

공백행 추출에 의한 기울어진 문서 영상의 보정

°정 재영*, 김 문현**

동양대학교 컴퓨터공학부, 성균관대학교 전기전자컴퓨터공학부**

A Skew Correction for Document Images by the Extraction of Blank Lines

°Jae Young Jung*, Moon H. Kim**

School of Computer Eng., Dongyang Univ., School of EEC Eng., Sungkyunkwan Univ.**

요약

본 논문에서는 선형적으로 기울어진 문서 영상의 기울기를 검출하기 위한 단순하면서도 효과적인 알고리즘을 제안한다. 문서 내의 인접한 두 행 사이에는 일정한 두께의 공백 행이 존재하며, 그 공백 행의 기울기는 실제 문서의 기울어진 정도를 반영한다는 사실에 기인한다. 먼저, 간단한 모폴로지 연산을 이용하여 문자행 영역과 공백행 영역을 분리한 후, 이를 일정 간격으로 수직 샘플링하여 수직선 상에 있는 모든 공백행의 중심점(행간점)을 찾는다. 전체 영상으로부터 동일한 공백 행 상에 있는 임의의 두 행간점 간에 계산된 기울기들의 분포를 보면 실제 문서의 기울기에서 최대 값을 가진다. 제안한 알고리즘을 다양한 형태의 가로쓰기 문서(검출 가능한 최대 기울기 : $\pm 45^\circ$)에 적용하여 0.5° 의 오차범위 내에서 정확한 결과를 얻을 수 있음을 보인다.

1. 서론

최근 디지털 라이브러리의 구축 및 광학 문자 인식 등의 필요성이 증대함에 따라 문서 영상의 자동 해석 및 인식에 대한 연구가 활발히 진행되고 있다. 이러한 시스템의 주요 과정으로는 문서의 입력, 이진화, 영역 분할, 문자 인식, 철자 교정 등의 주요 단계를 거치게 된다. 이때 입력되는 문서가 기울어져 있는 경우, 나머지 전 과정에 영향을 미치게 되게 인식률을 떨어뜨리는 커다란 장애 요소가 된다. 따라서 기울어진 문서가 입력되는 경우 이를 복원하는 기술은 문서 영상의 자동 해석을 위해 필수적이다.

문서의 기울어짐을 보정하기 위한 연구는 크게 투영에 의한 방법[1~2], 허프 변환을 이용하는 방법[3], 퓨리에 변환을 이용한 방법[1], 클러스터링을 이용한 방법[4], correlation을 이용한 방법[5~6] 등으로 분류된다. 이들 방법의 대부분은 연결 문자 추출 및 영역 변환에 따른 많은 계산 시간을 요구한다거나[1~4], 텍스트 영역이 많은 문서에 대하여 적용이 가능하다[5~6]는 문제점을 안고 있다.

본 논문에서는 행간점 추출을 이용하여 빠르게 문서의 기울어짐을 보정할 수 있는 알고리즘을 제안한다. 2장에

서는 알고리즘의 과정을 설명하고, 3장에서는 다양한 형태의 문서 영상에 적용한 결과를 보인다.

2. 기울어짐 보정 알고리즘

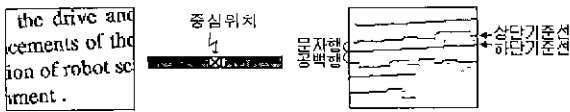
제안하는 기울어짐 보정 알고리즘의 전체적인 과정은 표 1과 같다.

