

PDA에서의 명함 영상의 기울기 보정

박 준 효, *장 익 훈, 김 남 철
경북대학교 전자공학과, *경운대학교 전자공학과
전화 : 053-950-5530 / 핸드폰 : 011-9355-2858

Skew Correction of Business Card Images for PDA Application

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

Department of Electronic Engineering, Kyungpook National Universty

*Department of Electronic Engineering, Kyungwoon University

E-mail: withlove74@vcl.knu.ac.kr

Abstract

We present an efficient algorithm for skew correction of business card images obtained by a PDA 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 effects of irregular illumination and shadows over the input image. In the stripe generation, character string clusters are generated merging character strings and their inter-spaces, 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 correction rates of 97% for business card images.

I. 서 론

현대인은 일상생활에서 문서 형식을 갖춘 매체, 신문, 잡지, 명함 등을 접하고 있다. 그 중 명함은 흔히 접할 수 있는 문서의 한 형태이다. 명함은 중요한 영업 수단 및 홍보 수단으로 사용될 뿐 아니라 사람을 만났을 때 인사와 함께 명함을 주고받는 것은 일상적이 되었다. 명함 사용이 일상적이 됨에 따라 많은 양의 명함을 받게 되고 이들을 수첩이나 메모를 통해 관리하는 것은 보관 및 검색 등의 측면에서 볼 때 상당히 비효율적이다. 이동성 및 휴대성이

중시되는 현대 사회에서 PDA 같은 휴대 단말기에서 부속 카메라를 통하여 디지털화하여 저장하거나 디지털화 한 후 인식을 거쳐 저장 및 관리하는 것이 효율적이고 편리하다. 그러나 이와같은 PDA 카메라를 통하여 디지털화하는 과정에서는 손의 흔들림이나 촬영 오류로 인하여 전자문서 내에서 기울어짐이 발생된다. 일반적으로 문서 영상의 기울어짐이 2~3도인 경우는 문서 분석 및 문자 인식의 정확도가 떨어지며, 5도 이상이 되면 문자 인식의 신뢰성이 급격히 떨어지는 것으로 알려져 있다[1]. 그러므로 문서 영상의 분석이나 문자 인식 이전에 문서 영상의 기울기를 검출하고 보정하는 과정은 필수적이다.

문서 영상에 대한 기울기 보정에 관한 많은 연구가 이루어져 왔다[2-6]. Dengel[2]은 왼쪽 가장자리 탐색(left margin search)을 이용하여 문서 영상의 기울기를 검출하는 방법을 제안하였다. 그리고, D. S. Le[3] 등은 Hough 변환을 이용한 방법, Avanindra[4] 등은 상관도(correlation)를 이용한 기울기 검출 방법을 제안하였다. 그 밖에도 투영(projection)을 이용하는 방법[5], 푸리에 변환(Fourier transform)을 이용한 방법[6] 등이 있다.

이상의 연구들은 문자열이 문서 전 영역에 걸쳐 많이 존재하는 문서 영상에 대한 연구이고 또한 스캐너로 얻은 영상에 대한 연구이기 때문에 최소 300dpi 이상의 고화질의 영상을 얻을 수 있어 문자열이 선명하게 나올 뿐 아니라 외부의 환경에 의한 영향이 없는 깨끗한 화질의 영상에 대한 것들이다. 그러나 명함 영상은 문서 영상에 비해 문자열의 수가 확연히 적을뿐더러 문자열의 위치가 일정하지 않고, 명함에 따라 배경 무늬나 사진, 용지의 질감 등 기울기 보정에 있어 여러 가변 요소도 포함하고 있다. 더욱이 PDA 카메라로 획득한 명함 영상의 경우, 보통 최대 640×480의 저해상도이기 때문에 문자열의 선명도가 스캐너로 얻은 영상에 비해 떨어질 뿐만 아니라 명함 획득시 조명, 그림자 등 외부환경에 영향을 받아 입력 영상의 화질이 저하된다. 따라서 PDA 카메라를 통해 얻은 기울어진

