

PDA 응용을 위한 명함 영상의 회전 보정

정희원 박준효*, 장익훈**, 김남철***

Skew Correction of Business Card Images for PDA Application

Jun-Hyo Park*, Ick-Hoon Jang**, and Nam-Chul Kim*** *Regular Members*

요 약

본 논문에서는 PDA에 장착된 카메라를 사용하여 획득한 명함 영상에 대한 회전 보정 알고리즘을 제안한다. 제안된 방법은 크게 블록적응 이진화부, 스트라이프 생성부, 회전각 계산부, 그리고 영상 회전부로 구성된다. 블록적응 이진화부에서는 입력 영상의 불균일한 조명이나 그림자의 영향을 감소시키기 위하여 입력 영상을 블록별로 이진화한다. 스트라이프 생성부에서는 인접한 문자와 이들로 이루어진 문자열들을 병합하여 문자열 클러스터를 생성하고, 생성된 클러스터 중에서 회전각 계산에 유용한 것들을 스트라이프로 분류한다. 회전각 계산부에서는 스트라이프들의 방향각을 중심 모멘트를 이용하여 계산하고 이들 방향각을 평균하여 입력 영상의 회전각으로 결정한다. 영상 회전부에서는 입력 영상을 계산된 회전각 만큼 회전시킨다. 실험결과 제안된 방법은 여러 종류의 명함을 다양한 주변 여건에서 PDA로 획득한 시험 영상에 대하여 약 93%의 회전 보정률을 보였다.

Key Words : skew correction; PDA; binarization; stripe; direction angle.

ABSTRACT

We present an efficient algorithm for skew correction of business card images obtained by a PDA (personal digital assistant) camera. The proposed method is composed of four parts: block adaptive binarization (BAB), stripe generation, skew angle calculation, and image rotation. In the BAB, an input image is binarized block by block so as to lessen the effect of irregular illumination and shadow over the input image. In the stripe generation, character string clusters are generated merging adjacent characters and their strings, and then only clusters useful for skew angle calculation are output as stripes. In the skew angle calculation, the direction angles of the stripes are calculated using their central moments and then the skew angle of the input image is determined averaging the direction angles. In the image rotation, the input image is rotated by the skew angle. Experimental results shows that the proposed method yields skew correction rates of about 93% for test images of several types of business cards acquired by a PDA under various surrounding conditions.

I. 서 론

현대인은 일상생활에서 신문, 잡지, 명함 등과 같은 문서 형식을 갖춘 매체를 접하고 있다. 이들 중에서 명함을 가장 흔히 접할 수 있는 문서 형식의

매체로서 영업 사원은 물론 교사나 엔지니어와 같은 보통 사람들에게도 중요한 홍보 수단으로 사용되고 있다. 그래서 사람을 처음 만났을 때 인사와 함께 명함을 주고받는 것은 일상적이 되었다. 이에 따라 대부분의 사람들이 많은 양의 명함을 보유

* 삼성전자(주) TN총괄 무선사업부 (junhyo.park@samsung.com), ** 경운대학교 디지털전자정보공학부 (ihjang@ikw.ac.kr)
*** 경북대학교 전자전기컴퓨터학부 (nckim@ee.knu.ac.kr)
논문번호 : 030343-0807, 접수일자 : 2003년 8월 7일
※ 본 연구는 삼성전자(주)의 연구비 지원으로 수행되었습니다.

하게 되었고, 이들을 직접 가지고 다니는 것보다는 그 정보를 효율적으로 저장하고 관리할 필요성이 대두되고 있다. 지금까지는 대부분의 사람들이 명함의 정보를 수첩에 메모하거나 컴퓨터에 입력하여 관리하여왔다. 하지만 이러한 명함 관리는 정보를 입력하는데 시간과 노력이 필요하고 보관 및 검색 등의 측면에서 볼 때 상당히 비효율적이다.

최근에 PDA와 같은 휴대 단말기가 이동성 및 휴대성이 증시되는 현대 사회에서 효용 가치가 높아짐에 따라 그 사용이 증가하고 있다. PDA는 부속 카메라를 통하여 명함을 쉽게 디지털화하여 명함 영상을 얻을 수 있다. 이러한 PDA 카메라로부터 얻어지는 명함 영상을 저장 관리하면 명함 정보의 효율적인 관리가 가능하게 되고 명함의 문자 인식이 더욱 효율적으로 이루어질 수 있게 된다. 그러나 PDA 카메라를 통하여 디지털화하는 과정에서는 손의 흔들림이나 촬영 오류로 인하여 영상이 회전된 경우가 발생된다. 일반적으로 문서 영상의 회전각이 2~3도인 경우는 문서 분석 및 문자 인식의 정확도가 떨어지며, 5도 이상이 되면 정확도가 급격히 떨어지는 것으로 알려져 있다^[1]. 그러므로 문서 영상의 분석이나 문자 인식 이전에 문서 영상의 회전각을 검출하고 보정하는 것은 필수적인 전처리 과정이라고 할 수 있다.

지금까지 문서 영상에 대한 회전 보정에 관한 많은 연구가 이루어져 왔다^{[2],[7]}. Dengel^[2]은 왼쪽 가장자리 탐색(left margin search)을 이용하여 문서 영상의 회전각을 검출하는 방법을 제안하였다. 이 방법에서는 문서 영상의 각 행(row)의 첫 번째 검은색 화소의 위치를 검출한 후 이들로부터 왼쪽 가장자리의 오른쪽 경계에 의한 직선을 정하고 이를 영상의 수직 축과 비교하여 영상의 회전각을 구하였다. 그러나 이 방법은 문서의 왼쪽 가장자리 영역에 그림이 있는 경우에는 회전각을 정확하게 검출하기가 어렵다.

D. S. Le^[3] 등은 Hough 변환을 이용한 문서 영상의 회전 보정 방법을 제안하였다. 이 방법에서는 영상의 회전각을 검출할 때 문자 영역이 아닌 영역(non-textual region)의 영향을 줄이기 위하여 먼저 영상을 특정한 크기의 블록으로 나누고 각 블록을 문자 영역 블록과 문자 영역이 아닌 블록으로 분류한다. 그리고 문자 영역 블록 내의 각 문자의 가장 하위 화소만 남기고 다른 화소들은 모두 제거한다. 그런 다음 Hough 변환을 이용하여 남은 화소들의 회전각을 구하였다. 이 방법은 영역 분류를 사용하

지 않는 경우에 비해서는 성능이 우수하지만 한글과 같이 하나의 문자가 하나의 연결된 화소들로 구성되지 않은 문자 영상의 회전각을 정확하게 검출하기가 어렵다.

Avanindra^[4] 등은 Yan^[5]이 제안한 문서 영상에 대한 줄간 상호 상관도(interline cross-correlation)를 이용한 회전각 검출 방법을 개선하여 문서 영상에서 랜덤하게 추출한 여러 개의 사각 창에 대하여 상호 상관도를 구한 후 통계적인 방법을 통하여 문서의 회전각을 검출하였다. 그러나 이 방법은 문자 영역이 문서에 많이 포함된 영상에 대해서는 좋은 결과를 보이지만, 그림이나 표가 있거나 문서의 배경에 질감(texture)이 있는 영상에 대해서는 회전각을 정확하게 검출하기가 어렵다. 이외에도 투영(projection)을 이용한 회전 보정 방법^[6], Fourier 변환을 이용한 회전 보정 방법^[7] 등이 있다.