표 1. 기울어짐 보정 알고리즘

입력 : 기울어진 이진 문서 영상
출력 : 기울어짐이 교정된 이진 문서 영상
1. Dilation에 의한 화소 widening
2. 전체 영상에 대해 수직 방향으로 샘플링을 수행
3. 각 수직선 상의 행간점 추출
4. 행간점간의 기울기 계산
5. 히스토그램 분석 => 최대치:기울어진 각도
6. 영상 회전

먼저, 입력되는 이진 문서의 각 행에서 인접한 자간의 공백 및 문자내의 공백을 메우기 위하여 그림 1의 (b)와 같은 $(2w+1) \times 1$ 크기의 structuring element를 이용하여 원영상(그림 1 (a))에 dilation 연산을 수행함으로써 문서 내의 검은 화소의 굵기를 증가시킨다 이 때 w 는 검은

화소를 수평 방향으로 팽창시키기 위한 굵기를 조절하는 변수로 너무 작을 경우 인접 문자간에 빈 공간이 계속 남을 수 있으며, 너무 클 경우 인접 행이 서로 합쳐져서 공백 행이 사라지게 된다. 따라서 문서의 해상도 및 문서의 기울어짐이 가능한 최대각도 등을 고려하여 설정하여야 한다. 실험에서는 w 를 5로 설정하였다. 이와 같이 검은 화소를 수평 방향으로 팽창시킴으로써 그림 1(c)와 같이 문자행과 공백행을 분리할 수 있다. Dilation 결과 영상에서 검은 영역의 상단부를 연결한 다각선을 상단 기준선, 하단부를 연결한 다각선을 하단 기준선이라 할 때, 인접한 상단 기준선과 하단 기준선 사이의 영역은 문서 상의 문자 행에 해당하고, 인접한 하단 기준선과 상단 기준선 사이의 영역은 공백 행에 해당한다.



(a) 원 영상 (b) structuring element (c) Dilation 결과
 그림 1. Dilation 연산에 의한 기준선 생성

전체 문서 영상에 대하여 일정 간격으로 수직 샘플링을 수행할 경우 각 수직선은 문자행과 공백행을 교차하면서 통과하게 되는데, 각 공백 행의 중심점을 추출하여 이를 행간점이라 한다. 이 때 공백 행의 높이에 대한 범위를 가정하여 이웃하지 않은 행간에 추출되는 행간점과 문자 행 내에서 추출되는 행간점 등을 삭제하므로써 전체 처리 시간을 줄이는 것이 가능하다. 공백 행의 높이를 가정할 수 없다 하더라도 불필요한 행간점들은 나중에 삭제되므로 기울기 추정 결과의 정확도에는 아무런 영향을 미치지 못한다.

일정 간격 떨어진 두 수직선상에서 행간점을 하나씩 추출하여 쌍을 만들고 이들이 다음의 두 조건을 만족하는 경우에 동일한 공백 행 상의 두 점으로 간주하여 기울기를 계산한다.

- 두 행간점을 잇는 직선은 검은 화소와 만나지 않는다.
- 두 행간점간의 기울기는 (세로쓰기 문서를 고려할 때) 검출 가능한 문서의 최대, 최소 기울기인 $\pm 45^\circ$ 의 범위 내에 있다.

이 때 두 수직선 간의 수평거리 d 에 따라 계산되는 기울기 θ 의 오차 범위는 식 (1)과 같이 결정된다

$$-\frac{1}{2} \tan^{-1} \left(\frac{1}{d} \right) \leq \theta_{error} \leq \frac{1}{2} \tan^{-1} \left(\frac{1}{d} \right) \quad (1)$$

문서 내에서 일정 수평거리를 가지는 모든 수직선 쌍에 대하여 앞서의 과정을 반복한다. 이는 일정 수평거리를 가지면서 동일한 공백 행 상에 있는 모든 행간점 쌍에 대해서 기울기를 계산하는 의미를 가지는데, 이들 중 최대 분포를 가지는 기울기가 실제 문서의 기울기이다.

계산된 기울기들의 분포를 보기 위하여 문서의 최대 기울기 범위인 $-45^\circ \sim 45^\circ$ 사이의 1차원적인 각도 선분을 $\frac{1}{2} \tan^{-1} \left(\frac{1}{d} \right)$ 의 간격으로 분할한 후 대응되는 위치 및 좌우의 셀에 누적시킨다. 좌우의 셀에까지도 누적시키는 이유는 계산된 기울기 θ 가 $\pm \frac{1}{2} \tan^{-1} \left(\frac{1}{d} \right)$ 범위의 오차 한계를 가지기 때문이다.

