

MSER-b 이진화 기법을 이용한 스마트폰 문서 이미지 보정 기법

유영중¹ · 문상호² · 박성호^{3*}

Rectification of Document Image on Smartphone Using MSER-b Binarization

Young-jung Yu¹ · Sang-ho Moon² · Seong-ho Park^{3*}

¹Department of Computer Engineering, Busan University of Foreign Studies, Busan 609-815, Korea

²Department of Computer Engineering, Busan University of Foreign Studies, Busan 609-815, Korea

^{3*}Information Technology Center, Pusan National University, Busan 609-735, Korea

요 약

스마트폰 카메라로 생성한 문서 이미지는 촬영 방법에 따라 일반 스캐너에 비해 회전 왜곡과 원근 왜곡이 발생한다. 본 논문에서는 MSER-b 통해 조명에 영향을 적게 받는 이미지를 생성하고 텍스트 이미지의 특성을 고려한 텍스트 영역 윤곽선 검출 기법을 제안하고 이를 통해 왜곡된 문서 이미지를 보정하여 프린터 품질의 이미지로 복원하였다. 그리고 제안한 기법의 성능 평가를 위해 현재 서비스되고 있는 타사의 제품과 비교하였으며, 다양한 왜곡에 대하여 효과적으로 처리가 가능함을 실험을 통해 보였다.

ABSTRACT

The smartphone with camera can easily generate an image instead of a scanner. However the document image through a smartphone can have distortions related rotation or perspective. In this paper, we proposed a method to generate the document image in that distortions are reduced from the captured document image through a smartphone. For this, the original document image through a smartphone is preprocessed using the MSER-b technique to reduce the light effect. Then, the text area contour is extracted using the characteristics of the document image. Lastly, rotation or perspective distortions are reduced using the extracted text area contour. For experiments, the proposed method is compared two other products. Through experiments, we show that the distortions within the captured document image through smartphone can be effectively reduced.

키워드 : 스마트폰, 왜곡, 보정

Key word : smartphone, distortion, rectification

접수일자 : 2014. 11. 21 심사완료일자 : 2014. 12. 12 게재확정일자 : 2014. 12. 26

* **Corresponding Author** Seong-ho Park(E-mail:shpark@pusan.ac.kr, Tel:+82-51-510-2785)
Information Technology Center, Pusan National University, Busan 609-735, Korea

Open Access <http://dx.doi.org/10.6109/jkiice.2015.19.1.201>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

카메라를 탑재한 스마트폰이 널리 사용되고 있는 오늘날에는 종이 문서를 디지털 이미지로 변환하는데 접근성이 떨어지는 스캐너보다는 스마트폰의 카메라를 이용하는 사용자가 증가하고 있다. 카메라를 탑재한 스마트폰은 스캐너를 대체하여 편리하게 문서 이미지를 생성할 수 있다. 그러나 스마트폰 카메라로 생성한 문서 이미지는 그림 1과 같이 촬영 방법에 따라 일반 스캐너에 비해 회전 왜곡이 발생하거나 원근 왜곡이 발생한다. 이러한 이미지 왜곡은 촬영자의 노력에 따라 왜곡을 줄일 수 있지만 촬영자의 노력만으로 완전히 해소하기는 매우 어렵다. 또한 목표로 하는 문서 이외의 다양한 잡음 영상이 이미지에 포함되는 경우도 발생한다.

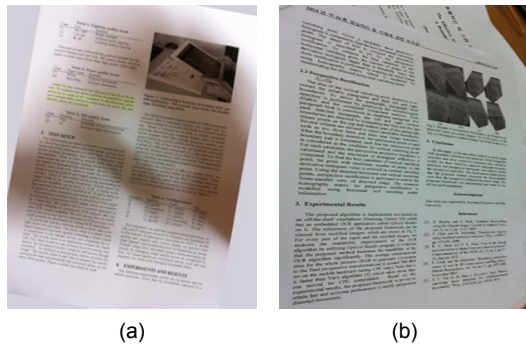


그림 1. 스마트폰 문서 이미지 : (a) 회전 왜곡 이미지, (b) 원근 왜곡

Fig. 1 Document Images on smartphone: (a) rotation distortion, (b) perspective distortion

