(Optical Character Recognition)을 위한 문서 이미지에서의 텍스트 추출 연구로부터 시작하여 정해진 환경 또는 제약이 없는 환경에서 취득한 자연이미지에서의 장면(Scene) 텍스트 추출 연구로 꾸준히 발전하고 있다[1-14]. 최근에는 복잡한 배경을 포함하는 자연이미지에서의 장면 텍스트 추출에 관한 연구가 활발하게 진행되고 있으며, 제안된 다양한 방법들에 대한 객관적인 비교도 체계적으로 진행되고 있다[1,2].

자연이미지에 포함된 텍스트를 찾기 위한 다양한 방법들이 제안되고 있으며, 본 논문에서는 이와 관련된 연구를 두 방향으로 나누어 소개한다. 텍스트 영역에서 화소의 밝기 변화와 일정한 명도 유지를 특징으로 사용하는 방법과 색 변화 정도 및 동일한 색 정보 유지를 특징으로 사용하는 방법으로 분류한다. 이는 본 논문에서 사용한 텍스트 영역의 특징이 밝기 또는 색 변화 정도를 측정하여 획의 조밀함을 표현하는 것으로서, 이미지 부분 영역의 밝기 또는 색 변화 정도를 색 분산 특징으로 표현하여 사용하였다.

밝기 값을 텍스트 특징으로 사용한 초기의 대표적인 연구로서 Ohya 등은[3] 이미지 내에서 동일한 영역의 텍스트들은 명도 값이 거의 일정하다는 특징을 이용하였다. 지역적 이진화 방법을 적용한 후 연결요소의 유사성을 계산한 후 완화(relaxation) 방법을 적용하여 텍스트 후보 영역의 요소들을 병합하였다. 텍스트의 배열이 기울어지지 않고 배경과 뚜렷하게 구분되는 비교적 단순한 이미지에서는 좋은 결과를 얻었지만, 복잡하거나 조명 변화가 심한 이미지에서는 오류가 많이 발생하는 단점을 보였다.

Jain 등은[4] 정지된 이미지, 웹 이미지 및 동영상을 대상으로 Multi-value 이미지 분해에 의한 지역적 편차 이미지를 구하여 BAG(Block Adjacency Graph) 알고리즘과 연결요소 분석을 이용하여 텍스트 영역을 추출하였다. 이 연구는 뒤에 수행된 많은 연구에 영향을 준 중요한 연구가 되었다. Ezaki 등은[5] 시각장애인을 위한 보행 보조시스템을 만들기 위해서 자연이미지에서의 텍스트 추출을 연구하여, 소벨

(Sobel) 에지 검출, 오투(Otsu) 이진화, 연결요소 추출 및 규칙기반 연결요소 필터링으로 구성된 텍스트 추출 방법을 제안하여, 단순한 이미지에서는 우수한 추출 결과를 만들었지만 복잡한 배경과 다양한 색으로 표현된 문자 영역의 추출은 고려하지 않은 단점이 있다.

Dubey는[6] 복잡한 이미지를 대상으로 에지 기반 텍스트 검출 방법을 제안하였다. 수직 소벨 에지 검출기를 적용하고, 새롭게 제안한 모폴로지 필터를 이용하여 에지들을 결합하고 후보 영역을 결정하였다. 제안한 방법은 텍스트 영역에 대하여 뚜렷한 질감(texture)을 제공하는 장점이 있으며, 다양한 입력 도구와 환경으로부터 취득한 여러 종류의 이미지를 대상으로 한 실험에서 높은 성공률을 보여주었다.

박종천 등은[7] 문자를 기반으로 한 에지 맵의 패턴 히스토그램을 고안하여 텍스트 영역을 추출하는 새로운 방법을 제안하였다. 8종류의 문자-에지 맵 특징을 이용하여 텍스트 후보 영역을 결정하고, 문자-에지 맵의 패턴 히스토그램과 텍스트 영역의 구조적 특징을 이용하여 후보 영역을 검증하였다. 복잡한 배경, 다양한 글꼴 및 문자 색으로 구성된 자연이미지에 제안한 방법을 적용하여 텍스트 영역을 효과적으로 추출하였다. Liu 등은[8] Niblack 알고리즘을 적용하여 이미지를 이진화한 후 기하학적인 특징을 사용하여 연결요소들을 결합하였다. 강도(intensity) 및 형태(shape) 필터를 함께 사용하여 연결요소를 검증한 후 후보 영역을 결정하였다. 249개의 ICDAR 2003 콘테스트에서 사용한 이미지를 대상으로 [1] 제안한 방법을 평가하였으며, 실내외 이미지 및 배경이 복잡한 이미지에 대해서도 우수한 연구 결과를 보고하였다.

텍스트 추출 특징으로서 색을 이용한 관련 연구로서 Zhong 등은[9] 자연스럽게 취득한 색 이미지를 대상으로 색 히스토그램을 계산한 후 비슷한 색 영역별로 분할 및 합병(Split and Merge) 알고리즘을 적용하여 전처리 과정에서 지워진 영역을 복원하며 텍스트 영역을 추출하는 방법을 제안하였다. 또한, Wang은[10] 색 이미지에 대해서 먼저 색 정보를 얻기 위해 다그룹분해(multi-group decomposition) 방식을 제안하여 각 색상 계층 이미지에 대한 처리와 명도를 이용한 처리를 병행하였다. 위의 두 방법은[9,10] 색 정보를 이용한 텍스트 추출 연구에 많은 공헌을 하였다.