최종적으로 히스토그램을 분석하여 최대 분포를 가지는 셀을 선택하고, 그 셀 내에 있는 기울기 값들을 평균하여 문서의 기울어진 각도로 추정한다. 결과를 보기 위하여 식 (2)와 같이 역방향 사상에 의하여 영상을 회전시켰다. 이 때, (x, y) 는 원 영상내의 좌표이며, (x', y') 는 원 영상의 중심점 (C_x, C_y) 를 기준으로 점 (x, y) 를 각 α 만큼 회전한 좌표이다.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x' - C_x \\ y' - C_y \end{bmatrix} \quad (2)$$

3. 실험 결과

스캐너(모델: HP scanjet 5P)를 통하여 100dpi의 해상도로 문서를 입력받아 실험하였으며, 몇 가지의 결과를 수작업으로 계산된 실제 기울기와 비교하여 보인다. 문서는 $\pm 45^\circ$ 범위 내에서 기울어진 이전 영상으로 영한 혼용 문서(그림 2), 그래픽스, 사진, 도표 등을 큰 비율로 포함하는 문서(그림 3,4), 거래명세표 등과 같은 전표(그림 5), 편지 봉투(그림 6) 등을 대상으로 하였다.

그림 2(a)는 한글과 영문이 혼합된 2단 문서이며, (b)는 문서 (a)를 문서행과 공백행으로 분리하고 행간점을 추출한 후, 동일한 공백행 상의 행간점들을 연결한 결과이다. 잘못된 행간점을 잇는 선분이 존재하지만 전체적으로 문서의 실제 기울기와 유사한 선분들이 가장 많이 존재함을 알 수 있다. 실험에 사용된 대부분의 문서에서 0.5° 의 오차 한계 내에서 정확히 기울어짐을 보정하고 있다. 특히, 그림 4는 텍스트 영역이 부족하여 Avanindra[6]등이 기울기 검출에 실패하였으나, 제안한 방식에서는 기울기를 정확히 검출하고 있다.

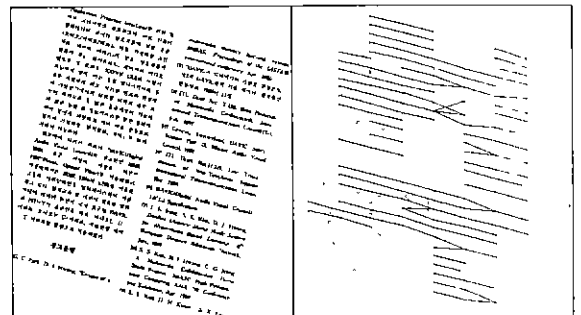


그림 2 한글과 영문이 섞인 문서(실제 14.75° , 계산 15.04°)

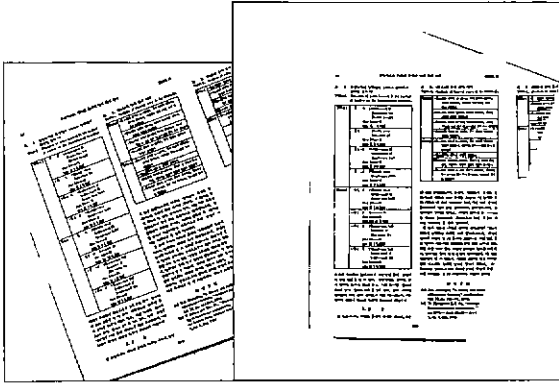


그림 3. 도표를 포함하는 문서(실제: -19.93° , 계산: -20.55°)

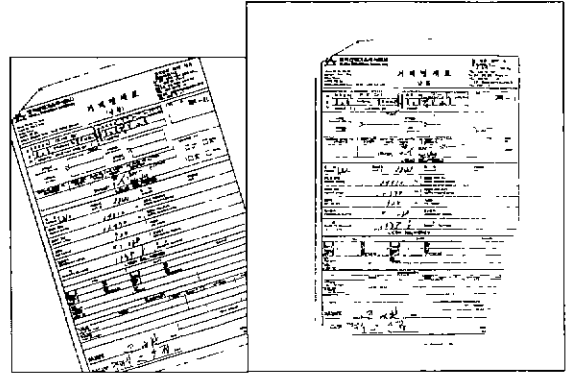


그림 5. 거래명세표 영상(실제: -16.98° , 계산: -17.35°)