카메라로부터 캡처된 이미지를 보정하는 연구는 광학 문자 인식(OCR)을 위한 전처리 과정으로 진행되어 왔다. 광학 문자 인식을 위해 카메라 캡처 문서를 입력받아 이미지의 왜곡을 보정하여 정규화 함으로써 문자의 인식률을 높이는 역할을 수행한다. 이러한 연구는 자동차 번호판, 바코더, QR코더 인식과 같은 특수한 분야에서 특정한 색상의 외곽선이나 위치 패턴을 이용하여 이미지를 보정하는 기법을 활용하고 있다 [1]. 그러나 그림 1과 같이 스마트폰 카메라로 입력된 일반문서 이미지는 특정한 색상의 외곽선이나 위치 패턴 등을 이용할 수 없으므로 직사각형 형태의 정규화된 이미지로 보정하기 위해서는 추가적인 연구가 필요

하다.

오늘날 스마트 폰이나 휴대용 카메라로부터 입력받은 문서 이미지의 직사각형 정규 이미지로의 자동 변환에 대한 요구가 증가하고 있으며, 관련 연구도 진행되고 있다. Jagannathan[2] 등은 카메라로부터 입력받은 왜곡된 문서 이미지를 투영변환(perspective transformation)을 통해 직사각형의 정규화된 문서 이미지로 변환하는 방법을 제시하였다. 그러나 투영변환에 필요한 문서의 왜곡정보를 자동으로 찾아내는 방법은 명확하게 제시하지 못하였다. Yin[3,4] 등은 이진화된 문서 이미지를 이용하여 각 텍스트의 수평 라인을 단서로 하여 소멸점(vanishing point)을 검출함으로써 문서의 왜곡 정도를 자동으로 탐색하는 기법을 제시하였다. 그러나 사용된 이진화 기법인 Block-Otsu 알고리즘은 조명이 일정하지 못한 카메라 입력 이미지에 대한 이진화 기법의 한계와 잡음 영상을 처리하는 부분에 대한 기법이 없어 대상 문서 이미지 이외에 다른 문서 이미지가 중첩되어 하나의 이미지에 포함되어 있을 경우 문서의 외곽선을 검출하지 못하여 문서를 정규화하는데 있어 한계를 가지고 있다.

이러한 문제를 해결하기 위한 다른 접근 방식은 사용자가 스마트 폰으로 촬영한 문서 이미지에서 문서의 모서리를 사용자가 지정하여 보정하는 수동방식이 있다. 또한 Chen[5]등이 제안한 방법으로 사용자가 스마트폰의 위치를 조절하여 문서 이미지가 직사각형에 가깝게 조절되었을 때 자동으로 촬영되어 정규화 문서 이미지를 캡처하는 반자동 기법이 있다. 그러나 현재 많은 스마트폰 앱에서 채택하고 있는 수동방식의 경우 스마트폰의 작은 화면에서 사용자가 일일이 문서 이미지의 모서리를 지정하는 것은 정밀하지 못할 뿐 아니라 불편한 작업이 되고 있다. 반자동 기법 또한 사용자의 정밀한 카메라 조절 능력을 요구할 뿐 아니라 조명이나 주변의 잡음에 대하여 효과적으로 대처하지 못한다.

본 논문은 스마트 폰을 통해 획득한 문서 이미지를 사용자의 추가적인 작업 없이 자동화된 보정 알고리즘을 통하여 정규화된 문서 이미지를 생성하는 방법을 제안하고 있다. 특히, 스마트 폰 환경에서 발생할 수밖에 없는 불규칙한 조명 환경과 잡음 이미지 등을 처리하여 문서의 여백과 문서를 구분하는 윤곽선을 검출하여 문서를 보정하는 기법을 제안한다.

