
에지를 이용한 문서영상의 기울기 보정

주재현* · 오정수**

Skew Correction of Document Images using Edge

Jae-hyon Ju* · Jeong-su Oh**

요 약

본 논문에서 선명 화질은 물론 저화질 문서영상에서 에지를 이용해 문서영상의 기울기를 검출하고, 보정하는 알고리즘을 제안한다. 제안된 알고리즘은 영상 복잡도를 이용해 문자영역을 선택하여 에지를 검출하고, 이들을 다양한 방향으로 투영하여 투영 히스토그램들을 생성한다. 그리고 히스토그램들에서 에지 집중도를 평가하여 문서영상의 기울기를 검출하고 기울어진 문서 영상을 보정한다. 고속 기울기 검출을 위해 부표본화와 3단 coarse-to-fine 탐색 알고리즘을 사용한다. 선명 화질과 저화질 영상의 기울기 검출에서 제안된 알고리즘의 최대 검출 오차와 평균 오차가 기존 유사 알고리즘의 50% 정도이고, 기울기 검출 시간은 25% 정도로 감소된다. 모바일기기로 취득된 밝기 불균일 영상에서는 기존 알고리즘은 적절한 2진 영상을 얻을 수 없어 기울기 검출이 어려우나 제안된 알고리즘의 평균 검출 오차 0.10 이하로 기울기를 검출한다.

ABSTRACT

This paper proposes an algorithm detecting the skew of the degraded as well as the clear document images using edge and correcting it. The proposed algorithm detects edges in a character region selected by image complexity and generates projection histograms by projecting them to various directions. And then it detects the document skew by estimating the edge concentrations in the histograms and corrects the skewed document image. For the fast skew detection, the proposed algorithm uses downsampling and 3 step coarse-to-fine searching. In the skew detection of the clear and the degraded images, the maximum and the average detection errors in the proposed algorithm are about 50% of one in a conventional similar algorithm and the processing time is reduced to about 25%. In the non-uniform luminance images acquired by a mobile device, the conventional algorithm can't detect skews since it can't get valid binary images, while the proposed algorithm detect them with the average detection error of 0.10 or under.

키워드

문서영상, 기울기보정, 에지, 저화질영상

Key word

document image, skew correction, edge, degraded image

* 준회원 : 부경대학교 이미지시스템공학과
** 종신회원 : 부경대학교 이미지시스템공학과(ojs@pknu.ac.kr)

접수일자 : 2012. 02. 06
심사완료일자 : 2012. 03. 12

I. 서 론

문서영상에서 문자를 인식하여 추출하는 문서인식은 정보화 사회에서 중요한 요소이고, 최근 스마트폰처럼 카메라를 포함한 모바일기기가 대중화되며 이를 통한 문서인식이 실생활에 광범위하게 활용되고 있다. 그러나 모바일기기는 작은 렌즈로 광량확보가 힘든 구조적 문제, 주변 사물에 의해 불균일한 조명 같은 촬영 환경적 문제, 손떨림 같은 촬영기술 문제들로 인해 빈번히 저화질 문서영상을 생성한다.

모바일기기를 통한 문서처리 시 가독성 향상이나 인식율 향상을 위해 기울기 보정은 필수적이다. 기존 기울기 보정 알고리즘은 주로 문자와 배경이 잘 구분되는 선명한 문서영상을 대상으로 연구되어 왔다. 따라서 모바일기기로 취득될 가능성이 있는 저화질 문서영상의 기울기 보정은 기존 알고리즘과 다른 접근이 필요하다. 기존 기울기 보정 알고리즘은 행간 여백의 교차 상관관계를 이용해 기울기를 보정하는 알고리즘[1,2,3], 주파수변환에 의한 푸리에공간에서 최대 밀도를 검출하여 기울기를 보정하는 알고리즘[4], 허프변환에 의한 변환공간에서 최대 분포점을 검출하여 기울기를 보정하는 알고리즘[5], 최근접 클러스터를 이용해 기울기를 보정하는 알고리즘[6], 투영 히스토그램을 이용해 기울기를 보정하는 알고리즘[7] 등이 있다. 이 알고리즘들은 문자주도적이고 선명한 영상을 대상으로 하고, 특히 주파수변환이나 허프변환을 이용한 알고리즘들은 많은 계산량이 요구되고 최근접 클러스터를 이용한 알고리즘은 *text line*을 가져야 하는 제한을 갖고 있다. 투영 히스토그램을 이용하는 *Ishitani* 알고리즘[7]은 선명한 문서영상에서 유효하고 문자영역과 비문자영역을 판단할 수 있는 장점이 있으나, 저화질 영상에서는 문자와 배경의 경계가 분명하지 않아 적절한 이진 영상을 얻을 수 없어 기울기 검출의 정확도가 떨어지고, 문자영역의 크기, 고려되는 문자영역의 수, 기울기 범위에 따라 매우 많은 계산량이 요구될 수 있다.

