

## 모바일 카메라용 컬러영상 인식을 위한 문서영상 해석 알고리즘

김진호<sup>†</sup>

경일대학교 전자정보통신공학부  
경북 경산시 하양읍 부호리 33번지  
(2004년 5월 21일 받음, 2004년 6월 12일 채택)

## Document image analysis for mobile camera based color image recognition

Jin Ho Kim<sup>†</sup>

*Department of Electronic & Telecommunication, Kyungil University  
33 Buho-ri, Hayang-up, Gyeongsan-si, Gyeongsang buk-do, Korea  
(Received 21 May 2004, accepted for publication 12 June 2004)*

**Key words** selective local thresholding method, edge operator

**요 약** 본 논문에서는 모바일 카메라기반 컬러문서영상에서 선택적 국소영역 임계치 설정 방식의 이진화 기법을 이용한 문서영상 해석 알고리즘을 제안하였다. 일반적으로 모바일 카메라기반 컬러문서영상은 다양한 형태의 조명 불균형으로 인해 이진화가 매우 어려운 실정이다. 따라서 모바일 카메라기반 컬러문서영상으로부터 문자영역을 추출하고 이를 인식하기 위해 많은 연구들이 진행되고 있다. 선택적 국소영역 임계치 설정방식의 이진화 기법은 국소영역을 대상으로 최적의 임계치를 구하므로 조명의 불균형 현상을 보상할 수 있는 방법으로 활용될 수 있다. 국소영역 내의 화소들의 명도값 분포 정보를 이용한 이진화 임계치 설정 방식을 제안하고 에지연산자를 이용한 문자영역과 개별 문자를 추출하는 문서영상 해석 알고리즘을 구현하였다. 제안된 방식을 이용하여 한국전자통신연구원의 모바일 카메라기반 문서영상 데이터베이스에 적용 실험해 본 결과 비교적 우수한 문자추출 성능을 얻을 수 있었다. 또한 처리속도가 빠르고 요구되는 상용 모바일 기기에 탑재할 수 있음을 확인하였다.

**Abstract** This paper presents a document image analysis algorithm by selective local thresholding method for mobile camera based color document image recognition. It is very difficult to binarize and segment a mobile camera based color document image due to inconsistent lighting condition. A selective local thresholding method in character region can reduce illumination effect and extract characters properly. Edge operator is used to detect character string and sub-window size for applying local thresholding method is decided from the height of character string. And then local thresholding method is applied in the selected character region. Distribution of intensity value of pixels is used to binarize the character region. It is an efficient method to binarize camera image and to extract characters from background image. The proposed method can be implemented in mobile device because of high speed processing time and a considerable memory. To evaluate the proposed method, we have experimented with camera document image of the ETRI database. An encouraging results have been obtained.

<sup>†</sup> e-mail : ho@kiu.ac.kr

## 1. 서론

PDA나 셀룰러폰과 같은 모바일 휴대장치에 소형 카메라가 탑재되면서부터 카메라기반 컬러 문서영상의 인식에 대한 연구가 매우 활발히 진행되고 있다[1-4]. 모바일 카메라기반 컬러 문서영상은 저해상도로서 변화와 왜곡이 심하고 조명의 불균형 및 잡음 등으로 인해 관심영역을 추출하고 인식하기가 매우 어려운 것으로 알려져 있다. 특히 일관성이 없는 조명 상황에서 영상을 획득하게 되므로 관심영역 추출 이전에 거쳐야 할 이진화가 매우 어렵고 또한 잡음이 많이 포함된 이진화 결과영상으로부터 인식대상 개별 문자들을 추출해 내는 것도 매우 어려운 과정으로 알려져 있다.

최근 국소영역 임계치 설정기법에 의한 컬러영상 이진화 연구 결과들이 많이 발표되었다[3-4]. 제안된 이진화 방법에서는 이웃하는 화소들의 명도 레벨 분포에 따라 문서영상을 이진화하였다. 모든 화소들에 대해 배경 명도 값들을 계산하고 저해상도 문자 영역을 추정하여 주변 배경 명도값들을 상호간섭시킴으로써 이진화를 수행하였다. 주변 화소들의 배경 명도값만으로 저해상도 문자영역을 찾아내는 것은 매우 어려우며 Niblack의 국소 이진화 방법에 비하여 이진화 수행 시간은 줄일 수 있었지만 여전히 실시간 처리가 어려운 문제점을 안고 있다.