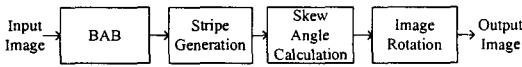
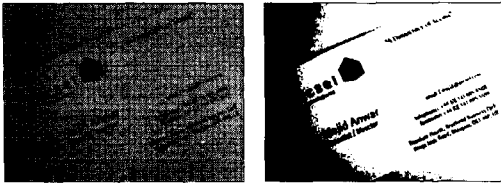


그림 1. 제안한 기울기 보정 방법의 블록도



(a) (b)

그림 2. 원 영상과 전역 이진화에 의한 이진 영상

명함 영상에 대해서는 기존의 방법을 적용하여 좋은 결과를 기대하기는 어렵다.

본 논문에서는 열악한 조건의 문서 영상에 대한 기울기 보정이라 할 수 있는 PDA 카메라를 사용하여 다양한 환경에서 얻은 명함 영상에 대하여도 강인하게 적용될 수 있는 기울기 보정 알고리즘을 제안한다.

II. 제안한 명함 영상의 기울기 보정

일반적으로 명함 영상에는 주소, 전화번호, email 주소 등의 문자열이 포함되어 있고, 문자열들은 명함의 가로축과 평행하다. 제안하는 기울기 보정 방법은 이러한 사실을 바탕으로 문자열의 각을 명함의 기울기로 간주하고 문자열의 기울기를 통하여 명함 영상의 기울기를 구한다. 제안한 기울기 보정 방법의 구성은 그림 1과 같다.

2.1 블록 적응 이진화(BAB)

제안한 기울기 보정 방법에서는 이진 영상에서 문자열들의 각을 구하기 때문에, 명함의 기울기를 구하기 위해서는 먼저 이진화 단계가 필요하다. 그림 2에서 보는 바와 같이 카메라로부터 입력받은 명함 영상을 Otsu의 방법[7]으로 역치를 정하여 전역 이진화(global binarization: GB)할 경우 불균일한 조명과 그림자 등에 의하여 문자열이 그림자에 의하여 나타나지 않을 수 있다. 이러한 문제에 대한 해결책으로 조명과 그림자 및 배경에 대해서도 문자열을 잘 추출할 수 있는 블록 적응 이진화 방법을 제안하였다. 그림 3은 블록 적응 이진화 절차를 보여준다.

먼저 입력된 영상을 8×8 크기의 블록들로 나누고, 각 블록을 문자 블록(character block: CB)과 배경 블록(background block: BB)으로 분류한다. 블록 분류에서는 문자 블록이 배경 블록보다 신호의 활동도가 큰 특성을 이용하여, 에너지가 큰 블록은 문자 블록으로 분류하고 에너지가 작은 블록은 배경블록으로 분류한다. 본 논문에서는 저주파 영역의 DCT 계수의 절대 합을 블록의 에너지로 정하였다. k번째 블록의 에너지 E^k 는 다음과 같이 구한다.

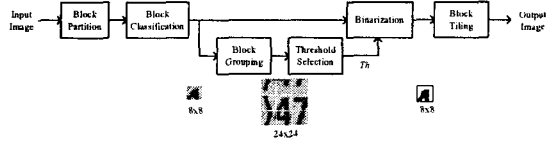


그림 3. 블록 적응 이진화 절차



(a) (b)

그림 4. 블록 분류 영상 및 BAB에 의한 이진 영상

$$E^k = \sum_i \sum_j |D_{ij}^k| \quad (1)$$

여기서 D_{ij}^k 는 k번째 블록에서의 $0 < i+j \leq 3$ 을 만족하는 (i, j) 위치에서의 DCT 계수를 나타낸다. k번째 블록은 다음과 같이 분류된다.

$$\begin{aligned} \text{If } E^k \geq Th_B \quad \text{then } k \in \text{CB} \\ \text{else } k \in \text{BB} \end{aligned} \quad (2)$$

여기서 Th_B 는 문자 블록과 배경 블록을 분류하는 역치로써 본 논문에서는 영상 전체에 대한 E^k 값의 평균을 사용하였다. 그림 4(a)에서 보면 문자 부분이 잘 분류되었음을 알 수 있다. 분류된 각 문자 블록은 인접한 8개의 블록들과 그룹핑하여 24×24의 크기의 블록을 생성하고 이 블록에 대하여 Otsu의 역치 결정 방법으로 이진화를 위한 역치 Th^k 를 찾는다. 그다음 Th^k 를 기준으로 원 8×8 크기의 블록을 다음과 같이 이진화 한다.