본 논문은 선명 화질은 물론 저화질 영상에 대응하기 위해 *Ishitani* 알고리즘의 장점을 취하면서 단점을 보완하는 알고리즘을 제안한다. 제안된 알고리즘은 가로세로 영상 복잡도를 이용해 문자영역을 선택하여 밝기변화를 볼 수 있는 에지를 검출하고 다양한 방향의 투영 히스토그램을 생성한다. 그리고 투영 히스토그램에서

에지 집중도를 평가하여 기울기를 검출하고, 문자영상이 갖는 특성을 이용해 검출된 기울기에 대해 신뢰성을 평가한다. 또한 기울기 검출의 고속 처리를 위해 부표본화와 3단 *coarse-to-fine* 탐색 알고리즘을 적용하고 있다. 다양한 영상을 이용한 실험 결과는 제안된 알고리즘이 선명 화질은 물론 저화질 영상에서 기울진 문서영상을 고속으로 잘 보정해 주고 있는 것을 확인시켜 주고 있다.

II. 문서영상의 특성

본 장에서는 효율적인 문서영상 기울기 보정 알고리즘을 개발하기 위해 활용될 수 있는 문서영상들의 특성들을 분석한다.

2.1. 고해상도 문서영상

최근 영상 취득 장치들은 고해상도를 지원하고, 다양한 해상도의 영상을 원하는 대로 선택할 수 있으나 대부분 최고 고해상도로 영상이 취득되고 있다. 이 영상은 문서영상의 기울기 검출을 위해 필요 이상으로 많은 화소를 갖고, 영상처리에 있어 계산 부담만 더한다. 따라서 취득된 영상의 해상도에 따라 적절한 부표본화(*down sampling*)를 수행하여도 문자가 갖고 있는 정보를 유지시키며 계산량을 크게 줄여준다. 그림 1에 2496×3196 해상도의 영상을 469×600 해상도 영상으로 부표본화를 해도 문자 정보를 유지시키는 것을 보여주고 있다. 원의 영상은 469×600 해상도의 일부 영상을 확대한 영상이다.

2.2. 높은 영상 복잡도

문자는 일반적으로 다수의 가로세로 획에 점이나 대각선 획이 첨가되어 구성된다. 영문 'w', 'e'의 경우 가로/세로 방향으로 최대 4/3개의 획을 갖고, 한글 '뽀', '홀'의 경우 가로/세로 방향으로 최대 6/8개의 획을 갖고 있다. 이들은 영상에서 획수의 두배 에지(*edge*)를 형성하여 영상 복잡도를 높여 문서영상에서 문자영역을 선택하는 중요한 정보가 된다. 본 논문에서 영상 복잡도(C)는 임의의 영역에 포함된 밝기 변화량으로 그라디언트(*gradient*, G)를 이용해 다음 식과 같이 표현한다.

$$G(x,y) = |f(x,y) - f(x,y+1)| + |f(x,y) - f(x+1,y)| \quad (1)$$

$$C(x_1 : x_2, y_1 : y_2) = \sum_{x=x_1}^{x_2} \sum_{y=y_1}^{y_2} G(x,y) \quad (2)$$

여기서 $f(x,y)$ 는 (x,y) 에서 화소 밝기이고, (x_1, y_1) 와 (x_2, y_2) 은 각각 영역의 시작과 끝 지점이다.

2.3. 문서영상의 기울기

문서영상은 일반적으로 문자들이 가로 방향으로 정렬되어 기록된 문서를 영상화한 것이다. 따라서 문서영상의 기울기는 문자행의 기울기와 같고, 모든 영역에서 같은 기울기를 갖는다. 그림 1에 문서영상의 기울기(θ_1)와 임의의 문자영역의 기울기(θ_2)를 보여주고 있고 이들은 서로 같다. 그래서 본 논문에서는 잘 정렬된 문자영역을 찾고 문자영역에서 기울기를 검출한다.