배경숙 등은[11] 자연이미지에서 텍스트를 추출하기 위해서 RGB 색 공간에서 각각의 색 계층별 이미지에 대한 이진화를 수행하여 텍스트를 추출하는 방법을 제안하였다. 다해상도 웨이블릿(Wavelet) 변환과 SVM(Support Vector Machine)을 이용하여 추출된 텍스트 영역을 검증하였다. 제안한 방법은 밝기 변화를 이용한 텍스트 추출 방법의 단점을

어느 정도 극복할 수 있음을 보여주었는데, 이는 화소 수준에서 표현하기 어려운 문자의 구조적인 정보를 문자의 획 수준으로 검증하였기 때문에 가능하였다.

김지수 등은[12] 스팸메일 이미지에 포함된 텍스트를 추출하기 위해서 색상 계층 기반 텍스트 추출 알고리즘을 제안하였다. 이미지를 8개의 색상 계층으로 나눈 후 각 이미지의 연결요소를 추출하여 텍스트 영역을 추출하였다. 또한 추출된 영역의 이진화 과정에서 빈번하게 발생하는 획 소멸과 자소 안이 매워지는 문제를 해결하는 방법도 함께 제안하여 이진화의 어려움을 부분적으로 해결하였다. 200개의 스팸메일 이미지를 사용한 실험에서 제안한 알고리즘의 우수성 및 실용화 가능성을 보여주었다.

송영자 등은[13] 다양한 종류의 자연이미지로부터 색 분산을 이용한 텍스트 추출 방법을 제안하였다. 이 과정에서 고정된 크기의 가로 및 세로 방향의 분산 추출 윈도우를 사용하여 분산 이미지를 만든 후 논리적 AND, 모폴로지 연산 및 영역 검증 과정을 거치면서 다양한 종류의 자연이미지에 대한 실험에서 비교적 좋은 결과를 보고하였다. 또한 추출된 영역에서 문자를 분리하는 방법으로서 일반적으로 사용하는 전역적, 지역적 이진화 방법 대신에 K-means 클러스터링 방법을 개선한 문자 분리 방법을 제안하여 문자 분리의 정확성을 향상시켰다.

Lee 등은[14] 간판과 같은 이미지를 대상으로 모바일 디바이스에서 사용하기 위한 텍스트 영역 검출 및 문자 분리 방법을 제안하였다. 이미지의 수평 에지 프로파일과 한글이 갖는 특징을 지식화하여 텍스트 영역을 찾았으며, 수직 에지 프로파일을 이용해서 단어 단위로 분리하였다. L^*a^*b 색 공간에서 3-means 클러스터링 방법으로 이진화를 수행하여 문자를 분리하였다. 비교적 복잡하지 않은 다양한 종류의 간판 이미지에 제안한 방법을 적용하여 좋은 결과를 얻었으며, 제한된 용량의 모바일 디바이스에서 유용하게 사용될 수 있음을 보여주었다.

본 논문은 [13]에서 제안한 색 분산 추출 과정에서 발생하는 미 추출된 색 분산의 복원 방법을 제안한다. 색 분산을 추출하기 위해서 1×21 크기의 수평 및 17×1 크기의 수직 방향의 고정된 윈도우를 사용하였는데, 간판이미지와 같이 큰 글자를 포함하는 이미지에서는 색 분산을 측정하는 윈도우가 문자 획의 폭 또는 길이와 비슷하거나 짧아져서 분산이 찾아지지 않거나 낮은 분산 값으로 찾아지는 경우가 빈번하게 발생하였다. 따라서 본 논문에서는 손실된 분산 정보를 획 단위로 복원하는 방법을 제안하여 색 분산 특징의 정확성을 높이고자 하였으며, 이는 일반적인 크기의 글자를 포함하는 이미지뿐만 아니라 큰 글자를 포함하는 이미지에 대해서도 이전보다 안정적인 검출 결과를 얻기 위해서다.

본 논문의 구성으로 2절에서는 이전에 제안한 색 분산을 이용한 텍스트 추출 방법을 간단히 소개한 후 본 논문의 출발점인 분산 추출의 문제점을 제시한다. 3절에서는 이러한 문제점을 해결하기 위한 분산 복원 방법을 제안하고, 추출된 영역의 제거 및 결합 방법의 일부 개선된 사항을 추가로 설명한다. 4절에서는 다양한 종류의 이미지에 대한 실험 결과를 언급하며, 5절에서 결론 및 향후 연구를 기술한다.

II. 기 제안한 색 분산 추출 및 문제점

여기서는 이전에 제안한 색 분산 추출 방법 및 텍스트 추출 과정을 전반을 간략히 소개하고, 이 과정에서의 문제점을 언급한다. 색 분산을 이용한 텍스트 영역 추출에 관한 자세한 내용은 논문 [13]을 참조하기 바란다.