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

여기서 α 와 β 는 각각 배경 및 문자 화소에 할당되는 값을 의미한다. 그 후 이진화된 각 8×8 블록들을 원 위치에 붙이는 블록 붙이기(block tiling)을 하여 이진화된 영상을 얻는다.

그림 4(b)는 그림 2(a)의 영상을 BAB을 사용하여 이진화한 결과를 보여준다. 그림 4(b)를 그림 2(b)와 비교해 볼 때 BAB에 의하여 이진화된 결과 영상은 문자열이 전역 이진화에 의한 이진화의 결과보다 훨씬 더 잘 드러나 있음을 알 수 있다.

2.2 스트라이프 생성(stripe generation)

문자열에 대한 방향각을 구하기 위하여 문자열을 스트라이프라 정의되는 하나의 긴 모양의 블록으로 만든다. 먼저 스트라이프가 잘 생성되게 하기 위하여 입력 영상에

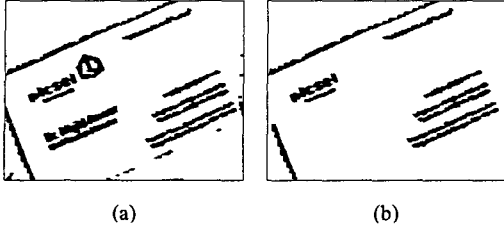


그림 5. 클러스터 생성 및 스트라이프 생성 영상

대하여 수평 방향 서브샘플링을 수행한다. 그 후 문자열 내의 문자를 접하게 하여 클러스터가 생성되도록 하기 위하여 형태학적 팽터링인 확장 연산(dilation operation)을 수행한다[8]. 확장 연산 후 상하가 접하게 된 문자열을 분리하기 위하여 수축 연산(erosion operation)을 수행하고 영상의 가로 대 세로 비를 원 영상과 같게 하기 위하여 수직 방향 서브샘플링을 수행한다. 그림 5(a)는 클러스터가 생성된 영상을 보여준다. 그림 5(a)에서 보면 대부분의 문자열들은 길쭉한 모양의 클러스터로 변하지만 마크와 잡음 블록 등과 같은 클러스터도 생기는 것을 볼 수 있다. 따라서 이러한 클러스터들 중에서 회전각 계산에 유용한 스트라이프를 구분해 내는 단계가 필요하다. 대개 스트라이프들은 다른 클러스터들과 비교해 볼 때 일정 범위의 크기와 긴 모양을 가지게 된다. 이 특징을 이용한 스트라이프 구분을 위하여 중심모멘트를 이용한 클러스터들의 크기와 클러스터들의 긴 정도를 나타내는 기준인 이심률(eccentricity)을 이용한다. 중심 모멘트[8]를 이용한 크기 μ_{00} 와 이심률 e 는 아래의 식과 같이 정의된다.

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

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

여기서 \bar{x} 및 \bar{y} 는 물체의 수평 및 수직방향의 중심이다. (4)식과 (5)식에 의하여 각각 구해지는 클러스터 크기 μ_{00} 및 이심률 e 를 실험적 통계에 의하여 구해진 e_{TH} 및 μ_L, μ_H 와 비교하여 클러스터들 중에서 $\mu_L \leq \mu_{00} \leq \mu_H$ 및 $e \geq e_{TH}$ 인 경우에 해당하는 클러스터를 스트라이프로 분류한다. 그림 5(b)는 그림 5(a) 영상에 대한 스트라이프 분류 결과 영상이다.