2.4. 문자의 가로 정렬

앞에서 언급한 것처럼 문자는 가로세로 방향으로 에지를 많이 발생시킨다. 그리고 문서 내의 문자들은 세로 방향에 대해 정렬되어 있지 않지만 가로 방향에 대해 정렬되어 있다. 그래서 문자영역에서 임의의 방향으로 밝

기변화를 나타내는 에지의 수를 합산할 때, 모든 방향에 대해 많은 에지를 발생시키지만 문자행과 수평인 방향에 대해 더 높은 집중도를 보인다.

따라서 기울어진 문서영상에서 에지 집중도가 가장 높은 방향이 문서영상의 기울기에 해당한다. 그림 1의 방사형 직선들은 임의의 방향에 따른 에지의 발생을 시각적으로 확인할 수 있다. 직선 l 을 제외한 모든 직선들에서 비슷한 에지를 발생시키지만 직선 l 에서 분명히 구분되는 에지 집중도를 보여주고 있고 이 직선의 기울기가 문서의 기울기이다.

2.5. 저화질 영상에서 문자 단위 에지

저화질 영상에서 나타나는 대표적인 현상이 블러(blur)이고, 이는 문자 획과 배경의 경계가 모호한 상태로 이상적인 2진 영상을 얻을 수 없게 한다. 따라서 문자의 에지도 문자 획이 아닌 문자 형태의 경계로 나타난다. 두 번째 문서영상 특성에 위배되어 기존 알고리즘의 성능을 저하시킨다. 그러나 문자의 수직 경계가 거의 문서 기울기에 맞추어 정렬되어 있어 기울기를 검출하는 중요한 자료이다.

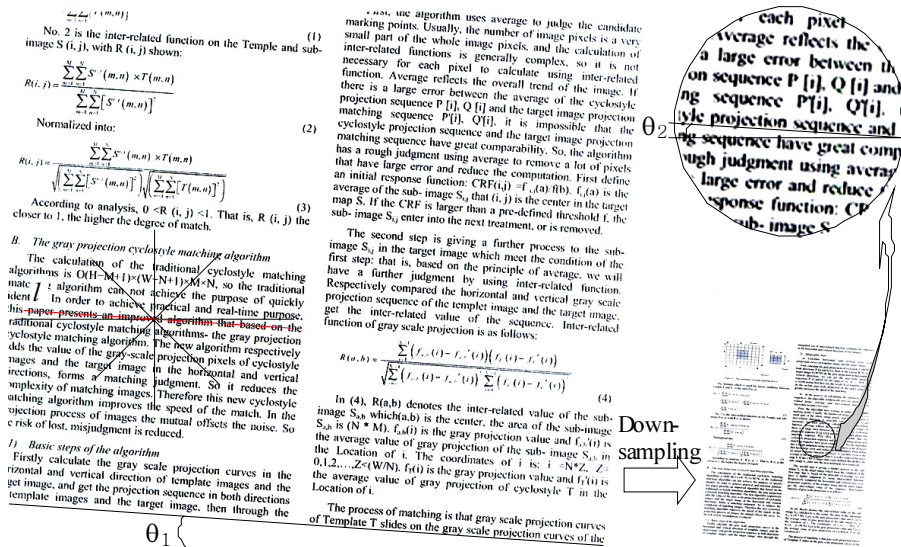


그림 1. 문서영상의 예
Fig. 1 Example for a document image

III. 에지를 이용한 문서영상 기울기 보정

본 논문에서 제안하는 문서영상의 기울기 보정 알고리즘은 그림 2과 같이 수행된다. 입력된 영상에서 문자영역을 선택하고, 문자영역에서 에지추출, 에지투영, 에지 집중도 평가를 통해 기울기를 검출하고, 기울기의 신뢰성을 평가하여 영상을 보정한다.