II. 문서 이미지 보정 기법

본 논문에서 제안한 스마트폰 문서 이미지 보정 기법의 전체 과정은 그림 2와 같다. 스마트폰을 이용해 촬영된 문서 이미지(I_S)를 그림 2와 같은 과정을 통해 왜곡 현상을 줄인 문서 이미지(I_T)로 보정한다.

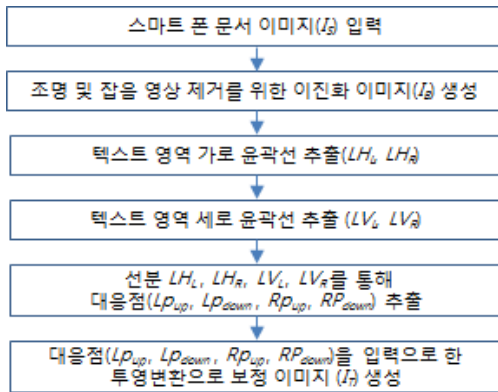


그림 2. 문서 이미지 보정 절차

Fig. 2 The rectification process of a document image

2.1. 이진 이미지 생성

스마트폰으로부터 입력된 문서 이미지(I_S)를 보정하기 위해서는 I_S 의 왜곡 정도를 검출해야 한다. 왜곡된 정도를 검출하기 위해서 문서에 포함된 텍스트와 배경을 분리하는 이진화 작업을 수행한다. Yin[3,4]등은 에지 검출 이진화 기법을 적용하였으나, 그림 3의 (a), (e)에서 보듯이 이진화 결과 이미지에 텍스트 영역이 아닌 잡음 이미지가 포함되어 용지 내에서 텍스트 영역을 검출하기 위한 전처리로 효과적이지 못했다. 또한 널리 알려진 이진화 기법인 Otsu 기법은 그림 3의 (b)에서 보듯이 조명에 따른 그림자 처리가 효율적이지 못하여 I_S 의 텍스트 영역이 손실되는 이진화 결과 이미지(I_B)를 생성함으로써 스마트폰과 같이 다양한 조명 환경에서 입력되는 스마트폰 문서 이미지의 전처리 이진화 기법으로는 적절하지 못하다. 본 연구에서는 스마트폰 문서 이미지의 효율적인 이진 이미지를 생성을 위해 기존에 제시된 다양한 기법을 적용해 보았으며, 그중에서 이미지의 객체 추출에 많이 사용되는 MSER(Maximally Stable Extremal Region)기법에 기반한 문서 이진화 기법인

MSER-b기법[6]이 본 연구의 환경에서 발생하는 잡음 제거 및 음영처리에서 우수한 성능을 나타내어 본 연구의 전처리 단계의 이진화 알고리즘으로 사용하였다. 그림 3의 (c), (g)는 MSER-b 이진화 기법을 적용한 이진 이미지를 보여준다. 이 기법은 조명 변화, 용지 배경 잡음 이미지 등을 적절하게 처리하여 텍스트 영역의 윤곽선 검출을 용이하게 한다.

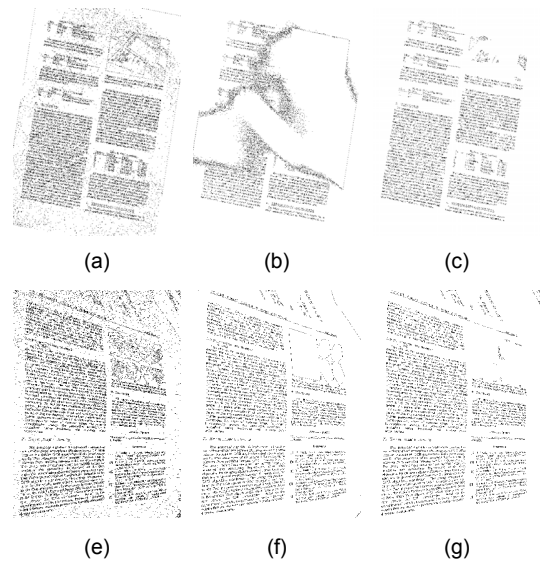


그림 3. 이진화 기법을 적용한 이미지 : (a)(e) 에지 이진화 이미지, (b)(f) Otsu 이진화 이미지, (c)(g) MSER-b 이진화 이미지
Fig. 3 Images applied binarization algorithm: (a)(e) image applied edge detection algorithm, (b)(f) image applied Otsu algorithm, (c)(g) image applied MSER-b algorithm