2.3 회전각 계산(skew angle calculation)

회전각 계산부에서는 스트라이프들의 기울기를 계산한다. 각 스트라이프의 각은 중심 모멘트를 이용하며 다음과 같이 구한다.

$$\theta = \frac{1}{2} \arctan \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right), 0 \leq \theta < 180 \quad (6)$$

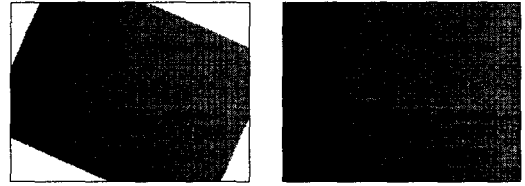


그림 6. 회전된 영상 및 모서리 채움 결과 영상

다음에는 θ 를 반올림하여 정수화 한 후 N 개의 스트라이프에 대해 히스토그램 $h(l), l=0, 1, \dots, 179$ 을 구한다. 그리고 $h(l)$ 을 이용하여 최종 방향각의 예상각 θ_r 을 다음과 같이 찾는다.

$$\theta_r = \arg \max_{0 \leq l < 180} \left\{ \frac{1}{3} \sum_{z=l-1}^{l+1} h(z \bmod 180) \right\} \quad (7)$$

예상각 θ_r 이 주어지면 최종 회전각 θ_c 는 $\theta_r - 0.5 \leq \theta < \theta_r + 0.5$ 범위내의 방향각 θ 를 평균하여 구한다.

2.4 영상 회전(image rotation)

영상 회전부에서는 기울어진 원 영상을 $-\theta_c$ 만큼 회전시킨다. 영상은 다음과 같이 원 영상의 좌표를 회전된 영상의 좌표로 매핑[9]함으로써 회전될 수 있다. 그러나 이러한 방법에 의하여 영상을 회전하면 원 영상과 화소들이 1:1로 매핑이 되지 않아서 회전된 영상내에서 홀이 발생할 수 있다. 따라서 본 연구에서는 다음과 같이 좌표 (x, y) 의 회전 보정된 결과 영상의 각 화소의 값들을 기울어진 원 영상의 좌표 (x, y) 의 화소의 값들로 역매핑[10]한다.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \quad (8)$$

이때 회전 보정된 결과 영상에 대응하는 원 영상의 화소의 위치가 삼각함수 연산으로 인하여 정수로 존재하지 않을 수가 있다. 이를 해결하기 위하여 실수형의 원 영상의 위치가 주어지면 가장 가까운 네 개의 화소의 값을 이용하여 쌍선형 보간법(bilinear interpolation)[8]을 이용하여 구한다. 그림 6(a)는 역매핑 회전 및 쌍선형 보간법을 이용하여 그림 2(a)의 영상을 회전한 결과 영상이다. 그림 6(a)에서 보는 바와 같이 회전 변환을 하게 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 6(b)는 그림 6(a) 영상에 모서리 채움을 수행한 영상이다.

III. 실험 결과

제안한 방법의 성능 평가를 위해 Nivicom사의 Nexicam 카메라를 Compaq사의 PDA인 iPAQ h3590에 장

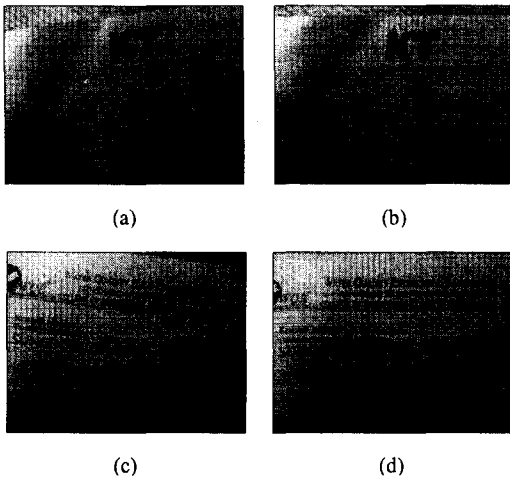


그림 7. 다양한 명함 영상과 회전보정 결과. (a) 나뭇결 바탕 위의 배경이 있는 명함 영상, (b) (a)에 대한 회전 보정 결과, (c) 불균일한 조명이 있는 명함 영상, (d) (c)에 대한 회전 보정 결과.