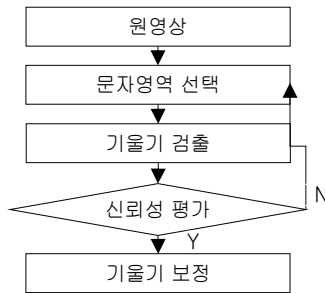


그림 2. 기울기 보정을 위한 블록도
Fig. 2 Block diagram for skew correction

3.1. 문자영역 선택

최근 고해상도의 문서영상에서 영상 전체를 대상으로 한 기울기 검출은 많은 계산 낭비를 가져온다. 본 논문에서는 문서영상의 특성을 고려하여 문서영상에 부표본화를 수행하고, 부표본화된 영상에서 문서영상의 기울기를 잘 표현할 수 있는 문자영역을 선택하여 전체 계산량을 줄인다. 실험적으로 A4 크기의 문서영상일 경우, 장축은 600화소가 적절하고, 확대 촬영 혹은 모바일 기기를 통해 근접 촬영하는 경우, 장축은 300~400화소가 적절하였다. 따라서 본 논문에서 문서영상의 부표본화는 스캔된 영상은 장축을 600 화소로 조정하고, 확대 촬영된 영상은 장축을 화질에 따라 300 혹은 400 화소로 조정한다. 저화질 영상을 저해상도로 부표본화할 경우 문자와 배경이 경계가 불분명하여 측정 오차가 다소 발생하여 저화질일수록 더 많은 화소를 요구한다.

문서영상의 기우러짐과 관계없이 문자영역은 복잡도가 높으므로 문서영상에서 일정 화소 간격으로 이동하면서 일정 영역의 영상 복잡도를 평가하여 최대 영상 복잡도를 갖는 영역을 문자영역으로 선택하였다. 후보 문자영역의 이동 간격과 크기는 각각 부표본화된 영상의 장축의 10%와 36%이다. 영상 복잡도 계산은 각 화소

에서 식 (1)의 그래디언트를 이용하여 식 (3)과 같은 누적 그래디언트 (A)를 계산한 후 식 (4)를 이용해 수행된다.

$$A(X, Y) = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} G(x, y) \quad (3)$$

$$C(x, y) = A(x+l_w, y+l_h) - A(x+l_w, y) - A(x, y+l_h) + A(x, y) \quad (4)$$

여기서 l_w 와 l_h 는 각각 문자영역의 폭과 높이이다. 누적 그래디언트는 영상 복잡도 계산에서 중복될 수 있는 계산을 크게 줄여준다.

3.2. 기울기 검출

문자영역에서 에지를 검출하고, 임의의 방향으로 투영시켜 에지 발생 집중도로 문서영상의 기울기를 결정한다. 단 문자영역은 정사각형이지만 에지 검출은 투영 방향에 따른 투영 에지 데이터의 일관성을 위해 정사각형의 내접원으로 제한하고 있다.

소벨 에지 검출기나 라플라시안 에지 검출기 같은 에지 검출기는 두터운 에지를 생성하며, 잡음에도 민감하다. 특히 저화질 영상에서 두터운 에지는 다방향성 에지로 표현되고, 불규칙적인 잡음 에지는 방향성이 없어 기울기 검출을 어렵게 한다. 따라서 본 논문에서는 방향성이 고려되고 가는 에지를 발생시키는 케니 에지 검출기를 사용한다. 이는 많은 계산량을 요구하나 에지 검출 영역이 문자영역의 내접원으로 제한되어 문제가 되지 않는다. 그림 3은 저화질 문서영상에서 검출된 에지영상을 보여주고 있다. 그림 3(b)는 소벨 에지 검출기로 에지를 구하고 레벨 24로 2진화한 영상으로 불규칙한 잡음 에지와 두터운 에지는 문자의 방향성에 혼란을 줄 것이다. 그림 3(c)는 소벨 에지 검출기로 에지를 구하고 레벨 84로 2진화한 영상으로 불규칙한 잡음 에지는 제거하였으나 방향성이 강한 라인 에지까지 제거하고 있다. 그림 3(d)는 케니 에지 검출기로 검출된 에지 영상으로 밝기 불균일과 상관없이 에지가 한 화소로 구성되고 방향성이 강한 라인 에지를 유지시켜 주고 있다.

검출된 에지는 투영되어 에지의 집중도가 계산된다. 에지 투영의 결과는 문자영역의 행수만큼의 빈 (bin, i)을 갖는 히스토그램이 생성되고 각 빈 (H_i)은 에지 화소의 발생 횟수이다. 투영은 $-45^\circ \sim +45^\circ$ 에서 다음과 같이 3 단계 coarse-to-fine 탐색 알고리즘을 사용한다.