이러한 이진화 방법들은 조명이 불규칙하고 잡음이 많은 컬러영상의 이진화에서 전역 임계치 설정 방법보다는 비교적 우수한 결과를 얻을 수 있지만 문자의 작은 획 정보가 중요한 문서영상의 이진화에서는 배경영상을 완전히 제거하는데 여전히 문제점을 가지고 있다. 모바일 카메라기반 컬러 문서 영상에서는 불균일한 조도에 따라 문자 획을 이루고 있는 화소들의 명도값에 작은 차이가 발생해도 문자 추출에 상당한 어려움이 따르게 된다. 즉, 어떤 문자들은 배경에 포함되어 사라지거나 어떤 문자 획은 배경영상을 포함해서 나타나는 경우가 흔하게 발생하게 된다. 또한 국소영역의 크기를 결정하는 윈도우 크기에 따라 문자 자체가 배경

영상으로 간주되거나 배경영상이 문자영역으로 간주되는 경우가 나타나기 때문에 최적의 윈도우 크기를 결정하는 데도 많은 어려움이 따르게 된다.

본 논문에서는 모바일 카메라기반 컬러 문서영상의 문자영역에 대해 선택적으로 국소영역을 설정하고 임계치를 구하는 방식에 의한 이진화를 수행하고 개별문자를 추출하는 문서영상 해석 알고리즘을 제안하였다. 컬러 문서영상에서 배경영상으로부터 문자영역을 분리해 내기 위해 에지연산자가 이용되었고 문자열의 높이 정보를 예측하여 윈도우의 크기를 설정하는 알고리즘을 개발하였다. 국소영역 이진화 임계치는 영역 내에 포함된 화소들의 명도값 분포를 기반으로 최적에 가까운 설정될 수 있도록 하였다. 제안된 방식을 이용하여 한국전자통신연구원의 모바일 카메라기반 문서영상 데이터베이스에 적용하여 실험해 보고 그 결과를 분석하였다. 2단원 및 3단원에서는 문서영상 해석을 통한 문자영역 추출 및 개별 문자 추출에 대하여 나타내었으며, 4단원 및 5단원에서는 제안한 방법에 대한 실험 결과 및 결론을 다루었다.

## 2. 문서영상 해석을 통한 문자영역 추출

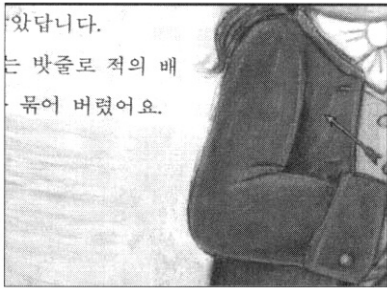
모바일 카메라로부터 입력된 컬러 문서영상은 트루컬러로서 256레벨 컬러영상으로 양자화(quantization)하고 256레벨 컬러영상은 다음과 같이 256레벨 명도영상으로 변환한다.

$$G(B, G, R) = \alpha \times B + \beta \times G + \gamma \times R \quad (1)$$

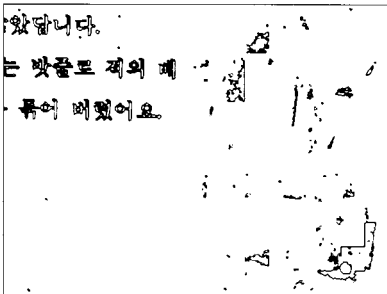
컬러화소값 R, G 그리고 B 각각에 적용되는 가중치는 인간의 시각 시스템에 비례하는 값으로 설정된다.

변환된 명도영상에 대해 소벨 에지연산자를 적용하여 영상내의 에지성분을 추출한다. Figure 1은 컬러 원본 영상과 이 영상에서 문자영역을 찾아내기 위하여 에지연산자를 적용한 결과를 도시한 것이다. 에지 연산자를 적용한 결

과 문자 획의 가장 자리에 강한 에지성분이 나타나는 것을 확인할 수 있다.



(a)



(b)

Figure 1. Experimental images, (a) an original color image and (b) an edge image of the grey level image.

에지성분을 추출한 영상에서 인접한 에지성분들에 대하여 문자열을 구성하기 위한 목적으로 스미어링(smearing) 방법을 적용하여 결합시키는 과정을 수행한다. 인접한 에지성분이 서로 결합되어 라인 형태의 문자열을 형성하게 됨으로 이 영역을 문자열에 해당하는 부분으로 간주하고 문자열 추출을 수행한다. 개별 문자들은 결합하여 문자열을 구성하기 위해서는 문자들이 비슷한 높이 정보를 가지고 있고 그 크기 또한 비슷하다는 정보를 이용하여 에지 문서영상 전체를 해석하는 과정을 수행한다. Figure 2는 인접한 에지성분들을 결합한 다음 에지문서영상 해석을 통해 문자열을 추출한 예를 나타낸 것이다.

