

## 카메라기반 문서인식 시스템을 위한 현장문서에 적합한 이진화 알고리즘 특징성능의 비교

지수영, 김계경, 유원필, 정연구, \*\*김태운  
한국전자통신연구원, \*\*고려대학교  
전화 : 042-860-5337 / 핸드폰 : 016-467-8747

### Comparison of Feature Performance of Binarization Methods for Character Recognition System Based on Digital Camera

Soo-Young Chi, Kye-Kyung Kim, Won-Pil Yu, Yun-Koo Chung, \*\*Tae-Yuen Kim  
Electronics and Telecommunications Research Institute  
\*\*Dept. of Computer Science, Korea University  
E-mail : chisy@etri.re.kr

#### Abstract

This paper represents a survey of a variety of thresholding techniques including both global and local thresholding. Several thresholding methods are examined in detail to evaluate their performance based on a given set of test images. We also attempt to evaluate the performance of several thresholding methods for construction field documents image recognition system using a broken line structures, broken symbols and text, blurring of lines, symbols and text, noise in homogeneous areas measure as a criterion functions.

#### I. 서론

지능정보단말기의 발전에 따라 시간과 공간에 제약 없이 받지 않고 문서정보의 입력, 저장, 편집 및 전송을 효율적으로 하기 위하여 카메라 기반 문서 인식 연구가 필요하다. 카메라기술과 PDA기술이 결합되면서 문서인식에 대한 수요는 매우 커져가고 있다. 카메라를 이용하여 스캐너수준의 고화질 문서영상을 입력받는데는 조명의 문제나 블러링이나 해상도등 많은 제약이 따른다. 특히 현장 문서영상이 갖는 다양하고 복잡한 특징으로 인하여 더욱 정교한 인식기술개발을 필요로

한다. 여기에서 현상이 갖는 특징으로 인하여 조명이 균일하지 않고, 글자와 배경이 복잡한 현장문서영상에서 글자를 추출하고 인식하는 연구가 필요하다. 따라서 카메라기반 현장 문서영상 인식을 위하여 문서영상 향상/개선 기술, 이진화 기술, 잡영제거 기술, 영상 보강 기술 등의 전처리등에 관한 연구[1,2,3,4]들이 진행되어야 한다.

지능정보단말기의 사용 증대 및 이에 따른 컴퓨터 환경의 급격한 변화에 따른 카메라 기반 문서 인식 기술 개발은 다양한 현장 문서에 대한 정보 획득의 편리성, 정보 공유의 신속성에 대한 요구를 충족시켜주며 또한, 상용문서 인식 응용 시스템의 요소 모듈로도 활용할 수 있다. 이와 더불어 현장에서 직접 문서 영상을 텍스트로 변환하여 활용할 수 있는 현장 적용형 문서 인식 기술의 개발은 국내·외적으로 문서 인식 상용 시스템 시장을 활성화시키는 데 크게 기여할 수 있다. 카메라 기반 문서 영상의 분석 및 문자 인식플랫폼을 구현하기 위해서는 다양한 현장의 카메라로부터 입력되는 문서영상의 이진화 방법에 대하여 현장 적용에 적합한 새로운 연구가 필요하다.

본 논문은 카메라기반 문서인식 시스템을 위한 현장문서에 적합한 이진화 알고리즘의 연구를 수행하는 것으로서, 다양한 현장에 존재하는 현장문서를 분석하