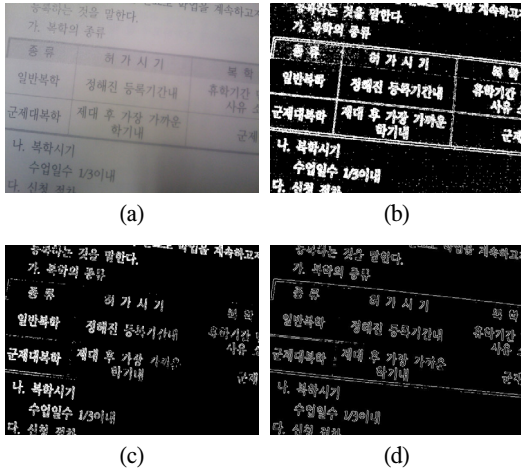


그림 3. 검출된 에지영상 (a) 문서영상 (b) 소벨 에지 24 (c) 소벨 에지 84 (d) 케니 에지
 Fig. 3 Detected edge images (a) Document image (b) Sobel edge 24 (c) Sobel edge 84 (d) Canny edge

단계1) $-45^{\circ} \sim +45^{\circ}$ 에서 3° 간격으로 최대 에지 집중도를 갖는 θ_{max1} 를 검출.
 단계2) θ_{max1} 을 중심으로 $-2^{\circ} \sim +2^{\circ}$ 에서 1° 간격으로 최대 에지 집중도를 갖는 θ_{max2} 를 검출.
 단계3) θ_{max2} 을 중심으로 $-0.5^{\circ} \sim +0.5^{\circ}$ 에서 0.1° 간격으로 최대 에지 집중도를 갖는 θ_{max3} 를 검출.

단계 1에서 투영각 변화를 크게 하면 국부 최대를 검출할 수 있어 본 논문에서는 선명한 문자영역에서 나타나는 에지 집중도의 분포특성을 고려해 3° 을 사용한다. 단계 3에서 최소 투영각 변화를 인식 가능한 0.1° 로 제한하였다. 그림 4는 투영각에 따른 에지투영과 투영 히스토그램을 보여주고 있다. 그림 4(a)는 문서 기울기와 투영 방향이 달라 모든 행이 투영 구간에 비례한 H_i 를 갖는 반면 그림 4(b)는 문서 기울기와 투영 방향이 일치하여 행에 따라 H_i 가 집중된 결과를 보여주고 있다. 한편 에지 집중도는 H_i 의 분산(V)으로 다음 식과 같이 계산한다.

$$V(\theta) = \frac{1}{n} \sum_{i=0}^n (A_i - M)^2, \quad M = \frac{1}{n} \sum_{i=0}^n A_i \quad (5)$$

여기서 n 은 빈의 개수로 문자영역의 라인수와 같다. 이상적인 문자영역에서는 문서영상의 기울기와 일치

하는 투영각을 중심으로 그림 5처럼 날카로운 산이 발생한다.

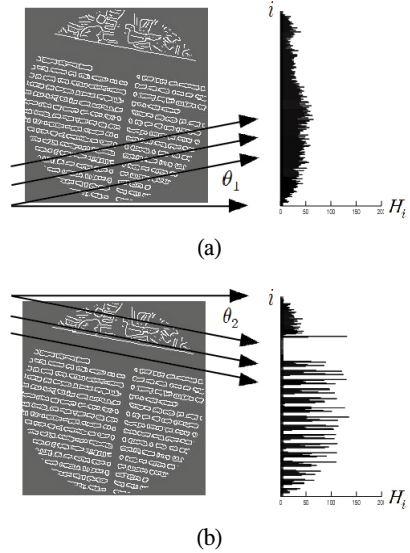


그림 4. 에지투영의 예
 (a) 투영각, θ_1 (b) 투영각, θ_2
 Fig. 4 Example for edge projection
 (a) Projection angle, θ_1 (b) Projection angle, θ_2

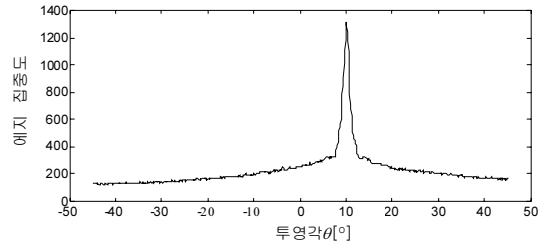


그림 5. 에지 집중도
 Fig. 5 Edge concentrativeness

3.3. 신뢰성 평가

일반적으로 문자영역이 정렬이 잘 된 문자들로 구성되어 있다면 투영각에 따른 에지 집중도 분포는 그림 5와 같이 다른 투영각과 분명히 구분되는 최대 집중도를 갖는 투영각을 중심으로 산을 형성한다. 그러나 복잡한 에지를 갖는 그림이나 그래프가 있는 영역이 문자영역으로 선택되면 에지들이 방향성 없이 불규칙적으로 발생하여 완만한 형태의 산을 형성하고, 최대 에지 집중도