2.2. 텍스트 영역 윤곽선 검출

문서 이미지를 보정하기 위해서는 이진 이미지(I_B)에서 텍스트 영역의 윤곽선을 검출하고 텍스트 영역의 꼭지 점을 찾아야 한다. 본 논문에서는 가로 윤곽선을 먼저 검출하고 이 정보를 이용하여 세로 윤곽선을 검출한다. 텍스트 영역은 여러 개의 텍스트 줄로 구성되며 텍스트 줄은 각각의 텍스트로 구성된다. 이들 텍스트 줄이 형성하고 있는 선분을 조사함으로써 텍스트 영역 왜곡 정도를 파악한다. 본 논문에서는 MSER-b 기법으로 이진화 되어 있는 이미지에 에지 이진화 기법을 추가 적용하여 허프변환 선분 검출에 필요한 대상 점의 개수를 줄였다.

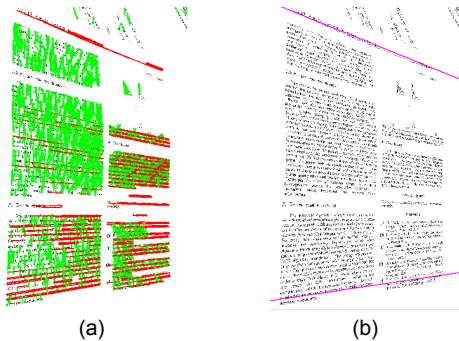


그림 4. 수평 윤곽선 검출: (a) 허프변환을 이용한 수평/수직 선분 검출, (b) 수평 윤곽선
Fig. 4 Horizontal contour line extraction: (a) horizontal/vertical line extraction using the Hough transform, (b) horizontal contour line

(1) 텍스트 영역 수평 윤곽선 검출

그림4의 (a)는 이진화 단계를 통해 생성된 I_B 에 허프 변환 수행을 통해 검출한 선분들을 보여준다. 허프 변환에 의해 검출된 선분들에는 텍스트 줄의 특성을 잘 나타내 주는 선분들과 그렇지 못한 선분들이 혼재되어 있다. 그러므로 검출된 선분 집합에서 텍스트 줄을 구성하는 선분을 추출하기 위하여 선분의 기울기를 이용해 (b)와 같은 수평 선분 집합을 생성한다.

$$L_{hough} = \{ \text{허프변환으로 검출된 모든 선분의 집합} \}$$

$$L_h = \{ L_{hough} \text{ 중 기울기가 } 45^\circ \text{보다 작은 선분의 집합} \}$$

$$x_{mid} = \text{이미지의 가로 픽셀 수} / 2$$

$$y_{mid} = \text{이미지의 세로 픽셀 수} / 2$$

L_h 을 이용하여 그림4의 (b)에 표시된 텍스트 영역 수평 윤곽선을 검출하는 알고리즘은 다음과 같다.

- ① L_h 에 속한 각 선분에 대하여 $f_i(x) = ax + b$ 의 직선 방정식을 구하고 $f_i(x_{mid})$ 를 구한다.
- ② 각 선분의 $f_i(x_{mid})$ 를 크기 순서로 정렬하여 y_s 집합을 구한다.
- ③ y_s 집합의 각 원소에 대하여 인접한 두 원소 사이의 거리 $dist_{i,i+1} = |y_s(i) - y_s(i+1)|$ 를 구하고 $dist_{i,i+1}$ 값이 줄 간격 임계값(T_{space})보다 작으면 i 선분과 $i+1$ 선분을 같은 텍스트 라인 그룹(H_G)으로

분류한다.