고 이에 적합한 이진화 알고리즘들에 대한 성능평가를 함으로 카메라기반 문서인식 시스템 구현을 위한 토대를 마련하고자 한다. 다양한 현장을 갖는 문서영상 인식처리 시스템에서 그들의 현장 배경으로부터 정확한 문서 영상을 추출하기 위하여 적합한 이진화 방법을 찾는 것은 매우 중요하다. 이 분야에서 다양한 방법들이 제안되어 왔다. 그러나 다양한 현장 문서가 갖는 특징들로 인해 정확하고 정밀하게 배경과 문서영상을 분리하는 것은 그리 쉽지가 않다. 또한 스캐너 기반의 기존의 인쇄체 문서영상 인식에 사용된 이진화 방법들이 디지털 카메라 기반의 이진화에 적용할 경우 여러 가지 복잡하고 어려운 문제들로 인하여 쉽게 적용하기가 곤란한 부분이 있는 것 또한 사실이다. 따라서 다양한 현장 문서가 갖는 특징들을 분석하고 이에 적합한 이진화 방법에 대한 연구는 필수적이다.

본 논문은 기존의 이진화 방법들의 문서영상분야에서의 성능평가 결과를 기준으로 분석하여 결과를 정리하였고 이를 토대로 현장 문서에 적합한 이진화 방법들을 제시함이 목적이다. 실험에서 사용된 현장 문서 영상들은 다양한 현장에서 디지털 카메라로 획득한 문서영상들을 사용하였다.

## II. 이진화 방법의 성능비교

현재까지 매우 다양한 이진화 방법이 제안되고 이용되었다. 이진화 기술에 대한 국내의 기술동향을 보면 최적의 영상 분할을 얻어 오분할에 의한 인식을 저하를 최소화하기 위한 다중 이진화 기법[8]과 가우시안, 미디어 필터 사용기술[10,11], Coplanar 필터를 이용한 문서영상의 이진화 기술[9], 국부(global) 또는 국소(local) 임계치 적용기술이[12,13,14,15,16] 있다. 특히 문서영상에서 글자와 배경을 분리하기 위하여 사용되는 이진화 방법의 성능비교와 평가는 [5,6,7]에서 찾을 수 있다.

본 논문에서는 기존의 이진화 방법들과 그 성능평가 결과를 정리하고 이로부터 현장 문서영상에 적합한 이진화 방법들을 제시한다.

### 2.1 전역적 이진화 방법

Wong[12]의 논문은 이진화 방법론을 체계적으로 분류하였으며 9개의 전역적 이진화 방법들의 3개의 일반 영상에 대한 결과를 비교, 분석하였다. 성능비교에 사용한 척도는 uniformity와 shape measure이다. 이 논문의 결과를 볼 때 모든 영상에 대하여 좋은 결과를 보이는 것은 Johannsen and Bille, Kapur et al.,

moment-preserving, Ostu method이다.

### 2.2 국소적 이진화 방법

Trier의 연구[17,18]에서는 이진화 방법의 성능을 평가하기 위하여 문서영상을 사용하였다. 이 논문에서는 11개의 국소적 이진화 방법과 4개의 전역적 이진화 방법이 비교대상으로 선택되었다. 평가 기준으로는 직선이 끊어진 정도, 끊어진 심볼이나 텍스트가 있는지, blurring의 정도, 없어진 물체가 있는지, 잡영이 얼마나 남아있는지 등을 이용하였다. 평가는 사람들에 의하여 주관적으로 이루어졌으며, 다른 사람의 평가 결과를 검증하는 단계를 거쳤다. 이 결과에서 전역적 이진화에 비하여 국소적 이진화 방법이 좋은 결과를 보인다는 결론을 내릴 수 있다.

## III. 실험 방법 및 결과

### 3.1 실험 이미지

아래의 실험 이미지들은 SONY 디지털카메라로 영상을 획득하여 256 gray-scale image로 저장한 현장 실험영상들이다.