를 갖는 투영각은 문서영상의 기울기라고 할 수 없다. 그래서 검출된 기울기에 대한 신뢰성 평가가 수행된다. 신뢰성 평가는 검출된 문서영상 기울기 θ_{max3} 의 최대 에지 집중도 $V(\theta_{max3})$ 와 단계 1에서 계산된 31개의 에지 집중도를 이용하여 다음 조건식을 만족하면 검출된 기울기는 문서영상 기울기로 신뢰된다.

$$V(\theta_{max3}) > n \times \overline{V(\sim[\theta_{max3} \pm 10^\circ])} \quad (6)$$

여기서 $\overline{V(\sim[\theta_{max3} \pm 10^\circ])}$ 는 $\theta_{max3} - 10^\circ \sim \theta_{max3} + 10^\circ$ 의 밖에 존재하는 에지 집중도들의 평균이고 n 은 임의의 상수로 실험을 통해 선명한 영상과 저화질 영상에서 각각 3.5와 4를 사용할 때 매우 안정적이다. 만약 검출된 기울기가 신뢰되지 못하면 현재의 문자영역을 제외하고 다시 문자영역 선택을 수행한다.

IV. 실험 결과

제안된 알고리즘을 두 가지 측면에서 성능을 평가하였다. 첫 번째는 선명한 영상과 저화질 영상에 대해 **Ishitani** 알고리즘과 제안된 알고리즘을 비교하였고, 두 번째는 스마트폰 카메라로 취득된 저화질 영상에서 제안된 알고리즘의 성능을 평가하였다. 실험은 **Inter core i5(2.8GHz) CPU, 2G RAM**이 장착된 PC에서 **Visual C++**로 수행되었다.

첫 번째 성능 평가를 위해 웹매거진인 **CJ 매거진 2010년 11월호 1~14페이지**에서 문자와 그림이 적절하게 섞여 있는 페이지를 **860×1177** 해상도로 취득하여 10장의 선명한 실험 영상을 만들었고, 이들을 **146×200** 해상도로 축소시켰다가 다시 원영상 크기로 확대해 저화질 실험 영상을 만들고 다시 이들을 $\pm 30^\circ, \pm 20^\circ, \pm 10^\circ$ 로 회전하여 알고리즘들에 각각 적용하였다. 실험에서 **Ishitani** 알고리즘은 대상 영상을 원영상 크기 그대로, 제안된 알고리즘은 대상 영상의 장축을 **600** 화소가 되게 축소하여 사용하였고, 탐색영역은 모두 장축의 **36%**를 이용하였고, 투영 각도는 $-45^\circ \sim 45^\circ$ 범위로 한정하여 **0.1^\circ**의 정밀도까지 탐색하였다. 또한 2진 영상을 대상으로 하는 **Ishitani** 알고리즘은 2진화가 선행되었다. 그림 6와 그림 7은 실험 영상의 예를 보여주고 있다.

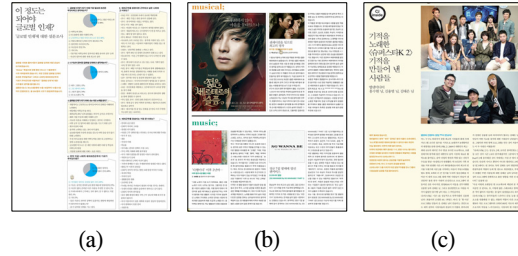


그림 6. 선명한 영상의 예
(a) 영상 1 (b) 영상 2 (c) 영상 3
Fig. 6 Example for clear images
(a) Image 1 (b) Image 2 (c) Image 3



그림 7. 저화질 영상의 예
Fig. 7 Example for a degraded image

회전된 문서영상들을 대상으로 한 실험 결과를 표 1에서 비교하고 있다. 기울기 검출에서나 계산 시간에서 제안된 알고리즘이 모두 좋은 성능을 보였다. 다소 크게 느껴지는 **0.4** 이상의 검출 오차가 **Ishitani** 알고리즘에서 선명한 영상과 저화질 영상에 각각 7개씩 보이는 반면 제안된 알고리즘에서는 전혀 없다. 이들은 주로 다단으로 구성된 문서영상에 발생하고 있고, 그림 8을 통해 원인을 예측할 수 있다. 그림 8(a)에 보이는 것처럼 문자행 내에서 수평 투영이나 대각선 투영은 유사한 결과를 가지므로 가로 변화만을 이용한 **Ishitani** 알고리즘은 근본적으로 다음과 같은 오차(δ)를 가질 수 있다.