2.1 색 분산을 이용한 텍스트 영역 추출

카메라로 취득한 이미지를 320×240 크기로 규격화한 후, 1×21 크기의 수평 윈도우와 17×1 크기의 수직 윈도우를 설정하여 식 (2.1)을 적용하여 각 픽셀의 수평 및 수직 분산을 구한다.

$$v_H = \frac{1}{21} \sum_{i=0}^{20} ((R_i - \mu_{HR})^2 + (G_i - \mu_{HG})^2 + (B_i - \mu_{HB})^2)$$

$$v_V = \frac{1}{17} \sum_{i=0}^{16} ((R_i - \mu_{VR})^2 + (G_i - \mu_{VG})^2 + (B_i - \mu_{VB})^2)$$

..... (2.1)

여기서 v_H 와 v_V 는 수평 및 수직 방향의 분산이며, R, G, B는 각 픽셀에 대한 색 값이다. μ_{HR} 은 R 색상의 수평방향 평균이며, μ_{VR} 은 R 색상의 수직방향 평균이다. 구해진 분산을 0과 255 사이의 값으로 재 분포시킨 후 분산이미지의 평균을 임계값으로 설정하여 전역적 이진화를 수행한다[15,16].

〈그림 1〉은 두 개의 이미지 샘플에 대해서 위의 식을 적용하여 얻은 분산이미지로서 텍스트 영역을 찾는 과정을 보여준다. 〈그림 1(b)〉와 〈그림 1(c)〉는 식 (2.1)을 적용한 수평 및 수직 분산이미지로서 텍스트 영역에 해당되는 부분은 분산이 커서 밝게 표현되며, 단순한 선들은 수평 또는 수직의 한 방향으로만 분산이 크게 나타난다. 두 분산 이미지의 논리적 AND 연산을 수행하여 〈그림 1(d)〉와 같이 선이나 불필요한 요소들의 일부를 제거한다. 논리적 AND 연산이 적용된 결과에 모폴로지 연산을 적용하여 간단한 잡영들을 추가로 제거한다. 모폴로지 연산은 수평 방향의 텍스트를 주로 강조하기 위

해서 2x5 크기의 불림연산자와 3x3 크기의 녹임 연산자를 적용한 opening 연산으로 진행한다. <그림 1(e)>는 후보 영역들에 대하여 휴리스틱 규칙을 적용한 후 최종적으로 텍스트 영역으로 판정된 영역들을 보여준다.

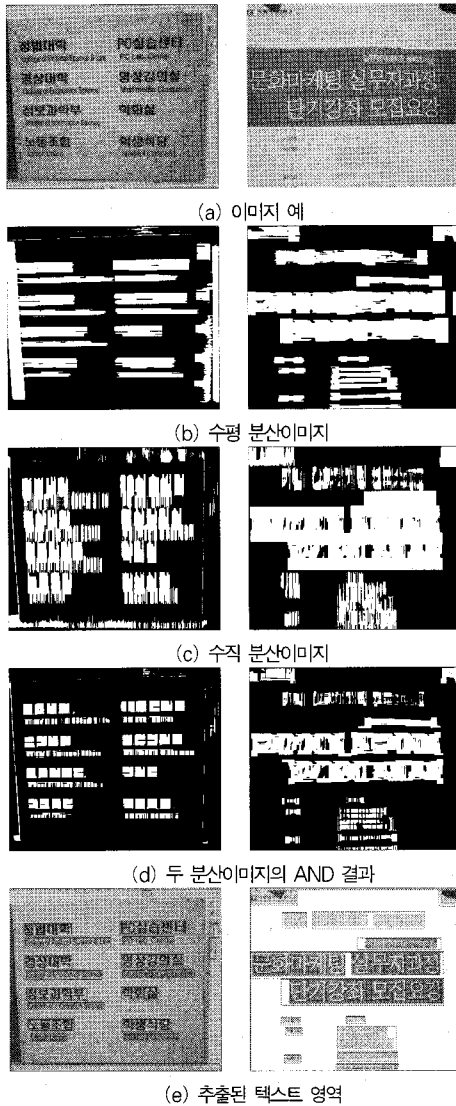


그림 1. 분산이미지 생성 및 추출된 텍스트 영역
Fig. 1. Variance Images and Detected Text Regions

2.2 미 추출된 분산

색 분산을 추출하는 과정에서 이미지에 포함된 글자의 가로 또는 세로획이 긴 경우에는 분산을 측정하는 윈도우의 길이가 문자 획의 길이보다 작아져서 문자 획 안에 윈도우가 완

전히 포함되거나 대부분 포함되어 색 변화가 거의 발생하지 않는 경우가 발생한다. 이는 분산이 검출되지 않거나 낮게 검출되는 결과를 만든다. 따라서 문자 획이면서도 분산이 검출되지 않는 경우로서 큰 글자를 포함하는 이미지에서 빈번하게 발생하여, 결국 텍스트 영역임에도 불구하고 분산이 검출되지 않거나 낮게 검출되어 영역을 찾지 못하는 결과를 만든다.

<그림 2(a)>는 큰 글자를 포함하는 이미지 예이며, <그림 2(b)>는 1x21 크기의 수평 윈도우를 적용하여 분산을 구한 결과로서 '드', '핀', '피', '외', '환', '은' 등의 긴 수평 획에서 분산이 추출되지 않거나 낮게 추출되었지만 분산 값의 이진화 과정에서 손실된 것을 보여준다. 또한 <그림 2(c)>는 17x1 크기의 수직 윈도우를 사용한 분산 결과로서 '레', '핀', '피', '자', '외', '환' 등의 긴 수직 획에서 분산이 나타나지 않음을 보여준다. 이는 문자 영역임에도 불구하고 분산이 나타나지 않은 결과로서 큰 글자를 포함하는 이미지에서 빈번하게 발생한 텍스트 추출 오류의 주원인이다.