착하여 다양한 환경에서 다양한 명함을 대상으로 시험 영상을 얻었다. 명함의 종류로는 일반 명함을 포함해 사진이나 배경 그림이 있는 명함, 질감이 있는 용지의 명함 등을 사용하였고, 취득 환경으로는 평범한 환경과 불균일한 조명으로 인하여 명함 영상에 그림자가 있는 환경, 명함의 외부에 나뭇결 등의 기물이 검출에 방해 요소가 있는 경우 등의 좋지 않은 환경으로 나누어 영상을 획득하였다. 회전 보정의 성공 여부는 시각적으로 기울임을 느낄 수 없는 정도를 기준으로 하였고 보정 성공률은 전체 영상의 개수와 회전 보정에 성공한 영상의 개수의 비로 정의하였다.

그림 7은 다양한 명함 영상과 회전 보정 결과를 보여준다. 그림 7(a)와 그림 7(b)는 나무결의 배경이 있는 명함 영상과 회전 보정 결과를 각각 보여주고, 그림 7(c)와 그림 7(d)는 불균일한 조명이 있는 명함 영상과 회전 보정 결과 영상을 보여준다. 표 1은 명함 영상의 종류와 취득 환경에 따른 회전 보정 성공률을 보여준다. 표 1에서 보던 제안한 방법은 전체적으로 약 97%의 보정 성공률을 보여준다는 것을 알 수 있다.

다음으로 제안한 방법의 성능을 Avanindra의 방법과 비교하였다. PDA 카메라를 통하여 얻은 명함 영상은 조명 등의 주위 환경에 영향을 받기 때문에 공정한 비교를 위하여 Avanindra 방법에 대한 입력 영상으로 GB와 BAB에 의한 두가지 이진 영상을 사용하였다. 표 2는 제안한 방법과 Avanindra 방법에 대한 성능 비교를 나타낸다. GB에 의한 이진 영상에 대한 Avanindra 방법은 배경이 없는 명함 영상과 배경이 있는 명함 영상에 대해서 25% 및 61%의 회전 보정 성공률을 보여주었다. 그리고 BAB에 의한 이진 영상에 대한 Avanindra 방법은 50% 및 32%의 회전 보정 성공률을 보인 것에 반해, 제안한 방법은 각 명함 영상에 대해 100% 및 93%의 회전 보정 성공률을 보였다. 이들 결과로부

표 1. 제안한 방법의 회전 보정 성공률

		Conditions	
		Normal	III
Types of business cards	Normal	100%(18/18)	100%(24/24)
	Abnormal	94%(16/17)	93%(13/14)

표 2. 제안한 방법과 Avanindra 방법에 대한 회전 보정에 대한 성능 비교

	Avanindra with GB	Avanindra with LAB	Proposed algorithm
No pattern in background	25%(2/8)	50%(4/8)	100%(8/8)
Patterns in background	61%(19/31)	32%(10/31)	94%(29/31)

터 명함 영상에 대해서 제안한 방법이 Avanindra 방법에 비해 훨씬 더 좋은 결과를 보임을 알 수 있다.

참고문헌

- [1] H. K. Kwag, S. H. Kim, S. 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," *Structrued Document Image Analysis*, H. S. Baird, et al., Eds. New York: Springer-Verlag, 1992.
- [3] D. S. Le, G. R. Thoma, and H. Wechsler, "Automated page orientation and skew angle detection for binary document images," *Pattern Recognition*, 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] G. Ciadilla et al., "An experimental system for office document handling and text recognition," in *Proc. 9th Int. Conf. Pattern Recognition*, 1988, pp. 739-743.
- [6] W. Postl, "Detection of linear oblique structures and skew scan in digitized documents," in *Proc. 8th Int. Conf. Pattern Recognition*, 1986, pp. 739-743.
- [7] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, pp. 62-66, Jan. 1979.
- [8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Upper Saddle River, NJ: Prentice-Hall, 2002.
- [9] L. G. Shapiro and G. C. Stockman, *Computer Vision*, Upper Saddle River, NJ: Prentice-Hall, 2001.
- [10] B. Jhne, H. Haubecker, and P. Geibler, *Handbook of Computer Vision and Applications*, San Diego, CA: Academic Press, 1999.