기존의 영상 회전 보정 방법들은 주로 문서 영상에 대한 방법들이다. 문서 영상은 대부분 고해상도의 스캐너로 얻어지고 문자열이 문서 전 영역에 걸쳐 많이 존재하기 때문에 회전 보정이 비교적 쉽게 이루어질 수 있다. 반면에 PDA에 장착된 저해상도의 카메라를 이용하여 얻어지는 명함 영상은 불균일한 조명이나 그림자의 영향을 받을 수 있고 문서 영상에 비하여 문자열의 수가 확연히 적을 뿐더러 그 위치가 일정하지 않고 명함에 따라 배경 무늬나 사진, 용지의 질감 등 회전 보정에 있어 여러 가변 요소도 포함하고 있다. 따라서 PDA 카메라를 통하여 얻은 회전된 명함 영상에 대해서는 기존의 방법을 적용하여 좋은 결과를 기대하기가 어렵다.

본 논문에서는 PDA 카메라를 사용하여 열악한 주변 여건에서 얻어지는 명함 영상에 대한 효율적인 회전 보정 방법을 제안한다. 제안된 방법은 블록적응 이진화(block adaptive binarization: BAB)부, 스트라이프 생성(stripe generation)부, 회전각 계산(skew angle calculation)부, 그리고 영상 회전(image rotation)부로 구성된다. 블록적응 이진화부에서는 불균일한 조명이나 그림자의 영향을 줄이기 위하여 입력 영상을 블록별로 이진화한다. 먼저 입력 영상을 8×8 크기의 블록으로 나누고 각 블록을 문자 블록(character block: CB)과 배경 블록(background block: BB)으로 분류한다. 각 문자 블록을 인접한 8개의 블록들과 그룹핑하여 24×24 크기의 블록을 생성하고 이 블록에 대하여 Otsu의 방법^[8]으로 이진화를 위한 역치(threshold)를 결정한다. 그런 다음 이 역치를 사용하여 원래의 문자 블록을

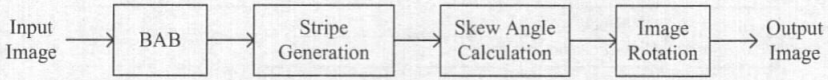


그림 1. 제안된 회전 보정 방법의 블록도
 Fig. 1. Block diagram of the proposed skew correction method.

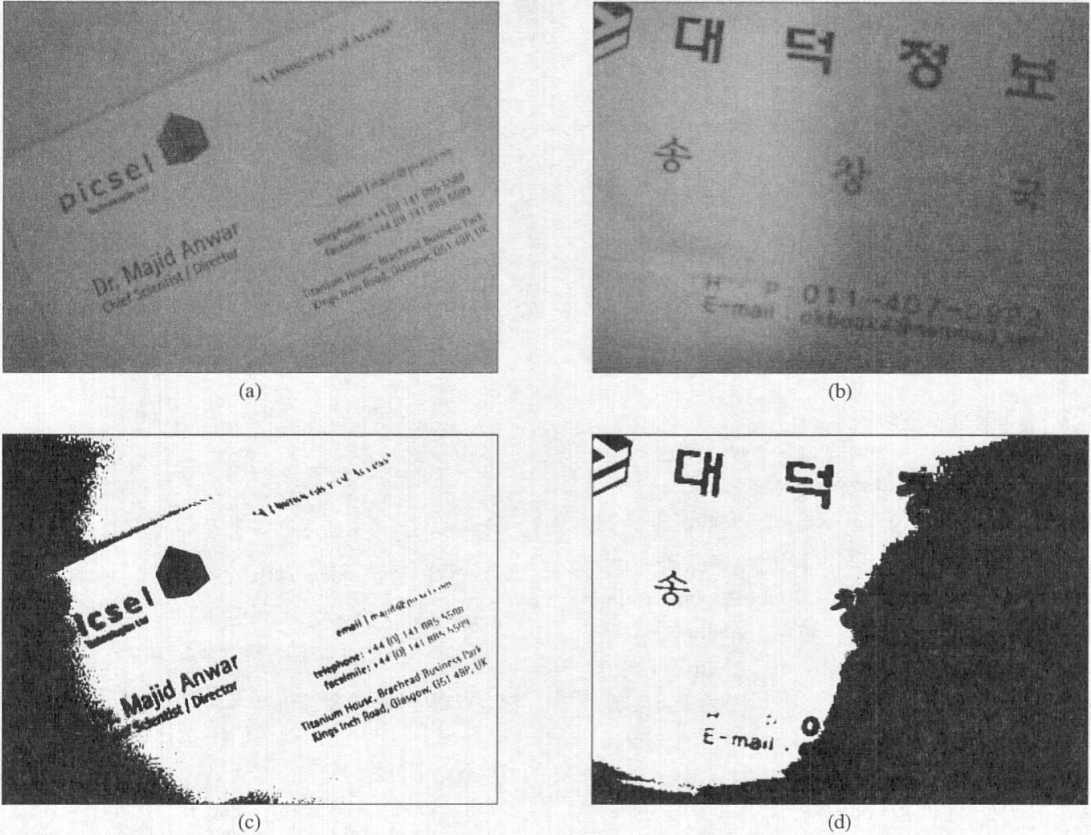


그림 2. PDA 카메라로 획득한 명함 영상과 전역 이진화에 의한 이진 영상: (a) 입력 영상 1, (b) 입력 영상 2, (c) 입력 영상 1에 대한 이진 영상, (d) 입력 영상 2에 대한 이진 영상
 Fig. 2. Business card images obtained by a PDA camera and the result images of GB: (a) input image 1, (b) input image 2, (c) binary image for input image 1, and (d) binary image for input image 2.

이진화한다. 스트라이프 생성부에서는 인접한 문자와 이들로 이루어진 문자열들을 병합하여 문자열 클러스터(cluster)를 생성하고 그 크기와 이심률(eccentricity)이 주어진 역치 이상인 클러스터를 스트라이프로 분류한다. 회전각 계산부에서는 스트라이프들의 방향각을 중심 모멘트를 이용하여 계산하고 이중에서 빈도수가 가장 많은 방향각들을 평균하여 영상의 회전각으로 결정한다. 영상 회전부에서는 입력 영상을 결정된 회전각만큼 회전한다. 제안된 방법의 성능을 평가하기 위하여 여러 종류의 명함을 다양한 주변 환경에서 PDA 카메라를 이용하

여 획득한 시험 영상에 제안된 방법을 적용한다. 그리고 그 성능을 Avanindra의 방법^[4]의 성능과 비교한다. 실험 결과에서는 제안된 방법이 명함 영상의 회전을 잘 보정한다는 것을 보인다.

II. 제안된 명함 영상의 회전 보정

일반적으로 명함 영상에는 이름, 주소, 전화번호, email 주소 등의 문자열이 포함되어 있고, 이 문자열들은 명함의 가로축과 평행하다. 제안된 회전 보정 방법에서는 이러한 사실을 바탕으로 회전된 명

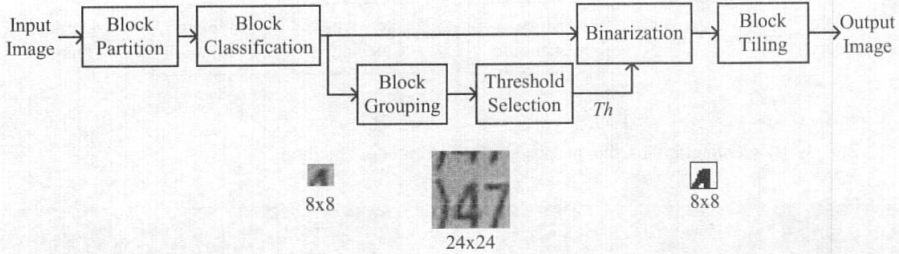


그림 3. 블록적응 이진화 과정
Fig. 3. Procedure of BAB.

함 영상의 문자열과 가로축과의 각을 그 명함 영상의 회전각으로 간주하고 그 각을 구한다. 그림 1은 제안한 회전 보정 방법의 블록도를 보여준다. 제안된 방법은 블록적응 이진화부, 스트라이프 생성부, 회전각 계산부, 영상 회전부의 4개 단계로 구성된다.