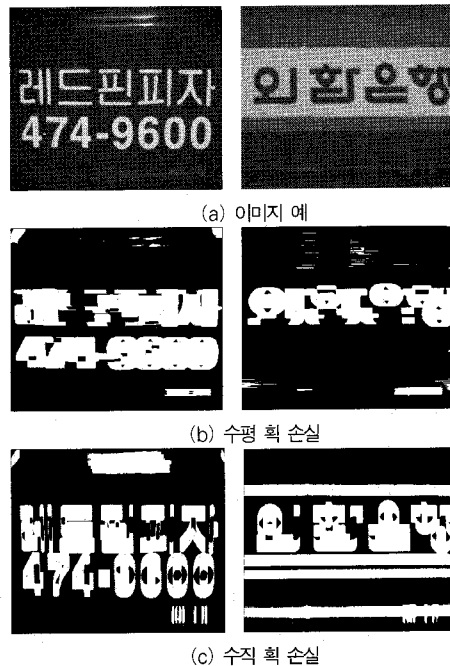


그림 2. 수평 및 수직 분산 정보 손실
Fig. 2. Loss in Horizontal and Vertical Variances

III. 손실된 색 분산 복원 및 영역 확장

여기서는 색 분산이 제대로 추출되지 않는 것을 개선하기 위하여 연결요소의 기하학적인 배치와 경험적인 지식을 함께 이용

해서 손실된 분산을 복원하여 문자의 손실된 획을 복원시키는 방법을 제안한다. 또한, 후보 영역을 확정하기 위해서 후보 영역의 제거 및 결합 과정에서 새롭게 개선한 방법을 설명한다.

3.1 분산 복원

〈그림 3〉은 손실된 분산의 복원 과정으로서 먼저 외곽사각형을 검사해서 복원 검토 외곽사각형을 선택한 후 수평 및 수직 복원 순서로 손실된 분산을 복원한다.

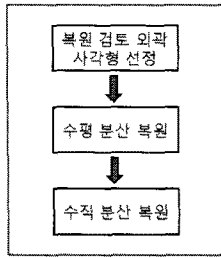


그림 3. 분산 복원 과정
Fig. 3. Procedures in Variance Recovery

전체 텍스트 추출 과정에서의 분산 복원은 모폴로지 연산이 수행된 이진이미지에 간단한 잡영들을 제거한 후 적용한다. 이진이미지는 각 연결요소의 외곽사각형으로 표현되며, 다양한 크기 및 형태로 구성된 사각형들로부터 복원할 외곽사각형을 찾는 것이 중요하다. 복원 검토 외곽사각형의 중요 특징은 수평 또는 수직 방향으로 분산 윈도우의 크기를 고려한 일정한 거리 이내에 마주보는 외곽사각형의 존재 여부이다. 또한 복원 검토 외곽사각형의 넓이는 대부분 일정한 값 이하이며, 문자 획의 일부로서 외곽사각형에 포함된 연결요소의 면적 비율이 전반적으로 크게 나타나는 특징이 있다. 따라서 복원 검토 외곽사각형은 면적이 일정한 값 이하로서 외곽사각형내의 연결요소의 비율이 일정한 값 이상이며, 수평 복원에서는 좌우, 수직 복원에서는 상하로 일정한 거리 이내에 마주보는 외곽사각형이 존재할 때 선정한다. 다양한 종류의 이미지를 대상으로 한 실험에서 찾은 임계값을 적용했으며, 외곽사각형의 면적은 380 픽셀 이하, 외곽사각형내의 연결요소의 면적 비율은 60% 이상, 수평 및 수직 복원에서의 좌우 및 상하 거리는 분산 윈도우 크기의 약 2배인 40 픽셀 이내에 다른 외곽사각형이 존재할 때 복원 검토 외곽사각형으로 선정한다.

선정된 외곽사각형의 복원은 아래와 같이 수행한다. 수평 분산이미지에서 선정된 외곽사각형은 동일한 위치의 수직 분산이미지의 분산 존재 여부를 검사하여 수직 분산이미지에 분산이 존재하면 이는 가로획에서 분산이 손실된 것으로 판정하

여 복원한다. 수직 복원도 선정된 외곽사각형에 대하여 동일한 방법을 적용하여 수평 분산이미지의 분산 존재 여부를 검사하여 세로획을 복원한다.