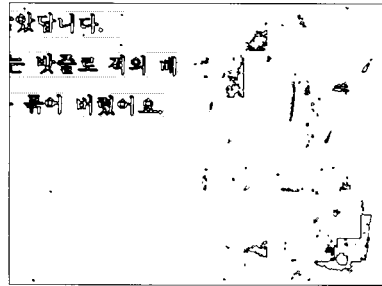


Figure 2. Selected character regions by merging of neighboring edges.

국소영역 이진화 임계치 설정 알고리즘은 선택된 문자열 후보 영역에 대해서 이루어진다. 국소영역 이진화를 수행할 부윈도의 크기는 에지 영상 내에서 문자열 후보들을 추출해서 평균한 값을 바탕으로 결정된다. 컬러문서영상의 이진화 임계치는 국소영역을 포함하는 부윈도우 내의 화소 명도에 대해 최대 및 최소값의 차이정보를 바탕으로 설정된다. 모바일 기반 컬러문서영상의 이진화 임계치는 다음과 같이 부윈도우내의 화소값들을 이용하여 설정한다.

$$T = \alpha \times (g_{\max} - g_{\min}) \quad 0.4 \leq \alpha \leq 0.7 \quad (2)$$

여기서  $g_{\max}$  및  $g_{\min}$  은 각각  $r \times r$  부윈도우 내의 화소들 중에서 최대 및 최소 명도값을 의미하고 가중치  $\alpha$ 는 영상 획득 조건에 따라 실험적으로 결정한다. 그림 1에 도시된 원본 영상에 대해 주어진 방법에 따라 이진화 한 결과를 Figure 3에 도시하였다.

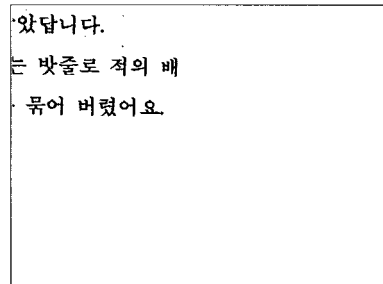


Figure 3. A binary image by local selective thresholding method.

Figure 3에 나타난 선택적 국소영역 이진화 결과에는 Figure 2에서 보였던 문장영역 이외의 많은 잡음들이 대부분 제거된 것을 확인 할 수 있다.

### 3. 개별문자 추출 알고리즘

문자열로 추정되는 화소들로 구성된 이진화 영상에 대해 가로방향으로 화소값들을 투영하고 그 결과를 해석하여 문자열을 추출한다. 추출된 문자열에 대해 다시 세로방향으로 화소값들을 투영하고 문자와 문자사이 정보 및 단어와 단어 사이 정보를 추출하여 최종적으로 단어 영역을 찾아낸다. 각 단어영역의 화소들에 대해 연결화소 성분들(connected components)을 추출하고 각 연결화소 성분들을 해석하여 개별 문자들을 추출한다. 문자열 내에 포함될 수 있는 한글 이외의 영문 및 숫자들을 구분해 내기 위해서 그림 4와 같은 한글 문자 구성 규칙을 활용한다.

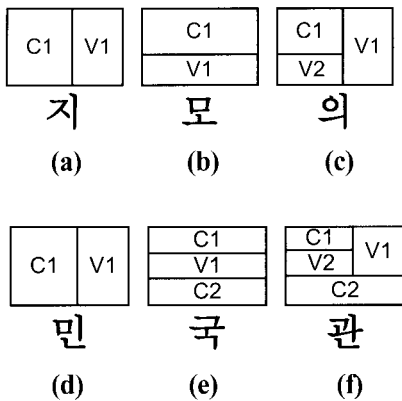


Figure 4. Six types of Korean characters, (a) type 1, (b) type 2, (c) type 3, (d) type 4, (e) type 5, and (f) type 6.

Figure 4에 도시된 규칙은 한글이 자음과 모음으로 구성되며 각 글자는 최소외접사각형(minimum boundary rectangle)내에 포함된다는 정보를 이용하여 설정되었다. 따라서 대상 연결화소들이 한글의 6가지 유형 중에 하나에 속하는지 여부를 판단하기 위해 단어영역 내의 각

연결화소들에 대해 결합 및 분리 규칙(merge and split rule)을 적용하여 개별 문자들을 추출한다. Figure 5는 Figure 3의 이진영상에 대해 개별 문자들을 추출한 결과를 도시한 것이다.