1. 블록적응 이진화

제안된 회전 보정 방법에서는 입력 명함 영상의 이진 영상에서 문자열들의 방향각을 구한다. 이를 위하여 먼저 명함 영상의 이진화 단계가 필요하다. 영상을 이진화할 때 1개의 역치만을 사용하여 이진화하는 경우를 전역 이진화(global binarization: GB)라고 한다. 그림 2는 PDA 카메라로 획득한 640×480 크기의 명함 영상과 이들을 Otsu의 방법^[8]으로 이진화를 위한 역치를 정하여 전역 이진화한 결과 영상을 보여준다. 그림 2(a)와 2(b)의 영상에서 보면 각각 불균일한 조명과 그림자가 있는 열악한 조건의 영상임을 알 수 있다. 그리고 이들을 전역 이진화한 그림 2(c)와 2(d)의 이진 영상에서 보면 불균일한 조명과 그림자의 영향으로 일부 문자열이 나타나지 않는다는 것을 알 수 있다.

본 논문에서 제안하는 블록적응 이진화는 이러한 문제에 대한 해결책으로 열악한 조건의 영상에서도 문자열을 잘 추출할 수 있도록 고려되었다. 그림 3은 블록적응 이진화 과정을 보여준다. 먼저 입력된 영상을 8×8 크기의 블록들로 나누고, 각 블록을 문자 블록과 배경 블록으로 분류한다. 블록 분류에서는 문자 블록이 배경 블록보다 신호의 활동도가 큰 특성을 이용하여, 블록의 활동도가 역치보다 큰 블록은 문자 블록으로 분류하고 그렇지않은 블록은 배경 블록으로 분류한다^[9]. 본 논문에서는 k번째 블록의 활동도를 다음과 같이 저주파 영역에서의 DCT 계수의 절대 합(absolute sum)으로 정하였다.

$$E^k = \sum_u \sum_v |D_{uv}^k|, \quad 0 \leq u, v \leq 7 \quad (1)$$

여기서 D_{uv}^k 는 k번째 블록의 (u, v) 주파수에서의 DCT 계수를 나타낸다. 따라서 k번째 블록의 분류는 다음과 같이 표현된다.

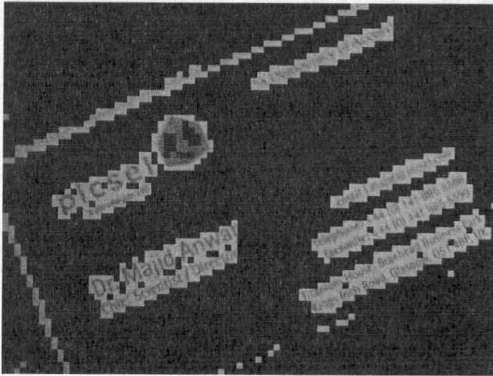
$$k \in \begin{cases} \Omega_C, & \text{if } E^k \geq Th_B \\ \Omega_B, & \text{otherwise} \end{cases} \quad (2)$$

여기서 Ω_C 와 Ω_B 는 각각 문자 블록과 배경 블록의 인자 집합(index set)을 나타내고, Th_B 는 블록 분류를 위한 역치를 나타낸다. 본 논문에서는 역치 Th_B 를 영상 전체에서의 E^k 의 평균값으로 정하였다. 그림 4는 그림 2(a)와 2(b)의 영상에 대한 블록 분류 결과 영상을 보여준다. 그림 4에서 그레이(gray) 부분은 문자 블록을 나타내고, 검은 부분은 배경 블록을 나타낸다. 그림 4에서 보면 블록들이 잘 분류되었음을 알 수 있다.

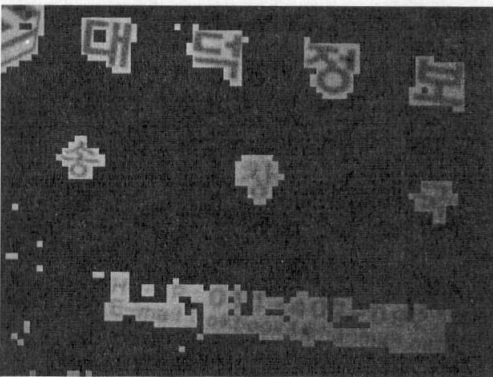
분류된 k번째 문자 블록은 인접한 8개의 블록들과 그룹핑하여 24×24 크기의 블록을 생성하고 이 블록에 대하여 Otsu의 방법으로 이진화를 위한 역치 Th^k 를 찾는다. 그런 다음 Th^k 를 기준으로 원래의 k번째 8×8 크기의 문자 블록 $u^k(i, j)$ 를 다음과 같이 이진화 한다.

$$b^k(i, j) = \begin{cases} 1, & \text{if } u^k(i, j) \geq Th^k \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

그런 다음 이진화된 각 8×8 블록들을 원 위치에 붙이는 블록 붙이기(block tiling)를 하여 이진화된 영상을 얻는다. 그림 5는 그림 2(a)와 2(b)의 영상을 블록적응 이진화를 사용하여 이진화한 결과 영



(a)



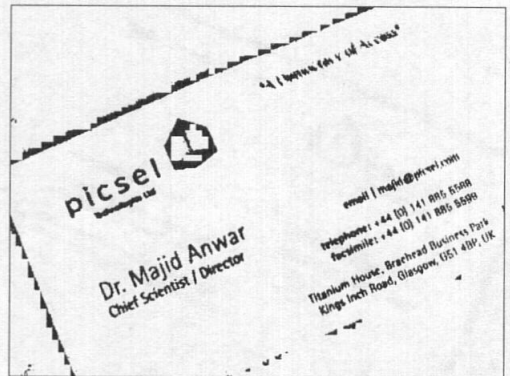
(b)

그림 4. 그림 2(a)와 2(b)의 영상에 대한 블록 분류 결과 영상: (a) 그림 2(a)에 대한 결과 영상, (b) 그림 2(b)에 대한 결과 영상, 여기서 그레이 부분은 문자 블록을, 검은 부분은 배경 블록을 각각 나타낸다
 Fig. 4. Result images of the block classification for the images in Figs. 2(a) and 2(b): (a) result image for Fig. 2(a) and (b) result image for Fig. 2(b), where gray parts represent CBs and black ones BBs.

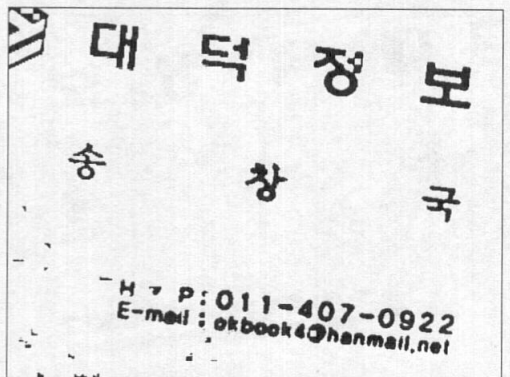
상을 보여준다. 그림 5와 그림 2(a), 2(b)의 이진 영상을 비교해보면 블록적응 이진화에 의하여 이진화한 결과 영상은 불균일한 조명과 그림자에 상관없이 전역 이진화에 의하여 이진화한 결과 영상보다 문자열이 훨씬 더 잘 드러나 있음을 알 수 있다.