$$\delta = \tan^{-1}\left(\frac{ch}{lw}\right) \quad (6)$$

여기서 ch 와 lw 는 각각 문자의 높이와 문자행 길이다. 특히 다단의 경우 lw 가 짧아 오차는 더 커진다. 반면 에지를 이용한 제안된 알고리즘은 그림 8(b)에서 보이는 것처럼 수직 에지가 활용되어 식 (6)의 오차가 발생하지 않는다.

표 1. 기울기 검출 오차와 시간
Table. 1 Skew detection error and processing time

영상	회전각	선명한 영상		저화질 영상	
		Ishitani	제안방법	Ishitani	제안방법
1	10°/-10°	0.6°/0.1°	0.2°/0.2°	0.1°/0.1°	0.3°/0.0°
	20°/-20°	0.5°/0.0°	0.0°/0.0°	0.6°/0.0°	0.3°/0.1°
	30°/-30°	0.5°/0.1°	0.0°/0.1°	0.4°/0.5°	0.1°/0.0°
2	10°/-10°	0.0°/0.1°	0.1°/0.0°	0.2°/0.2°	0.0°/0.0°
	20°/-20°	0.0°/0.0°	0.0°/0.0°	0.0°/0.1°	0.0°/0.1°
	30°/-30°	0.1°/0.1°	0.0°/0.1°	0.0°/0.2°	0.1°/0.0°
3	10°/-10°	0.0°/0.0°	0.1°/0.1°	0.7°/0.1°	0.1°/0.0°
	20°/-20°	0.1°/0.0°	0.1°/0.0°	0.1°/0.0°	0.0°/0.0°
	30°/-30°	0.5°/0.1°	0.0°/0.0°	0.6°/0.2°	0.0°/0.1°
4	10°/-10°	0.0°/0.1°	0.1°/0.0°	0.0°/0.1°	0.0°/0.1°
	20°/-20°	0.0°/0.0°	0.0°/0.1°	0.0°/0.1°	0.0°/0.1°
	30°/-30°	0.0°/0.1°	0.1°/0.0°	0.2°/0.1°	0.0°/0.0°
5	10°/-10°	0.0°/0.0°	0.0°/0.1°	0.3°/0.0°	0.0°/0.1°
	20°/-20°	0.0°/0.0°	0.0°/0.1°	0.0°/0.0°	0.1°/0.0°
	30°/-30°	0.2°/0.1°	0.0°/0.0°	0.2°/0.2°	0.0°/0.1°
6	10°/-10°	0.1°/0.1°	0.1°/0.1°	0.1°/0.2°	0.1°/0.2°
	20°/-20°	0.0°/0.1°	0.2°/0.1°	0.1°/0.1°	0.1°/0.1°
	30°/-30°	0.2°/0.0°	0.0°/0.1°	0.2°/0.2°	0.0°/0.0°
7	10°/-10°	0.4°/0.2°	0.1°/0.0°	0.1°/0.0°	0.1°/0.0°
	20°/-20°	0.4°/0.1°	0.0°/0.0°	0.1°/0.2°	0.2°/0.1°
	30°/-30°	0.1°/0.1°	0.0°/0.1°	0.3°/0.1°	0.0°/0.0°
8	10°/-10°	0.3°/0.2°	0.1°/0.2°	0.1°/0.4°	0.2°/0.1°
	20°/-20°	0.1°/0.4°	0.1°/0.1°	0.6°/0.5°	0.1°/0.2°
	30°/-30°	0.1°/0.1°	0.1°/0.2°	0.2°/0.5°	0.1°/0.2°
9	10°/-10°	0.0°/0.1°	0.0°/0.0°	0.0°/0.0°	0.1°/0.0°
	20°/-20°	0.0°/0.0°	0.1°/0.0°	0.1°/0.0°	0.1°/0.0°
	30°/-30°	0.1°/0.0°	0.1°/0.0°	0.1°/0.1°	0.1°/0.0°
10	10°/-10°	0.1°/0.1°	0.1°/0.0°	0.2°/0.1°	0.1°/0.3°
	20°/-20°	0.1°/0.0°	0.1°/0.0°	0.1°/0.2°	0.1°/0.1°
	30°/-30°	0.1°/0.1°	0.0°/0.0°	0.1°/0.2°	0.0°/0.0°
절대 평균오차		0.12°	0.06°	0.18°	0.07°
평균 계산시간		770ms	195ms	790ms	198ms