- ④ 생성된 텍스트 라인 그룹(H_G)에 대하여 그룹을 구성하는 원소의 개수가 선분 임계값(T_h)보다 큰 H_G 에 대하여 첫 번째 H_G 와 마지막 H_G 를 수평 윤곽선 후보로 선택한다.
- ⑤ 2개의 수평 윤곽선 후보 H_G 에 포함된 각 선분의 기울기 평균과 원소들의 y 값 평균을 이용해 수평 윤곽선을 확정한다.

(2) 텍스트 영역 세로 윤곽선 검출

그림4의 (a)에서 보듯이 I_B 에 허프 변환 수행을 통해 검출한 선분들은 오른쪽 수직 윤곽선을 거의 표시하지 못하고 수평 선분과 달리 연속적이어서 윤곽선 후보 선분을 특정하기 어렵다. 그러므로 그림5의 (a)와 같이 I_B 에서 세로 윤곽선 후보 점을 추출하여 이진 이미지(I_V)를 생성하고 I_V 를 대상으로 허프 변환 선분(L_v)을 생성하여 텍스트 영역 수직 윤곽선을 검출한다.

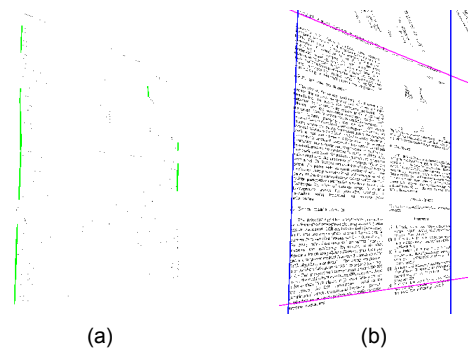


그림 5. 수직 윤곽선 검출: (a) 세로 윤곽선 후보 점과 허프 변환 선분, (b) 텍스트 영역 세로 윤곽선 검출
Fig. 5 Vertical contour line extraction: (a) hough transform line for vertical contour candidate points, (b) vertical contour line extraction in text area

그림5의 (b)와 같은 텍스트 영역 수직 윤곽선을 검출하는 알고리즘은 다음과 같다.

- ① I_B 의 왼쪽 방향과 오른쪽 방향을 각각 스캔하여 첫 번째 발견되는 점을 수직 윤곽선 후보 점으로 채택하여 이진 이미지 I_V 에 표시한다. 이때 앞서 구한 수평 윤곽선 바깥에 위치한 점은 후보 점에서 제외한다.

- ② 수평 윤곽선을 검출하는 알고리즘의 ① - ③ 단계와 유사한 과정을 거쳐 수직 선분 그룹(V_G)를 생성한다.
- ③ 생성된 V_G 에 대하여 그룹을 구성하는 원소의 개수가 선분 임계값(T_v) 보다 큰 V_G 에 대하여 첫 번째 V_G 와 마지막 V_G 를 수직 윤곽선 후보로 선택한다.
- ④ 2개의 수직 윤곽선 후보 V_G 에 포함된 선분의 기울기 평균과 원소들의 x 값 평균을 이용해 수직 윤곽선을 확정한다.

2.3. 대응점 검출 및 투시변환

그림5의 (b)와 같이 수평, 수직 윤곽선이 확정되면 각 윤곽선의 교차점을 이용하여 그림6의 (a)와 같이 텍스트 영역의 윤곽을 나타내는 사각형의 꼭지점(S_1, S_2, S_3, S_4)을 확정한다. 그리고 점 C_1 을 기준으로 직사각형을 형성하는 보정된 텍스트의 꼭지점(C_1, C_2, C_3, C_4)을 구하여 (S_1, S_2, S_3, S_4)를 (C_1, C_2, C_3, C_4)로 투시변환을 함으로써 그림1의 (b) 이미지를 그림6의 (c)와 같이 직사각형의 프린터 품질의 이미지로 보정한다.