2. 스트라이프 생성

스트라이프 생성부에서는 먼저 문자열을 클러스터로 변환하는 클러스터 생성 과정을 수행한다. 그런 다음 문자열에 의한 방향각을 계산하기에 유용한 길쭉한 모양의 클러스터를 스트라이프로 선택한다. 클러스터 생성부에서는 먼저 문자열 내의 문자들이 서로 접하게 하기 위하여 입력 이진 영상에 대하여 수평 방향으로 서브샘플링(subsampling)을



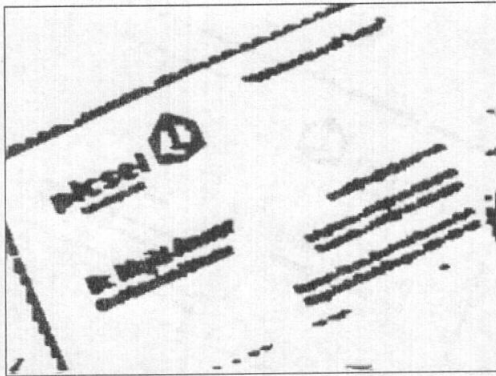
(a)



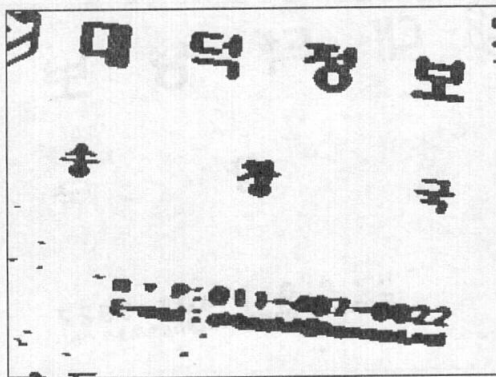
(b)

그림 5. 그림 2(a)와 2(b)의 영상에 대한 블록적응 이진화 결과 영상: (a) 그림 2(a)에 대한 결과 영상, (b) 그림 2(b)에 대한 결과 영상
 Fig. 5. Result images of the BAB for the images in Figs. 2(a) and 2(b): (a) result image for Fig. 2(a) and (b) result image for Fig. 2(b).

수행한다. 그리고 서로 접한 문자들이 뭉쳐져 길쭉한 모양의 클러스터로 변환되게 하기 위하여 형태학적 필터링(morphological filtering)인 확장 연산(dilation operation)^[10]을 수행한다. 확장 연산 후 상하가 접하게 된 문자열을 분리하기 위하여 수축 연산(erosion operation)^[10]을 수행하고, 결과 영상의 가로 대 세로의 비를 원 영상의 비와 같게 하기 위하여 수직 방향으로 서브샘플링을 수행한다. 그림 6은 그림 5의 이진 영상에 대하여 클러스터가 생성된 결과 영상을 보여준다. 그림 6에서 보면 대부분의 문자열들은 길쭉한 모양의 클러스터로 변하지만 마크와 잠음 블록 등과 같은 길쭉하지 않은 모양의 클러스터도 생기는 것을 알 수 있다. 따라서 이러한 클러스터들 중에서 방향각 계산에 유용한 문자열에 의한 스트라이프를 분류해내는 단계가 필요하다.



(a)



(b)

그림 6. 그림 5의 영상에 대한 클러스터 생성 결과 영상: (a) 그림 5(a)에 대한 결과 영상, (b) 그림 5(b)에 대한 결과 영상

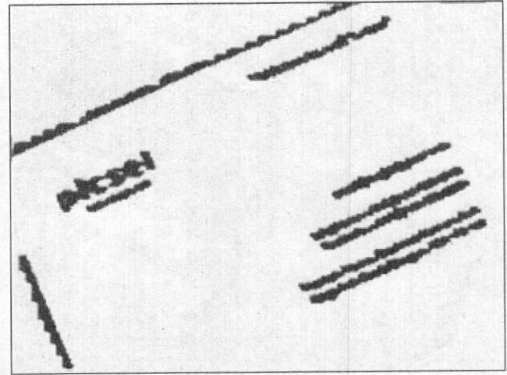
Fig. 6. Result images of the cluster generation for the images in Fig. 5: (a) result image for Fig. 5(a) and (b) result image for Fig. 5(b).

원하지 않는 클러스터들을 제거하기 위하여 다음과 같이 정의되는 중심모멘트를 이용한 클러스터들의 크기 μ_{00} 와 클러스터들의 긴 정도를 나타내는 기준인 이심률 $e^{[10]}$ 를 이용한다.

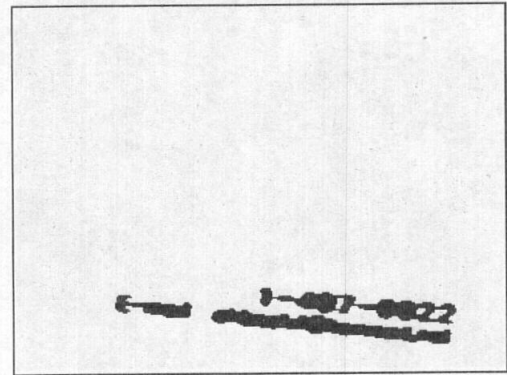
$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q, \quad p, q = 0, 1, 2, \dots (4)$$

$$e = \frac{4\mu_{11}^2 + (\mu_{20} - \mu_{02})^2}{(\mu_{20} + \mu_{02})^2} \quad (5)$$

여기서 \bar{x} 와 \bar{y} 는 각각 물체의 수평 및 수직 방향의 무게 중심을 나타낸다. 생성된 클러스터들 중에서 $\mu_L \leq \mu_{00} \leq \mu_H$ 와 $e \geq e_{th}$ 의 조건을 만족하는 클러스터를 스트라이프로 분류한다. 여기서 μ_L 과 μ_H 는 클러스터의 크기 μ_{00} 에 대한 역치들을 나타



(a)



(b)

그림 7. 그림 6의 영상에 대한 스트라이프 분류 결과 영상: (a) 그림 6(a)에 대한 결과 영상, (b) 그림 6(b)에 대한 결과 영상

Fig. 7. Result images of the stripe generation for the images in Fig. 6: (a) result image for Fig. 6(a) and (b) result image for Fig. 6(b).

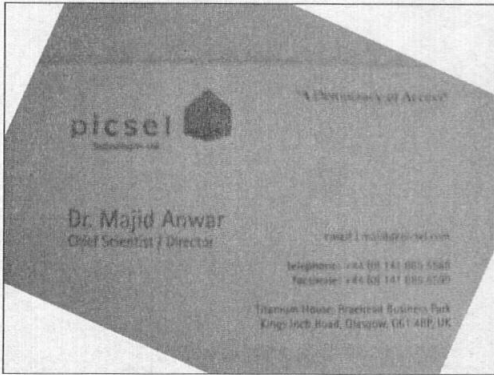
내고, e_{th} 는 이심률 e 에 대한 역치를 나타낸다. 그림 7은 그림 6의 영상에 대한 스트라이프 분류 결과 영상을 보여준다. 그림 6의 영상에서 보면 마크나 잡음 블록과 같은 원하지 않는 클러스터들이 잘 제거되었다는 것을 알 수 있다.

3. 회전각 계산

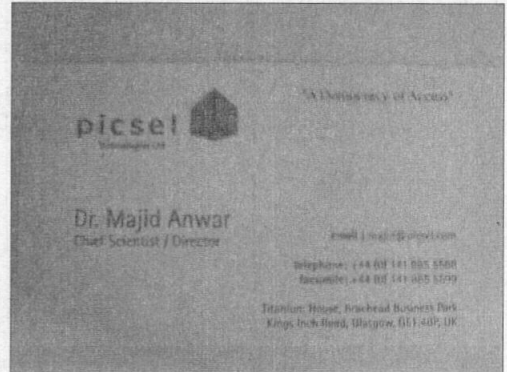
회전각 계산부에서는 각 스트라이프의 방향각을 다음과 같이 계산한다^[10].

$$\theta = \frac{1}{2} \arctan \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right) \quad (6)$$

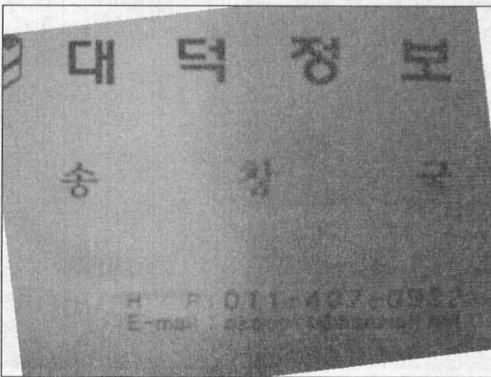
여기서 θ 는 그 범위가 $0 \leq \theta < 180$ 인 방향각을 나타낸다. 그리고 회전각을 추정하기 위하여 다음과 같이 방향각 θ 를 반올림하여 정수화한다.