그림3.1 명함영상      그림3.2 파노라마영상      그림3.3 컨퍼런스영상

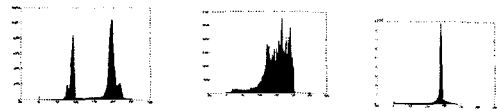


그림3.4 3.1히스토그램      그림 3.5 3.2히스토그램      그림3.6 3.3 히스토그램

### 3.2 성능평가 기준

본 논문에서는 현장 문서영상에 적합한 이진화 방법의 성능을 평가하기 위하여 3개의 전역적 이진화 방법과 5개의 지역적 이진화 방법이 비교대상으로 선택되었다. 성능평가 기준으로는 끊어진 문자나 기호가 있는지, blurring의 정도, 없어진 물체가 있는지, 잡영이 얼마나 남아있는지 등을 이용하였다. 평가는 사람들에 의하여 주관적으로 이루어졌으며, 다른 사람의 평가결과를 검증하는 단계를 거쳤다. 그 결과는 표 1

과 같다.

### 3.3 실험 결과

현장 문서영상에 적합한 전역적 이진화 방법은 실험의 결과를 볼 때 Ostu의 방법이 좋은 것으로 나타났고, 국소적 이진화 방법으로는 Eikvil방법이 좋은 것으로 나왔다. 특히 지역적 이진화 방법이 전역적 이진화 방법보다 이진화의 성능이 좋은 것으로 나타났다.

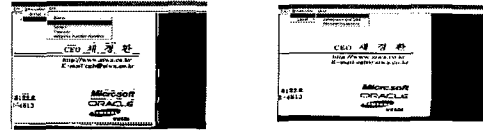


그림3.11 1의 Eikvil method      그림3.11 1의 ostu method

## IV. 결론

본 논문은 기존의 이진화 방법들의 문서영상분야에서의 성능평가 결과를 기준으로 분석하여 결과를 정리 하였고 이를 토대로 현장 문서에 적합한 이진화 방법 들을 제시하였다. 현장 문서영상의 특성상 다양한 조명, 낮은 질의 이미지, 디지털 카메라로 현장 문서 영 상을 획득하는 사람의 자세나 특성 그리고 디지털 카 메라의 성능등 다양한 현장의 상황들로 인하여 어떤 현장 문서영상에 대하여 하나의 이진화 방법이 절대적 으로 좋다고 할 수 없으며 그 상황에 따라 여러 가지 이진화 방법이 적용해야 한다는 어려움이 있다. 현장 문서영상에 적합한 전역적 이진화 방법은 실험의 결과 를 볼 때 Ostu의 방법이 좋은 것으로 나타났고, 국소 적 이진화 방법으로는 Eikvil방법이 좋은 것으로 나왔 다. 특히 지역적 이진화 방법이 전역적 이진화 방법보 다 이진화의 성능이 좋은 것으로 나타났다.

표1. 이진화 성능평가 순위

method	image number			execution time(sec)	rank
	1	2	3		
<b>global methods</b>					
ostu	20	3	22	0.2 0.2 0.1	2
J&B	5	3	20	0.3 0.2 0.1	7
moment preserving	3	20	2	0.2 0.1 0.1	8
<b>locally adaptive methods</b>					
Bersen	12	10	14	1.5 6 5	5
Eikvil	16	22	22	2 8 5	1
Niblack	10	14	8	3 15 9	6
Yanowitz	17	12	13	43 90 25	3
IFA	14	15	10	1.2 5 3	4

image num 1=명함 영상, 2=파노라마 영상  
3=컨퍼런스 발표영상

## V. ACKNOWLEDGEMENTS

이 논문은 1997년부터 1999년 3년 동안 정보통신부 의 출연금으로 수행한 지식기반 적응형 컴퓨팅 기술 개발의 연구결과로 나온 것이며 연구책임자인 오원근 박사님과 과제에 참여한 동료들께 깊은 감사를 드립니다. 그리고 이 논문에 설명된 알고리즘의 original version은 강경원 박사에 의해 조사되었습니다.