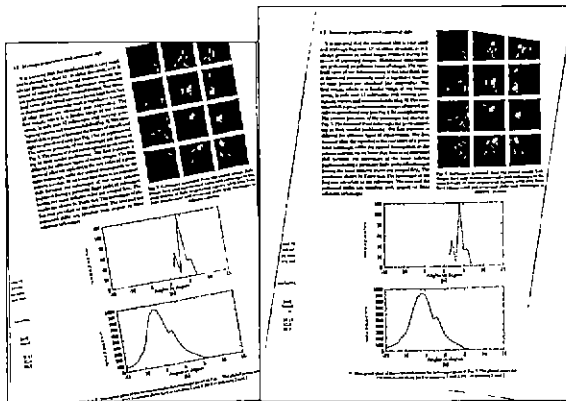


그림 4. 그림과 그래프를 포함하는 문서(실제: -9.23° , 계산: -9.23°)

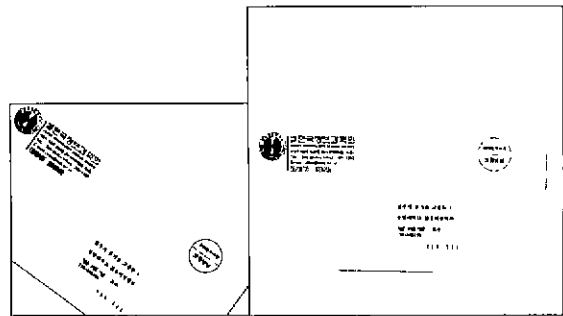


그림 6. 편지 봉투 영상(실제: 36.25° , 계산: 35.94°)

4. 결론

본 논문에서는 문서 영상의 해석 및 인식을 위해 입력되는 문서가 기울어져 있는 경우, 이를 보정할 수 있는 알고리즘을 제안하였다. 제안한 알고리즘을 다양한 형태의 문서, 다양한 내용물을 포함하는 문서에 적용하여 대부분의 문서에서 0.5° 의 오차 한계 내에서 정확한 결과를 얻을 수 있음을 보였다.

기존의 방법들과는 달리, 비텍스트 영역에 비해 텍스트 영역이 차지하는 비율이 많아야 한다면, 기울어짐의 정도가 작아야 한다면 하는 제약요구하지 않는다. 또한, 문서내의 연결화소를 찾는다면, 변환 함수를 적용한다면 하는 복잡한 과정을 요구하지 않으므로 처리 시간이 매우 빠르다.

참고문헌

- [1] W. Postl, "Detection of linear oblique structures and skew scan in digitized documents," *Proc. of the 6th Int. Conf. on Pattern Recognition*, pp. 1149-1152, 1986.
- [2] H. S. Baird, "The skew angle of printed documents," *Proc. of SPIE 40th Annual Conf. and Sym. on Hybrid Imaging Systems*, pp. 21-24, 1987.
- [3] S. Hinds, J. Fisher, and D. D'Amato, "A document skew detection method using run-length encoding and the Hough transform," *Proc. of the 6th Int. Conf. on Pattern Recognition*, pp. 1149-1152, 1990.
- [4] A. Hashizume, P. S. Yeh, and A. Rosenfeld, "A method of detecting the orientation of aligned components," *Pattern Recognition Letters*, vol. 4, pp. 125-132, 1986.
- [5] H. Yan, "Skew correction of document images using interline cross-correlation," *CVGIP Graphical Models and Image Processing*, vol. 55(6), pp. 538-543, Nov., 1993.
- [6] Avanindra and S. Chaudhuri, "Robust detection of skew in document images," *IEEE Trans. on Image Processing*, vol. 6, no. 22, pp. 344-349, Feb, 1997.