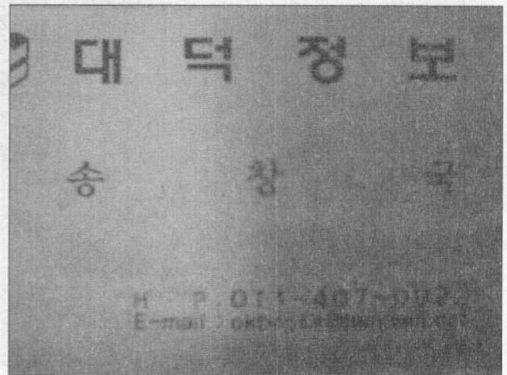
(a)



(a)



(b)



(b)

그림 8. 그림 2(a)와 2(b)의 영상에 대한 영상 회전 결과 영상: (a) 그림 2(a)에 대한 결과 영상, (b) 그림 2(b)에 대한 결과 영상

Fig. 8. Result images of the image rotation for the images in Figs. 2(a) and 2(b): (a) result image for Fig. 2(a) and (b) result image for Fig. 2(b).

$$\theta_n = \text{round}(\theta) \bmod 180 \quad (7)$$

여기서 θ_n 은 반올림된 정수형 방향각을 나타낸다. 다음으로 빈도수가 가장 큰 θ_n 을 구하기 위하여의 θ_n 히스토그램 $h(\theta_n)$ 을 구한다. 그리고 $h(\theta_n)$ 을 이용하여 빈도수가 가장 큰 정수형 방향각 θ_r 을 다음과 같이 구한다.

$$\theta_r = \arg \max_{0 \leq \theta_n < 180} \left\{ \frac{1}{3} \sum_{\theta_i = \theta_n - 1}^{\theta_n + 1} h(\theta_i \bmod 180) \right\} \quad (8)$$

최종적으로 방향각은 다음과 같이 추정한다.

$$\theta_s = \langle \theta \rangle, \quad \theta_r - 0.5 \leq \theta < \theta_r + 0.5 \quad (9)$$

그림 9. 그림 8의 영상에 대한 모서리 채움 결과 영상: (a) 그림 8(a)에 대한 결과 영상, (b) 그림 8(b)에 대한 결과 영상

Fig. 9. Result images of the corner filling for the images in Fig. 8: (a) result image for Fig. 8(a) and (b) result image for Fig. 8(b).

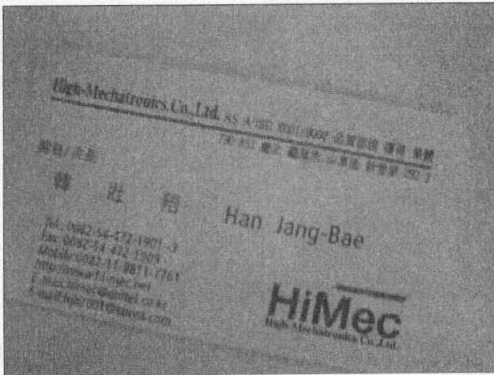
여기서 $\langle \cdot \rangle$ 는 평균 연산을 나타낸다.

4. 영상 회전

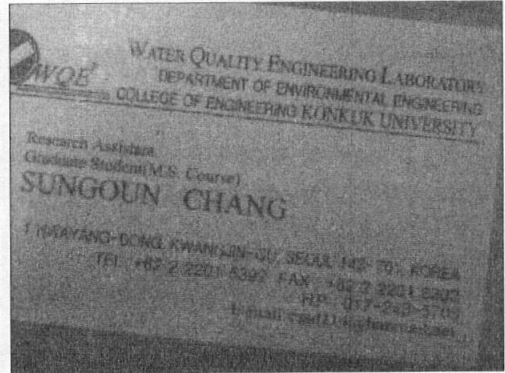
영상 회전부에서는 입력 영상의 추정된 회전각이 θ_s 일때 그 영상을 $-\theta_s$ 만큼 회전시킨다. 이러한 영상의 회전은 다음과 같이 입력 영상의 화소(pixel) (x, y) 를 출력 영상의 좌표(point) (x', y') 으로의 매핑^[11]으로 수행될 수 있다.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (10)$$

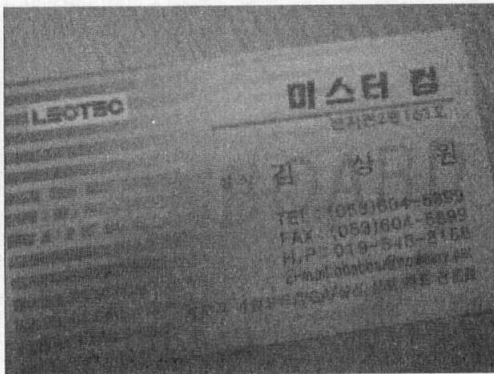
그러나 이러한 방법에 의하여 영상을 회전하면 입력 영상과 출력 영상의 화소들이 1:1로 매핑이 되지 않아서 회전된 영상 내에서 구멍(hole)이 발생할



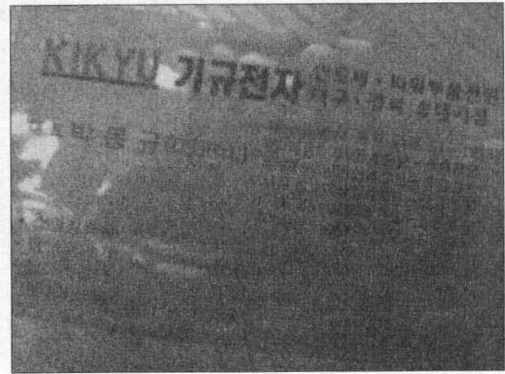
(a)



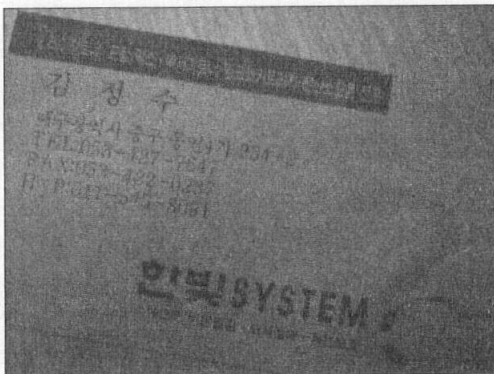
(b)



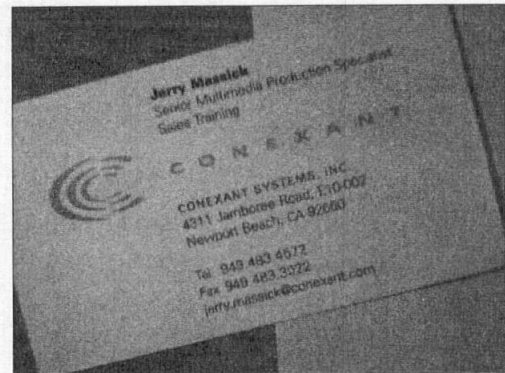
(c)



(d)



(e)



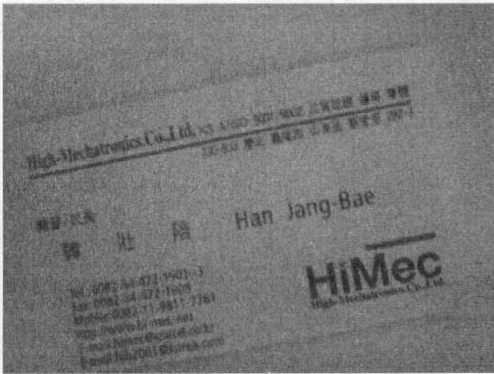
(f)