먼저 수평 복원은 수직 분산이미지를 참조하여 손실된 획 정보를 찾는다. 〈그림 4(a)〉에서 빨간색으로 표시한 왼쪽 원안의 외곽사각형이 복원 검토 대상으로 선정된 외곽사각형이며, 〈그림 4(b)〉의 파란색 원과 같이 선정된 텍스트 영역의 좌우 방향으로 수직 분산의 존재 여부를 확인하며 이 경우에는 좌측 방향으로 분산 정보를 복원한다. 복원은 손실된 획이라고 판정된 두 개의 사각형을 대상으로 동일한 영역의 동일한 방향으로 수직 분산이미지의 채워진 부분을 참조하며 채울 영역을 결정하여 복사한다. 수평 획의 두께를 이용하여 복원하는 영역의 두께로 사용한다. 〈그림 4(a)〉에서 오른쪽 빨간색 원도 복원 검토 대상으로 선정되었으며 〈그림 4(b)〉의 수직분산 이미지에서 좌우로 분산 정보의 존재를 확인하며, 여기서는 좌우로 분산 정보가 존재하지 않아서 복원하지 않은 결과이다. 〈그림 4(c)〉는 수평 복원 결과로서 '드', '핀', '피'에서 손실된 획의 분산이 복원된 것을 확인할 수 있다.

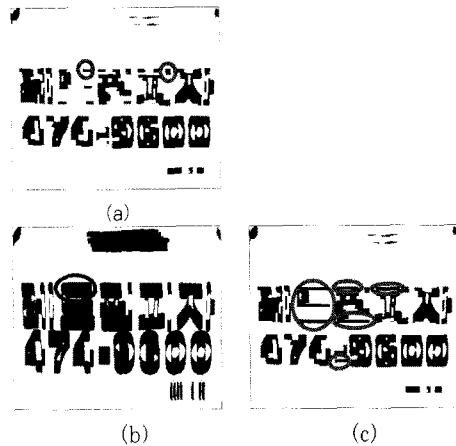


그림 4. 수평 복원
Fig. 4. Horizontal Recovery

수직 복원은 수평 복원을 수행한 결과 이미지에 대하여 앞에서 언급한 방법에 의해서 복원 검토 대상 외곽사각형을 선택하고 수평 분산 이미지를 참조하여 손실된 정보를 찾는다. 〈그림 5(a)〉는 수평 복원된 결과 이미지로서 빨간 원으로 표시된 영역이 위의 조건에 의해서 복원 검토 대상으로 선택된 영역의 예이다. 선택된 영역의 상하 방향으로 분산 정보의 존재를 〈그림 5(b)〉의 수평분산 이미지를 참조하여 결정한다. 〈그림 5(b)〉에서 파란색 원은 위쪽으로는 분산이 존재하지 않지만 아래쪽으로는 분산이 존재하여 분산이 존재하는 방향

으로 복원하며, <그림 5(c)>는 '레', '핀', '피', '자'에서 손실된 수직 획을 복원한 결과를 보여준다.

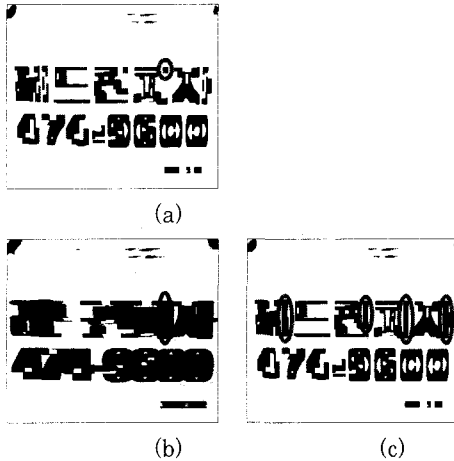


그림 5. 수직 복원
Fig. 5. Vertical Recovery

제한한 수평 및 수직 복원 알고리즘을 적용하여 손실된 획을 복원하였으며, 이에 따라 영역 추출 성능을 향상시킬 수 있었다. <그림 6>은 주어진 이미지에 대한 복원 결과로서 <그림 6(b)>는 분산 결과에 AND 및 모폴로지 연산이 적용된 결과로서 수평 및 수직 복원이 적용되기 전의 결과이며, <그림 6(c)>는 복원 과정이 수행된 결과 이미지이다. <그림 6(d)>는 복원된 결과를 바탕으로 최종 추출된 텍스트 영역을 보여주며, 복원이 수행되기 전에는 조각으로 나타난 영역들이 결합되어 찾아진 결과를 확인할 수 있다.



(b) 복원 전

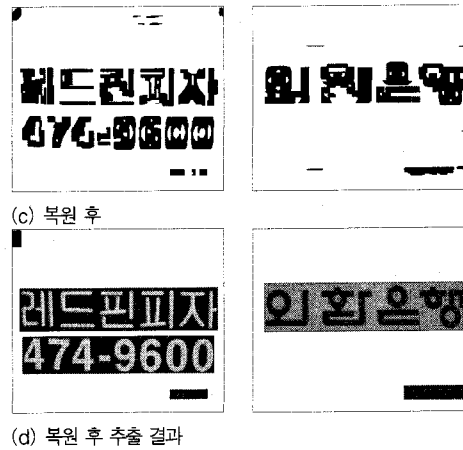


그림 6. 복원 결과
Fig. 6. Recovered Results

3.2 영역 제거 및 결합을 통한 확장

추출된 후보 영역들에 대한 최종 확정은 이전에 제안한 방법과 마찬가지로 두 단계의 영역 제거 및 영역 결합으로 이루어진다[13]. 첫 번째 단계로 연결요소 및 외곽사각형의 크기, 외곽사각형의 중횡비를 고려해서 제거하며, 외곽사각형의 상하 겹침 정도와 이와 비례한 상대적인 근접성을 고려해서 영역을 결합한다. 다음 단계에서는 외곽사각형의 크기에 따른 상대적인 겹침 정도와 근접성을 고려하여 제거 및 결합을 수행한다.