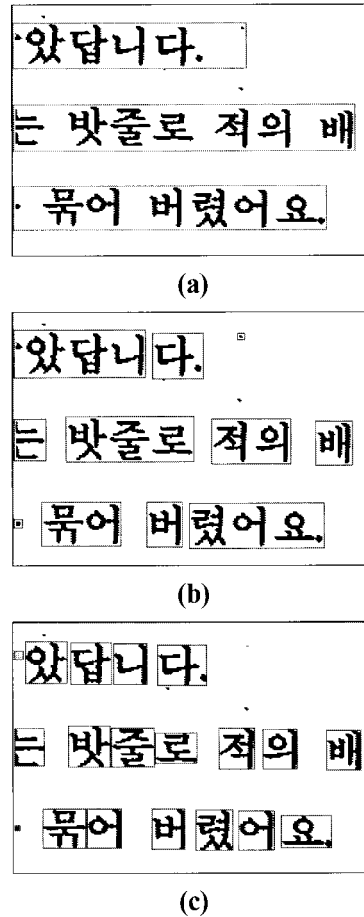


Figure 5. Character segmentation results (a) character string extraction, (b) word extraction, and (c) individual character segmentation.

### 4. 실험결과

제안된 선택적 국소영역기반 이진화 임계치 설정 방식의 문자추출 알고리즘을 한국전자통신 연구원의 모바일 카메라기반 컬러문서영상 데이터베이스에 대해 적용해보고 그 결과를 분석해

보았다. 제안된 방식의 성능 평가를 위하여 추출된 문자 영역에 대해서 국소 영역 이진화를 수행한 다음 문자 인식소프트웨어에 입력해서 인식 성능을 얻도록 하는 실험과 문서 영상 전체에 대하여 이진화를 수행한 다음 문자 인식소프트웨어에 입력하여 인식 성능을 얻어내는 실험을 통하여 인식 성능을 비교 분석해 보았다. 문서영상 인식을 위해 컬러 또는 명도영상으로 구성된 한글, 영문 및 숫자들을 인식할 수 있는 문자 인식소프트웨어를 이용하였다. 테스트용 모바일 카메라기반 문서영상은 320 x 240 크기로 구성된 25개의 영상 데이터를 사용하였다. 실험결과 제안된 방식을 적용하여 추출한 문자 영역에 대하여 국소 영역 이진화를 수행하여 문서 인식 소프트웨어에 입력한 결과 약 92.5%의 인식결과를 얻을 수 있었으며 이 실험 결과는 전체 문서 영상에 대해 이진화를 수행한 다음 문자 인식 소프트웨어에 그대로 입력했을 때 보다 약 3.5% 정도 인식성능에 있어서 증가한 결과를 얻을 수 있었다.

## 5. 결론

본 논문에서는 모바일 카메라기반 컬러문서영상에서 에지정보를 이용하여 문자영역을 추정하고 이를 바탕으로 선택적으로 국소영역을 설정하고 이진화 임계치를 구하는 방식에 의한 개별문자 추출용 문서영상 해석 알고리즘을 제안하였다. 컬러문서영상에서 배경영상으로부터 문자영역을 분리해 내기 위해 에지연산자가 이용되었고 문자열의 높이 정보를 예측하여 윈도우의 크기를 설정하는 알고리즘을 개발하였다. 국소영역 이진화 임계치는 영역 내에 포함된 화소들의 최대 및 최소값을 이용하여 설정하였다.

제안된 방식을 이용하여 모바일 카메라기반 컬러 문서영상 데이터베이스에 적용하여 실험해보고 그 결과를 분석하였다. 추출된 문자영상을 문서인식 소프트웨어에 입력시켜 인식해 본 결과 기존 방식에 비해 약 3.5% 정도의 인식을 향상을 얻을 수 있었다. 국소영역 이진화 및 문자

추출 알고리즘에 대해 처리속도를 빠르게 하고 소요되는 메모리 크기를 최대한 줄여서 모바일 장치에 탑재될 수 있도록 재구성하는 연구를 계속 수행할 예정이다.

## 참고문헌

1. H. Fujisawa, H. Sako, Y. Okada, and S. W. Lee, "Information capturing camera and developmental issues," Proc.of the 6th ICDAR. pp. 205-208, 2001.
2. M. Sawaguchi, K. Yamamoto, and K. Kato, "A proposal of character recognition method for low resolution images by using cellular phone," Proc.of the 9th Korean-Japan joint workshop FCV , pp. 216-221, 2003.
3. M. Seeger and C. Dance, "Binarizing camera images for OCR," Proc. of the 6th ICDAR, pp. 54-58, 2001.
4. L. Fan, L. Fan and C.L.Tan, "Binarizing document image using coplanar prefilter," Proc. of the 6th ICDAR, pp. 34-38, 2001.



김진호

1986년 경북대학교 전자공학과  
공학사

1988년 경북대학교 대학원  
전자공학과 공학석사

1992년 경북대학교 대학원  
전자공학과 공학박사

1992년~현재 경일대학교  
전자정보통신공학부 교수

Tel: 053-850-7167

Fax: 052-850-7779

E-mail: ho@kiu.ac.kr