그림 10. 여러 가지 형태의 명함을 다양한 주변 여건 하에서 PDA로 획득한 명함 영상: (a) 흰색 종이 위의 보통 명함, (b) 나뭇결 책상 위의 보통 명함, (c), (d) 나뭇결 책상 위의 패턴을 가진 특별한 명함, (e) 나뭇결 책상 위의 질감을 가진 특별한 명함, (f) 2가지 색 책상 위의 보통 명함

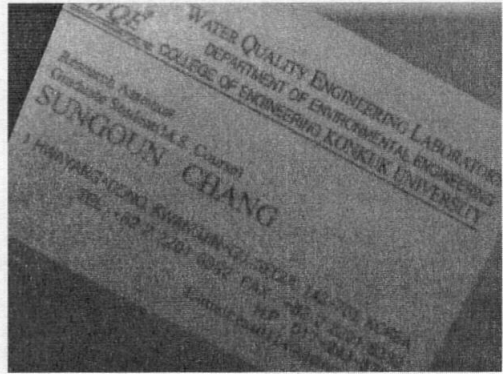
Fig. 10. Images of several types of business cards acquired by the PDA under various surrounding conditions: (a) ordinary business card on white paper, (b) ordinary business card on wood grain desk, (c), (d) special business cards with patterns on wood grain desk, (e) special business card of textured surface on wood grain desk, and (f) ordinary business card on two-tone desk.

수 있다. 따라서 본 논문에서는 다음과 같이 출력 영상의 좌표 (x', y') 를 입력 영상의 좌표 (x, y) 로 매핑하는 역매핑^[12]으로 영상을 회전한다.

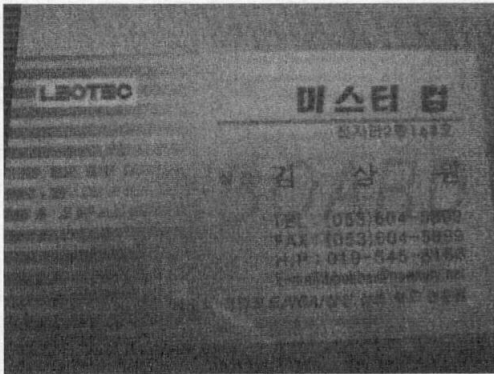
$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \quad (11)$$



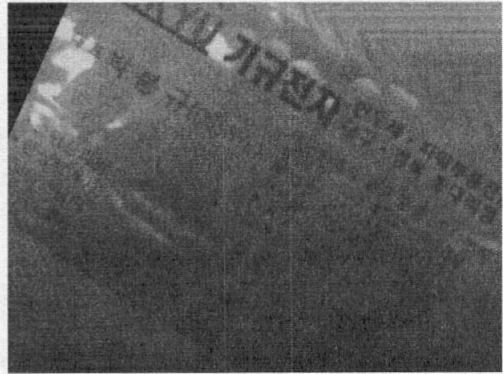
(a)



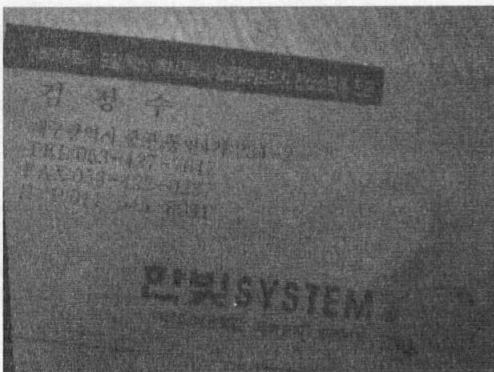
(b)



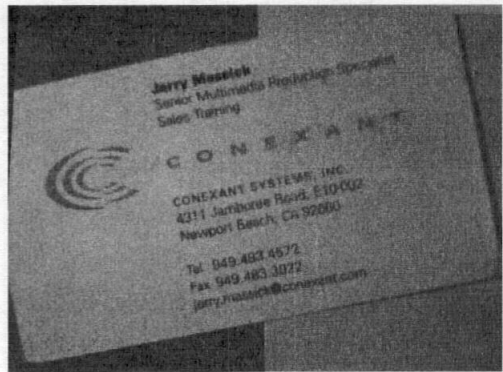
(c)



(d)



(e)



(f)

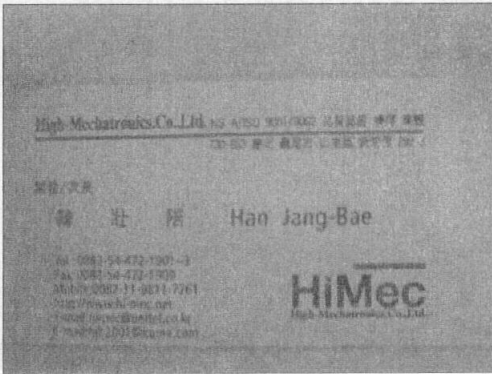
그림 11. 그림 10의 영상에 대하여 Avanindra의 방법으로 회전 보정한 결과 영상

Fig. 11. Result images of the skew correction for the images in Fig. 10 by Avanindra's method.

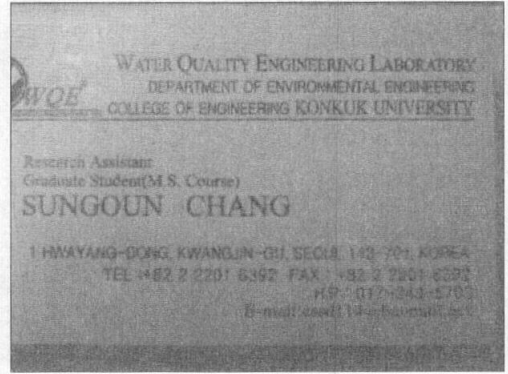
이러한 역매핑에 의한 영상의 회전은 화소 (x', y') 에 해당하는 좌표 (x, y) 가 삼각함수 연산으로 인하여 정수형의 화소가 아닐 수가 있다. 이러한 경우는 다음과 같이 정수형의 화소 (x', y') 에서의 출력 영상 g 의 밝기 값을 입력 영상 f 의 실수형의 좌표 (x, y) 에서 가장 가까운 4개의 화소의 밝기 값을

쌍선형 보간(bilinear interpolation)^[10]하여 구한다.

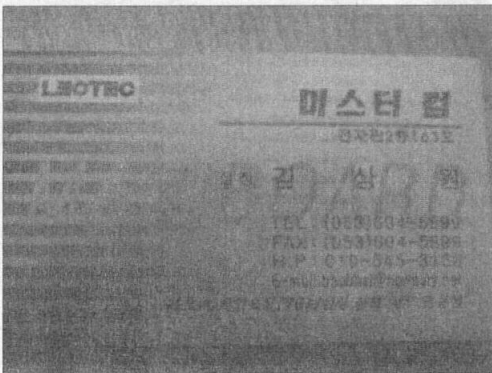
$$\begin{aligned}
 g(x', y') = & (1 - \Delta x)(1 - \Delta y) f(m, n) \\
 & + (1 - \Delta x)\Delta y f(m, n + 1) \\
 & + \Delta x(1 - \Delta y) f(m + 1, n) \\
 & + \Delta x \Delta y f(m + 1, n + 1)
 \end{aligned} \quad (12)$$