본 논문에서는 다양한 실험을 수행하여 영역 제거 파라미터 값들을 일부 조정하였으며, 특히 영역을 결합하는 과정에서 외곽사각형 사이의 절대적인 거리를 사용하던 이전의 방법에서 외곽사각형의 투영의 겹침 정도에 따른 상대적인 거리를 새로운 파라미터 값으로 적용하여 개선하였다.

영역 제거 1단계에서는 외곽사각형의 높이 또는 폭이 아주 작거나 외곽사각형의 높이가 큰 경우(70픽셀 이상) 제거하고, 외곽사각형의 가로 대 세로 비율과 외곽사각형에서 연결요소가 차지하는 면적을 고려하여, 외곽사각형의 면적이 작고(200 픽셀 이하), 가로 대 세로 비율이 낮으면 제거한다(0.2 이하). 외곽사각형의 크기가 일정한 수치(2,000 픽셀) 이하로서 문자일 가능성이 높지만, 연결요소가 차지하는 면적이 외곽사각형 면적의 40% 이하인 경우도 제거한다. 2단계 제거에서는 1차 결합이 수행된 후에 외곽사각형의 크기가 일정한 값 이하로 작거나(200 픽셀이하), 외곽사각형내의 연결요소의 면적 비율이 낮은 경우(0.3 이하) 제거한다. 또한, 1차 결합된 외곽사각형의 높이와 폭의 비율이 아주 큰(5 이상) 경우에도 함께 제거한다. <그림 7>은 이러한 제거 조건들이 적용된 결과를 보여준다.

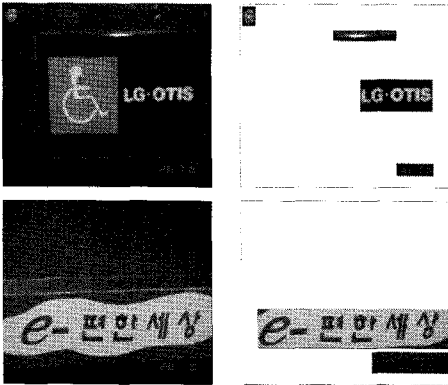


그림 7. 영역 제거 및 결합에 의한 영역 확정
Fig. 7. Detected Regions by Region Removal and Combining

영역 결합은 외곽사각형의 포함관계와 외곽사각형 투영의 수직 겹침 정도에 따른 상대적인 수평 거리를 고려하여 결합한다. 먼저 큰 외곽사각형에 작은 외곽사각형이 대부분 또는 완전히 포함된 경우에는 이전과 같이 큰 외곽사각형으로 통합하여 결합한다. 다음으로 두 외곽사각형의 X/Y축 방향으로 투영의 겹침 정도를 측정하여 겹쳐진 영역이 일정한 값 이상일 경우 두 외곽사각형을 결합하는데, 이 때 두 텍스트 후보 영역의 높이에 대한 겹침 정도에 따라 폭의 상대적인 거리를 설정하여 결합한다. 이 과정에서 두 후보 영역 가운데 외곽사각형의 높이가 큰 것을 기준으로 하여 두 영역의 Y축으로의 겹침 정도를 표현하고, 두 영역의 수평거리를 최소거리로 설정한다. 외곽사각형 높이의 겹침 정도가 크면 최소거리가 커도 결합하며, 높이의 겹침 정도는 작지만 최소거리가 작으면 마찬가지로 결합한다. 여기서는 <그림 8>과 같이 겹침 비율이 85% 이상이면 최소거리는 겹침 거리의 80% 이내에서 결합하고, 겹침 비율이 50% 이상이면 최소거리가 겹침 비율의 50% 이하일 때도 결합하였다. 또한, 겹침 비율이 30% 이하이지만, 최소거리가 20 픽셀 이하로 아주 가까운 경우에도 결합한다.

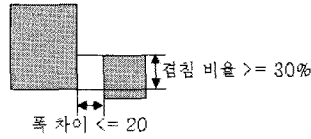
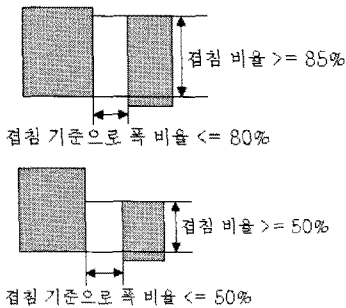


그림 8. 겹침 비율에 따른 최소거리의 상대적 적용
Fig. 8. Relative Applying of Minimum Distance According to Overlapped Ratio

<그림 9>는 위에서 제안한 제거 및 결합 기준을 적용한 결과로서, '아'는 '오'와 '나'의 겹침 비율이 50% 이하이지만 수평거리가 짧아서 결합되었다. 대부분의 텍스트가 수평방향으로 이루어진다는 전제하에 수평방향의 결합은 비교적 쉽게 허용하지만, 수직방향의 결합은 어렵게 설정하였다. 결합한 외곽사각형은 주변의 외곽사각형에 대하여 동일한 기준을 적용하여 반복적인 결합을 수행한다.