## Reference

- [1] T. Sato, T. Kanade, E. K. Hughes and M. A. Smith, "Video OCR for digital news archives," IEEE Workshop on Content-Based Access of Image Video Databases, pp. 52-60, India, 1998.
- [2] H. Li, O. Kia, and D. Doermann, "Text enhancement in digital video," Proceeding of SPIE Conf. On Document Recognition and Retrieval VI, Vol. 3651, pp. 2-9, 1999.

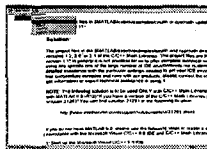


그림3.7 3의 Eikvil method

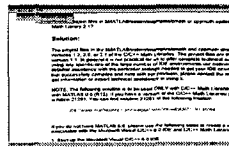


그림3.8 3의 ostu method

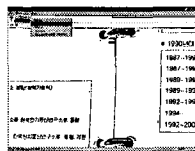


그림3.9 2의 Eikvil method

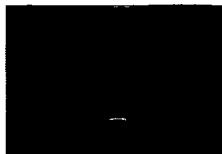


그림3.10 2의 ostu method

- [3] P. Natarajan, B. Elmieh, R. Schwartz and J. Makhoul, "Videotext OCR using Hidden Markov Models," Proceedings 6th ICDAR, pp. 947-951, 2001.
- [4] S. Madhvanath, G. Kim, and V. Govindaraju, "Chaincode contour processing for handwritten word recognition," IEEE Trans., on PAMI21, pp. 928-932, 1999.
- [5] M.Kanel and A. Thao, "Extraction of Binary character/graphics images from grayscale document images," CVGIP: Graphical Models and Image Processing, Vol.55, No.3, pp. 203-217, May 1993.
- [6] P.W. Palumbo, P. Swaminathan, and S. N. Srihari, "Document Image Binarization: Evaluation of Algorithms," SPIE Vol.697: Applications of Digital Image Processing IX, pp.278-285, 1986.
- [7] K. A. Moon, "Performance Comparison of Binarization Method for Document Image Processing," ETRI Technical Report, Sep. 1999.
- [8] 김동수, 진성일, 다중 이진화 문턱값 기반의 합성내 디지털 계기판 인식," 컴퓨터비전 및 패턴인식 연구회 추계워크샵, pp. 49-50, 2001.
- [9] Liying, Fan, and C. L. Tan, "Binarizing document image using coplanar prefilter," Proceedings 6th ICDAR, pp. 34-38, 2001.
- [10] C. Bouman and K. Sauer, "A generalized gaussian image model for edge-preserving map estimation," IEEE Trans. Image Processing, 2:pp. 296-310, July, 1993.
- [11] A.C.Bovik, T. Huang, and J. Munson, "A generalization of median filtering using linear combination of ordering statistics," IEEE Trans. Acoust. Speech, Signal Processing, 31:pp. 1342-1350, July, 1983.
- [12] P. Sahoo, S. Soltani, and Wong, "A survey of thresholding techniques," Computer Vision, Graphics, and Image Processing, 41:pp. 233-260, 1988.
- [13] A. Dawoud and M. Kamel, "Binarization of document image using image dependent model," Proceeding 6th ICDAR, pp. 49-53, 2001.
- [14] M. Seeger and C. Dance, "Binarizing camera images for OCR," Proceeding 6th ICDAR, pp. 54-58, 2001.
- [15] W. Niblack, "An Introduction to digital image processing," Prentice Hall, Englewood Cliffs, 1986.
- [16] Sang Uk Lee, Seok Yoon Chung and Rae Hong Park, "A Comparative Performance Study of Several Global Thresholding Techniques for Segmentation," Computer Vision, Graphics, and Image Processing 52. pp. 171-190, 1990.
- [17] Qivind Due Trier and Torfinn Taxt, "Evaluation of Binarization Methods for Document Images," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 17, No. 3, pp. 312-315, 1995.
- [18] Qivind Due Trier and Anil K. Jain, "Goal-Directed Evaluation of Binarization Methods," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 17, No. 12, pp. 1191-1201, 1995.