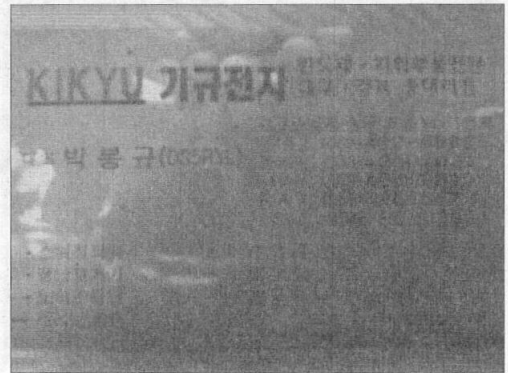
(a)



(b)



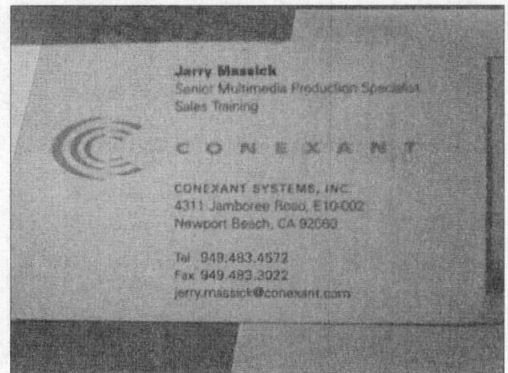
(c)



(d)



(e)



(f)

그림 12. 그림 10의 영상에 대하여 제안된 방법으로 회전 보정한 결과 영상
 Fig. 12. Result images of the skew correction for the images in Fig. 10 by the proposed method.

여기서 $m = \text{int}(x)$, $n = \text{int}(y)$, $\Delta x = x - m$, 그리고 $\Delta y = y - n$ 이다.

그림 8은 그림 2(a)와 2(b)의 영상을 역매핑 및 쌍선형 보간을 이용하여 회전한 결과 영상을 보여 준다. 그림 8에서 보면 영상을 회전하게 되면 네 귀퉁이에 빈 공백이 생긴다. 이는 영상의 보기에 좋지

않으므로 수평 방향으로 빈 공백의 값이 아닌 가장 가까운 화소의 값으로 채워 넣는 모서리 채움 (corner filling)을 수행한다. 그림 9는 그림 8의 영상에 대하여 모서리 채움을 수행한 결과 영상을 보여 준다. 그림 9에서 보면 모서리가 채워진 영상은 모서리를 채우기 전의 영상에 비하여 보기에 더 좋

표 1. Avanindra의 방법과 제안된 방법의 회전 보정률
Table 1. Skew correction rates of Avanindra's method and the proposed method.

Type of business card	Surrounding condition	Avanindra		Proposed	
		with GB	with BAB	with GB ¹	with BAB ²
Ordinary	Good	36%	46%	70%	96%
	Ill	34%	40%	66%	92%
Special	Good	30%	42%	60%	96%
	Ill	16%	38%	54%	88%
Total		29%	41.5%	62.5%	93%

- 1: the case that the BAB is replaced with the GB in the proposed method
- 2: the case of the proposed method itself

다는 것을 알 수 있다

IV. 실험 결과

제안된 회전 보정 방법의 성능을 평가하기 위하여 최대 800×600 크기의 영상을 얻을 수 있는 Necom사의 Nexicam 카메라를 400MHz의 Xscale 프로세서와 64MB의 RAM을 갖는 Compaq사의 PDA인 iPAQ h3590에 장착하여 다양한 주변 여건 하에서 여러 종류의 명함을 대상으로 시험 영상을 얻었다. 이때 사용한 명함의 종류로는 보통 명함을 포함하여 질감이 있는 용지를 갖는 명함과 사진이나 배경 그림이 있는 명함과 같은 특별한 명함이다. 그리고 주변 여건으로는 평범한 여건과 조명이 불균일한 경우, 명함 영상에 그림자가 있는 경우, 명함의 외부 영역에 나뭇결과 같은 배경이 있는 경우의 회전각 검출에 방해 요소가 있는 좋지 않은 여건이다. 회전 보정률은 전체 영상의 개수와 회전 보정에 성공한 영상의 개수의 비로 정의하였다. 회전 보정의 성공은 회전된 영상의 실제 측정된 회전각과 제안된 방법으로 계산된 각의 차이가 0.5도 이하인 경우로 하였다.

그림 10은 다양한 주변 여건 하에서 여러 가지 형태의 명함을 대상으로 얻는 640×480 크기의 명함 영상을 보여준다. 그림 10(a)와 10(b)의 영상들은 각각 보통 명함을 흰색 종이와 나뭇결의 책상위에 올려놓은 상태에서 얻은 영상들이다. 그림 10(c)와 10(d)의 영상들은 배경 그림이 있는 특별한 명함을 나뭇결의 책상위에 올려놓고 얻은 영상들이고,

그림 10(e)의 영상은 질감이 있는 용지를 갖는 특별한 명함을 나뭇결의 책상위에 올려놓고 얻은 영상이다. 그리고 그림 10(f)는 보통 명함을 2가지 색의 책상 위에 올려놓고 얻은 영상이다. 또한 그림 10(b)-(d)의 영상들은 불균일한 조명이나 그림자의 영향을 받은 영상들이다. 그림 11과 12는 각각 그림 10의 영상들에 대하여 전역 이진화에 의한 이진 영상을 입력 영상으로 하여 Avanindra의 방법^[4]으로 회전 보정한 결과 영상과 제안된 방법으로 회전 보정한 결과 영상들을 보여준다. 그림 11과 12에서 보면 제안된 방법은 Avanindra의 방법에 비하여 회전된 명함의 종류와 영상을 획득할 때의 주변 여건에 상관없이 회전 보정을 훨씬 더 잘 한다는 것을 알 수 있다.

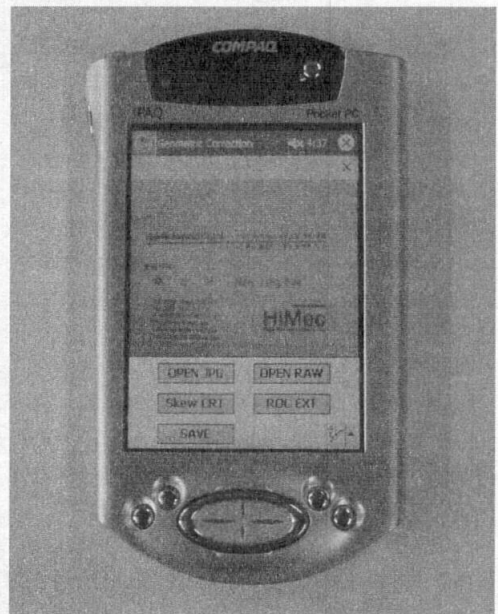
다음으로 제안된 방법의 회전 보정률 성능과 RMSE(root mean square error) 성능을 Avanindra의 방법^[4]의 성능들과 비교하였다. 서론에서 설명한 바와 같이 Avanindra의 방법은 해상도가 높은 스캐너를 통하여 얻은 문서 영상에 대한 회전 보정 방법이다. PDA 카메라를 통하여 얻은 명함 영상은 스캐너를 통하여 얻은 영상과는 달리 주변 여건의 영향을 받기 때문에 전역 이진화를 수행하면 그림 2의 이진 영상과 같이 일부 문자열이 나타나지 않을 수가 있다. 따라서 제안된 방법과 Avanindra의 방법과의 공정한 성능 비교를 위하여 Avanindra의 방법에 대한 입력 영상으로 전역 이진화와 블록적응 이진화에 의한 2가지 종류의 이진 영상을 사용하였다. 그리고 제안된 방법도 전역 이진화에 의한 경우와 블록적응 이진화에 의한 제안된 방법 그 자

표 2. Avanindra의 방법과 제안된 방법의 RMSE 성능
 Table 2. RMSE performances of Avanindra's method and the proposed method.