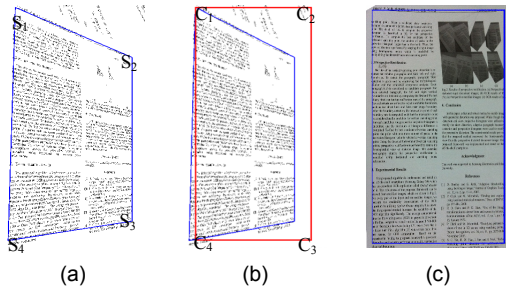


그림 6. 투시 변환: (a) 원본 이미지의 투시변환 꼭지점, (b) 투시변환 꼭지점에 대한 대응점, (c) 보정된 이미지
Fig. 6 Perspective transformation: (a) perspective transformation vertices of the original image, (b) corresponding points for the perspective transform vertices, (c) the corrected image

III. 실험 결과

본 논문에서는 실험을 위해 OpenCV 2.4에서 제공하는 히프변환 관련 함수와 투시변환 함수를 사용하

였다. 실험에서는 경험적 임계값($T_{space} = 40, T_h = 10, T_v = 3$)을 사용하였다.

그림 7은 현재 스마트 폰에서 서비스되고 있는 다른 제품과 본 논문의 결과를 비교하였다. 비교 프로그램은 앱 스토어를 통해 설치한 스마트폰용 스캔 프로그램으로 윤곽선 검출 기능을 포함한 제품 중에서 비교적 성능이 좋은 두 회사의 제품(A사, B사)을 선택하였다.

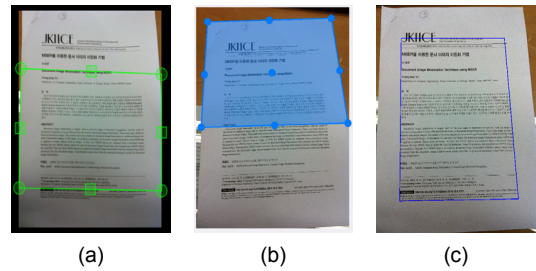


그림 7. 윤곽선 검출 결과: (a) A사 제품의 윤곽선 검출 결과, (b) B사 제품의 윤곽선 검출 결과, (c) 본 논문의 윤곽선 검출결과
Fig. 7 Contour extraction results: (a) a contour extraction result using production of A company, (b) a contour extraction result using production of B company, (c) a contour extraction result of this paper

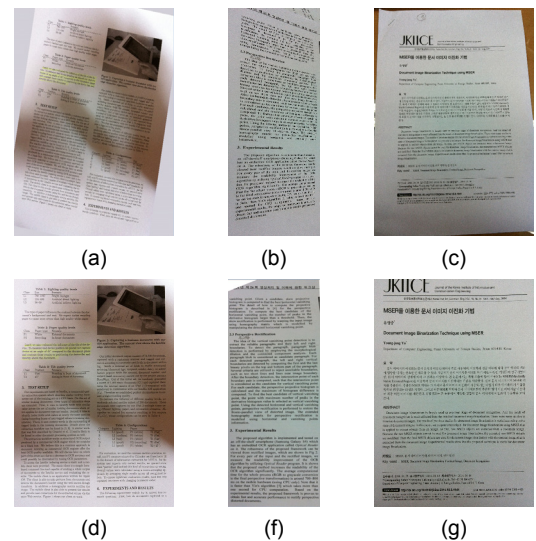


그림 8. 보정 결과: (a)(b)(c) 스마트폰을 통해 획득한 문서 이미지, (d)(e)(f) 본 논문의 보정 결과
Fig. 8 Rectification results: (a)(b)(c) a document image obtained through the smartphone, (d)(e)(f) the rectification results of this paper

그림 7에서 보듯이 A사 제품과 B사 제품은 텍스트 영역의 윤곽선을 정확하게 검출하지 못하는 반면 본 논문에서 제안한 방법은 텍스트 영역의 윤곽선을 정확하게 검출한다.