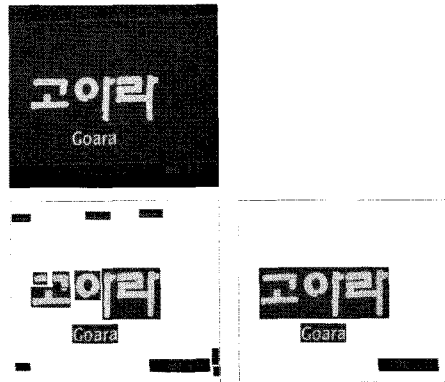


그림 9. 개선 전 및 개선 후 결과 비교
Fig. 9. Result Comparisons with Improved Approach

IV. 실험 및 결과

본 논문에서 제안한 방법은 Visual C++ 6.0을 이용하여 Pentium IV 2.8GHz 하드웨어에서 구현하였다. 디지털 카메라 및 휴대폰 카메라를 이용해서 간판, 표지판, 광고, 발표자료, 명함 등에서 직접 취득한 다양한 종류의 자연이미지와 ICDAR 콘테스트에서[1,2] 제공한 이미지를 함께 사용하였다.

텍스트 영역의 Ground-Truth 기준으로서 수평 방향으로 글자 간격이 크거나, 글자가 큰 텍스트 영역, 수직 방향의 텍스트 영역은 각 글자를 하나의 텍스트 영역으로 설정하였으며, 눈으로 식별할 수 있는 정도의 크기만을 추출 대상으로 설정하였다. 이전의 방법들과 마찬가지로 정확률(Precision)과 재현율(Recall)을 측정하였으며, 정확률은 텍스트로 검출한 영역 중에서 텍스트의 비율이며, 재현율은 전체 텍스트로

검출한 것 영역 중에서 텍스트 비율을 의미한다.

〈표 1〉은 비교적 단순한 동화책, 명함 등의 문서 유형 이미지를 대상으로 측정된 결과로서 개선 전은 분산 복원 및 영역 확정 과정이 개선되기 전을 의미한다. 제안한 방법을 적용한 결과를 이전 방법과 비교하면 재현율의 차이는 거의 없지만, 틀리게 찾은 개수를 많이 줄일 수 있어서 정확률이 향상된 것을 볼 수 있다. 이는 동화책, 명함 등의 비교적 단순한 이미지에 대해서는 개선 전에도 틀리게 찾은 영역은 많아도 못 찾은 영역은 거의 없어서 재현율에는 차이가 없지만, 틀리게 찾은 영역의 개수를 크게 줄여서 정확률에는 큰 차이가 있음을 보여주었다.

〈표 2〉는 직접 취득한 자연이미지 Set 1과 ICDAR 콘테스트에서 제공한 이미지 Set 2를 실험한 결과이다. 분산 복원 및 개선된 영역 확정 과정을 통해서 Set 1과 Set 2 모두에서 정확률과 재현율이 크게 향상된 결과를 보여준다. 이는 복잡한 이미지에 대해서도 제안한 방법의 유용성을 확인한 결과로서 추출한 결과의 예는 〈그림 10〉과 같다. (개인정보 보호를 위해서 명함이미지의 일부는 가린다.)

표 1. 문서유형 이미지에서의 영역 추출 결과
Table 1 Extraction Results with Document-style Images

이미지		A	B	C	D	정확률	재현율
동화책 (50개)	개선 전	881	868	13	235	79%	99%
	개선 후		880	1	49	94%	99%
명함 (50개)	개선 전	481	480	1	54	89%	99%
	개선 후		480	1	14	97%	99%

(A: 전체 텍스트 개수, B: 정확히 찾은 개수, C: 찾지 못한 개수, D: 틀리게 찾은 개수)

표 2. 자연이미지에서의 영역 추출 결과
Table 2 Extraction Results with Natural Scene Images

이미지		A	B	C	D	정확률	재현율
Set 1 (100개)	개선 전	418	358	60	513	41%	86%
	개선 후		408	10	197	67%	97%
Set 2 (111개)	개선 전	472	423	49	542	44%	90%
	개선 후		459	13	232	66%	97%

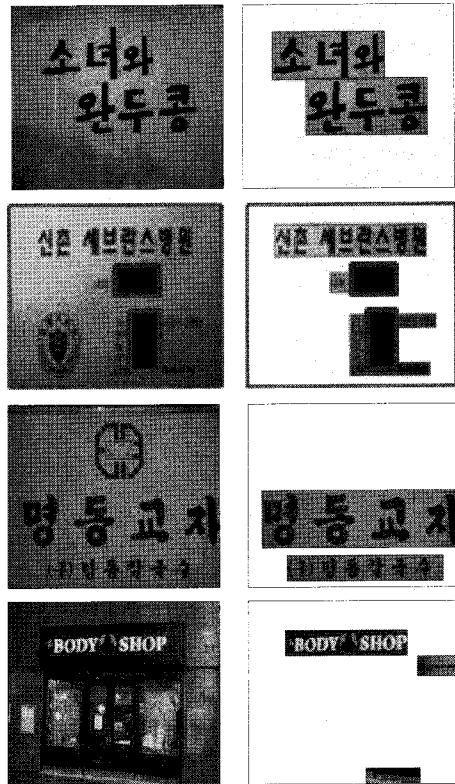


그림 10. 추출 결과 예
Fig. 10. Examples of Detection Results

V. 결론

본 논문에서는 기존에 제안한 색 분산을 이용한 이미지에서의 텍스트 추출 과정에서 발생할 수 있는 분산 미추출 문제를 보완하는 방법을 제안하였다. 이는 외곽사각형의 기하학적인 정보와 경험적인 규칙을 함께 이용하여 손실된 획 정보를 복원하는 것이다. 또한 후보 영역들을 제거하고 결합하는 과정에서 히스토그램 겹침 정도에 따른 상대적인 근접 거리를 이용하여 영역 결합 결과를 향상시켰다. 다양한 종류의 이미지에 제안한 방법을 적용하여 정확률 및 재현율이 모두 향상된 것을 확인할 수 있었다.