Type of business card	Surrounding condition	Avanindra		Proposed	
		with GB	with BAB	with GB	with BAB
Ordinary	Good	13.48°	12.69°	15.96°	0.32°
	Ill	13.64°	15.60°	14.04°	0.34°
Special	Good	14.75°	15.66°	12.84°	0.42°
	Ill	18.13°	15.74°	12.91°	0.60°
Total		15.11°	14.98°	13.99°	0.43°



(a)



(b)

그림 13. PDA에서 입력한 명함 영상과 회전 보정된 영상: (a) 입력 영상, (b) 회전 보정된 영상
 Fig. 13. An input business card image and its skew-corrected image on the PDA: (a) input image and (b) skew-corrected image.

체의 2가지 방법으로 회전 보정을 수행하였다. 표 1은 200개의 시험 명함 영상에 대한 제안된 방법과 Avanindra의 방법에 의한 회전 보정률 성능을 보여준다. 표 1에서 보면 제안된 방법 그 자체는 Avanindra의 방법에 비하여 약 51.5-64%의 회전 보정률 향상을 보인다는 것을 알 수 있다. 또한 제안된 방법 그 자체는 전역 이진화에 의한 제안된 방법에 비하여 약 30.5%의 회전 보정률 향상을 보인다는 것을 알 수 있다. 표 2는 200개의 시험 명

함 영상에 대한 제안된 방법과 Avanindra의 방법에 의한 회전 결과의 RMSE 성능을 보여준다. 표 2에서 보면 제안된 방법 그 자체는 Avanindra의 방법에 비하여 약 14.68° - 14.55°의 RMSE 향상을 보인다는 것을 알 수 있다. 또한 제안된 방법 그 자체는 전역 이진화에 의한 제안된 방법에 비하여 약 13.56°의 RMSE 향상을 보인다는 것을 알 수 있다.

마지막으로 제안된 방법을 Windows CE 환경에서 C++ 언어로 구현하여 시험 영상을 얻을 때 사

용했던 PDA에 탑재하여 명함 영상에 대한 회전 보정 실험을 수행하였다. 개인용 컴퓨터와는 달리 PDA는 프로세서에 FPU(floating-point unit)가 없어서 실수형 연산으로 알고리즘을 구현하면 수행 시간이 오래걸린다. 이러한 문제를 해결하기 위하여 모든 실수형 연산을 정수형으로 변환하였다. 실험 결과 PDA에서 수행한 결과는 개인용 컴퓨터에서 수행한 결과와 거의 동일하였고, 640×480 크기의 명함 영상 1장에 대하여 약 0.7초의 수행 시간을 보였다. 그림 13은 PDA에 명함 영상을 입력한 경우와 회전 보정이 완료된 경우를 보여준다.

V. 결 론

본 논문에서는 PDA 카메라로 획득한 명함 영상에 대한 회전 보정 방법을 제안하였다. 제안된 방법에서는 PDA 카메라로 획득한 명함 영상을 불균일한 조명이나 그림자 등의 영향을 줄이기 위하여 블록 분류에 의한 블록적응 이진화 방법을 사용하여 이진 영상을 얻었다. 블록적응 이진화로 얻어진 이진 영상에 대하여 서브샘플링과 형태학적 필터링을 이용하여 인접한 문자와 이들로 이루어진 문자열이 병합된 클러스터를 만들고, 그 중 모멘트를 이용한 클러스터 크기와 이심율을 이용하여 회전각 계산에 유용한 클러스터를 스트라이프로 선택하였다. 이때 선택된 스트라이프들의 방향각을 구하고 빈도수가 가장 큰 방향각들을 평균하여 최종 회전각을 구한 후 역매핑 및 쌍선형 보간법을 사용하여 입력 영상을 회전시켜 회전된 영상을 얻는다. 그런 다음 회전으로 인하여 생기는 모퉁이의 공백들을 가장 가까운 수평 화소의 값으로 채우는 모서리 채움을 수행하여 최종 회전 보정된 영상을 얻었다. 실험 결과 제안된 방법은 여러 가지 형태의 명함을 PDA 카메라로 다양한 주변 여건에서 획득한 시험 영상에 대하여 약 93%의 회전 보정률을 보여 명함의 종류와 주변 여건에 거의 상관없이 회전 보정을 잘 수행한다는 것을 알 수 있었다.

참 고 문 헌

[1] H. K. Kwag, S. H. Kim, S. H. Jeong, and G. S. Lee, "Efficient skew estimation and correction algorithm for document images," *Image and Vision Computing*, vol. 20, pp. 25-35, Jan. 2002.

[2] A. Dengel, "ANASTASIL: A system for low-level and high-level geometric analysis of printed documents," *Structured Document Image Analysis*, H. S. Baird et al. Eds., New York: Springer-Verlag, 1992.

[3] D. X. Le, G. Thoma, and H. Weschler, "Automated page orientation and skew angle detection for binary document images," *Pattern Recognition*, vol. 27, pp. 1325-1344, Oct. 1994.

[4] Avanindra and S. Chaudhuri, "Robust detection of skew in document images," *IEEE Trans. Image Processing*, vol. 6, pp. 344-349, Feb. 1997.

[5] H. Yan, "Skew correction of document images using interline cross-correlation," *CVGIP: Graphical Models and Image Processing*, vol. 55, pp. 538-543, Nov. 1993.

[6] G. Ciadilla et al., "An experimental system for office document handling and text recognition," in *Proc. Int. Conf. Pattern Recognition*, 1988, pp. 739-743.

[7] W. Postl, "Detection of linear oblique structures and skew scan in digitized documents," in *Proc. Int. Conf. Pattern Recognition*, 1986, pp. 687-689.

[8] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, pp. 62-66, Jan. 1979.

[9] 김종훈, 장익훈, 김남철, "블록 분류를 이용한 명함 영상에서의 블러링 판단," *대한전자공학회 2003년도 하계종합학술대회 논문집 IV*, 2003년 7월, pp. 1707-1710.

[10] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Upper Saddle River, NJ: Prentice-Hall, 2002.

[11] L. G. Shapiro and G. C. Stockman, *Computer Vision*, Upper Saddle River, NJ: Prentice-Hall, 2001.

[12] B. Jähne, H. Haußecker, and P. Geißler, *Handbook of Computer Vision and Applications*, San Diego, CA: Academic Press, 1999.

박 준 효(Jun-Hyo Park)

정회원



2001년 8월 : 경북대학교 공과
대학 전자공학과 (공학사)
2003년 8월 : 경북대학교 대학원
전자공학과 (공학석사)
2003년 8월~현재 : 삼성전자(주)
TN총괄 무선사업부 개발3그룹

S/W연구실 연구원

<주관심분야> 영상처리, 이동통신

장 익 훈(Ick-Hoon Jang)

정회원



1986년 2월 : 경북대학교 공과
대학 전자공학과 (공학사)
1988년 2월 : 경북대학교 대학원
전자공학과 (공학석사)
1998년 8월 : 경북대학교 대학원
전자공학과 (공학박사)

1988년 2월~1994년 2월 : 국방과학연구소 연구원

1998년 3월~현재 경운대학교 디지털전자정보공학부
조교수

<주관심분야> 영상처리, 멀티미디어, 컴퓨터 비전

김 남 철(Nam-Chul Kim)

정회원



1978년 2월 : 서울대학교 공과 대
학 전지공학과 (공학사)
1980년 2월 : 한국과학기술원
전기 및 전자공학과 (공학석사)
1984년 2월 : 한국과학기술원
전기 및 전자공학과 (공학박사)

1984년 3월~현재 : 경북대학교 전자전기컴퓨터학부
교수

1990년 1월~1994년 12월 : 생산기술연구원 HDTV
신호처리분과위원

1994년 1월~1996년 12월 : 한국통신학회 편집위원

1994년 1월~1996년 12월 : 대한전자공학회 편집위원

1996년 1월~현재 : 한국방송공학회 편집위원, 이사

1997년 1월~1998년 12월 : 한국통신학회 신호처리
연구회 위원장

<주관심분야> 영상처리, 영상압축, 영상복원, 컴퓨터
비전