그림 8은 본 논문에서 제안한 방법을 통해 회전 왜곡 문서 및 원근 왜곡 문서와 보정 결과를 보여준다. 그 결과 본 논문에서 제시한 방법이 스마트폰 촬영 시 발생하는 문서의 왜곡을 효과적으로 처리하고 있음을 확인할 수 있다. 그림8의 (b)-(f)의 보정되지 않은 왜곡은 렌즈 왜곡으로 본 논문에서는 처리하지 않았다.

IV. 결 론

본 논문에서는 스마트폰 촬영을 통해 획득한 문서 이미지를 프린트 품질의 이미지로 보정하기 위한 방법을 제시하고 성능을 현재 서비스되고 있는 타사의 제품과 비교하였으며, 다양한 왜곡에 대하여 효과적으로 처리가 가능함을 실험을 통해 보였다.

그러나 본 논문은 객체 기반 이진화 기법인 MSER-b 기법을 사용함으로써 처리시간이 타사의 제품보다 많이 소요되는 단점을 가지고 있다. 향후 이러한 단점을 보완하여 스마트폰을 통해 서비스할 수 있도록 추가 연구가 필요하다.

감사의 글

이 논문은 부산대학교 자유과제 학술연구비(2년)에 의하여 연구되었음.



유영중(Young-jung Yu)

1996 부산대학교 전자계산학과 이학사
1998 부산대학교 전자계산학과 이학석사
2002 부산대학교 전자계산학과 이학박사
2002 ~ 부산외국어대학교 컴퓨터공학과 교수
※관심분야 : 컴퓨터 애니메이션, 시뮬레이션

REFERENCES

[1] Hee-Jung Yeom, Sung-Jong Eun, Taeg-Keun Whangbo. "Recognition of Car License Plate Using Geometric Information from Portable Device Image." *Journal of t The Korea Contents Association*. Vol. 10 No. 10, 2010.

[2] Jagannathan, L., and C. V. Jawahar. "Perspective correction methods for camera based document analysis." *Proc. First Int. Workshop on Camera-based Document Analysis and Recognition*. pp. 148-154, 2005.

[3] Yin, Xu-Cheng, Jun Sun, and Satoshi Naoi. "Perspective rectification for mobile phone camera-based documents using a hybrid approach to vanishing point detection." *International Workshop on Camera-Based Document Analysis and Recognition*. Vol. 2, pp.37-44, 2007.

[4] Yin, Xu-Cheng, et al. "Robust vanishing point detection for MobileCam-based documents." *Document Analysis and Recognition (ICDAR), 2011 International Conference on. IEEE*, pp. 136-140, 2011.

[5] Chen, F., Carter, S., Denoue, L., & Kumar, J., "SmartDCap: semi-automatic capture of higher quality document images from a smartphone." *Proceedings of the 2013 international conference on Intelligent user interfaces. ACM*, pp. 287-296 2013.

[6] Young-jung Yu, "Document Image Binarization Technique using MSER", *Journal of the Korea Institute of Information and Communication Engineering*. Vol. 18 No. 8, pp. 1941-1947, 2014.



문상호(Sang-Ho Moon)

한국기계연구원 정보지원실 연구원
부산대학교 컴퓨터공학과 공학석사
부산대학교 컴퓨터공학과 공학박사
위덕대학교 컴퓨터공학부 조교수
부산외국어대학교 컴퓨터공학과 교수
※관심분야 : DB, GIS, 정보시스템 감리, 전산교육



박성호(Seong-ho Park)

1996년 부산대학교 전자계산학과 이학사
1998년 부산대학교 대학원 전자계산학과 이학석사
2002년 부산대학교 대학원 전자계산학과 이학박사
2002년 9월 ~ 현재 부산대학교 정보전산원 교수
※관심분야 : 멀티미디어 통신, 컴퓨터 비전, 2D/3D 영상 및 비디오 처리