향후 연구로는 텍스트 영역을 찾기 위한 새로운 특징을 개발하는 것과 추출된 영역의 텍스트 검증 과정에서 OCR을 활용해서 인식 전단계의 인식 거부 정확성을 향상시키는 것이다.

참고문헌

- [1] S. M. Lucas, A. Panaretos, L. Sosa, A. Tang, S. Wong and R. Young, "ICDAR 2003 Robust Reading Competition," International Conference on Document Analysis and Recognition(ICDAR), Vol. 2, pp. 682-687, August 2003.
- [2] S. M. Lucas, "ICDAR 2005 Text Locating Competition Results," Proceeding of International Conference on Document Analysis and Recognition, Vol. 1, pp. 80-84, August 2005.
- [3] J. Ohya, A. Shio, S. Akamatsu, "Recognizing Characters in Scene images," IEEE Transactions Pattern Analysis and Machine Intelligence, PAMI-16(2), pp. 67-82, February 1994.
- [4] Anil K. Jain, Bin Yu, "Automatic Text Location in Images and Video Frames," Pattern Recognition, Vol. 31, No. 12, pp. 2055-2076, December 1998.
- [5] N. Ezaki, M. Bulacu and L. Schomaker, "Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons," Proceedings of 17th International Conference on Pattern Recognition, Vol. II, pp. 683-686, August 2004.
- [6] Premnath Dubey, "Edge Based Text Detection for Multi-purpose Application," The 8th International Conference on Signal Processing, Volume 4, pp. 16-20, June 2006.
- [7] 박종천, 황동국, 이우람, 전병민, "문자-에지 맵의 패턴 히스토그램을 이용한 자연이미지에서 텍스트 영역 추출", 한국산학기술학회논문지 제 7권 제 6호, 1167-1174쪽, 2006년. 12월.
- [8] Zongyi Liu, Sudeep Sarkar, "Robust outdoor text detection using text intensity and shape features," 19th International Conference on Pattern Recognition, pp. 1-4, December 2008.
- [9] Y. Zhong, K. Karu and A. K. Jain, "Locating Text in Complex Color Images," Pattern Recognition, Vol. 28, No. 10, pp. 1523-1535, October 1995.
- [10] Hao Wang, "Automatic Character Location and Segmentation in Color Scene Images," Proceedings of 11th International Conference on Image Analysis and Processing, pp. 2-7, September 2001.
- [11] 배경숙, 장대근, 정연구, 소정, "색기반 이진화를 이용한 장면 텍스트 추출과 씨포트 벡터머신을 이용한 텍스트 영역 검증," 영상처리 및 이해에 관한 워크숍 발표 논문집, 368-373쪽, 2005년. 1월.
- [12] 김지수, 김수형, 한승완, 남택용, 손화정, 오성열, "색상 레이어를 이용한 스펀메일 영상에서의 텍스트 영역 추출", 한국정보처리학회논문지B, 제 13권, 제 4호, 409-416쪽, 2006년. 8월.
- [13] 송영자, 최영우, "카메라 획득 영상에서의 색 분산 및 개선된 K-means 색 병합을 이용한 텍스트 영역 추출 및 이진화," 한국정보처리학회논문지B, 제 13-B권 제 3호, 205-214쪽, 2006년. 6월
- [14] Toan N. Dinh, Jonghyun Park, Guesang Lee, "Korean Text Detection and Binarization in Color Signboards," IEEE International. Conference on Advanced Language Processing and Web Information Technology, pp. 235-240, July 2008.
- [15] Rafael C. Gonzalez and Richard E. Woods, Digital Image Processing, Addison Wesley, 1993.
- [16] D. H. Ballard and C. M. Brown, "Computer Vision," Prentice-Hall, 1982.

저자소개



최 영 우

1985년 : 연세대학교 전자공학과 졸업
(학사)

1986년 : University of Southern
California 컴퓨터공학과 졸업
(석사)

1994년 : University of Southern
California 컴퓨터공학과 졸업
(박사)

1994년 : 10월 - 1997년 2월 LG전자
기술원 선임연구원

1997년 3월 - 현 재 :
숙명여자대학교 컴퓨터학과 교수

2006년 9월 - 현 재 :
숙명여자대학교 정보통신처장, 도서관장
관심분야 : 영상처리, 패턴인식, 정보화시
스템 구축 등



조 은 숙

2005년 : 경희사이버대학교
정보통신학과 졸업(학사)

2007년 : 숙명여자대학교
컴퓨터학과 졸업(석사)

관심분야 : 영상처리, 패턴인식 등