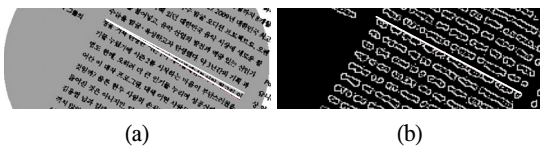


그림 8. 문자 크기의 영향 (a) Yasuto (b) 제안방법
Fig. 8 Effect of character size (a) Yasuto (b) Proposed

또한 저화질 영상에서는 그림 9에 보이는 것처럼 객체와 배경의 구분이 모호해서 문자를 구성하는 가로세로 획을 표현할 수 있는 이상적인 2진 영상을 얻을 수가 없다. 그 결과 Ishitani 알고리즘에서 검출 성능은 더 떨어지는 반면 제안된 알고리즘은 수직 에지가 유효하여 검출 성능이 거의 유지된다. 한편 계산 시간 측면에서 부표본화에 의한 영상처리 화소 수의 축소와 3단 coarse-to-fine 탐색은 제안된 알고리즘의 처리 시간을 Ishitani 알고리즘의 1/4 정도로 줄여 주었다.

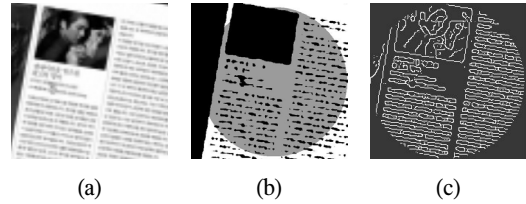


그림 9. 저화질 영상의 영향
(a) 원영상 (b) 2진 영상 (c) 에지 영상
Fig. 9 Effect of a degraded image
(a) Original image (b) Binary image (c) Edge image

두 번째 평가는 그림 10과 같이 문자영역과 그림, 차트같은 비문자영역이 존재하며 조명 불균형으로 그림자가 존재하는 스마트폰으로 취득된 1944×2592 해상도 영상들에서 수행되었다. 저화질 영상은 원영상을 243×324 해상도 영상으로 축소 후 원영상 해상도로 확대하여 생성되었다. 각 영상들은 ±30°, ±20°, ±10°로 회전하여 제안된 알고리즘에 적용되었다. 불균일한 밝기 분포로 적절한 2진 영상을 얻을 수 없는 Ishitani 알고리즘에는 적용할 수 없었다. 제안된 알고리즘은 대상영상의 장축을 400 화소로 축소하고 기울기 탐색영역은 기준축의 40%를 이용하였다. 탐색 각도는 -45°~45° 범위로 한정하였으며, 0.1°의 정밀도까지 탐색하였다. 실험 결과는 원영상과 저화질의 영상에서 평균 절대 오차는 각각 0.05°와 0.10°이고 계산 시간은 각각 98ms와 97ms로 보여주고 있다. 제안된 알고리즘은 실 생활에서 취득되는 이상적이지 않은 영상에서도 높은 신뢰도를 보여주고 있다.

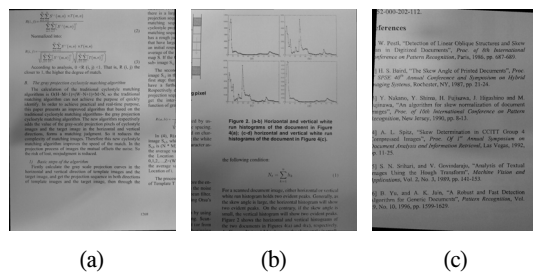


그림 10. 취득된 영상들 예
(a) 영상 1 (b) 영상 2 (c) 영상 3
Fig. 10 Example for acquired images
(a) Image 1 (b) Image 2 (c) Image 3

V. 결 론

선명 화질은 물론 저화질 문서영상을 대상으로 가로세로 밝기변화 정보를 갖는 에지를 이용한 고속 기울기 보정 알고리즘을 제안하였다. 영상 복잡도에 의해 선택된 문자영역에서 케니 에지 검출기를 이용해 가늘고 방향성 있는 에지를 생성하고, 투영 방향을 회전하며 에지의 투영 히스토그램들을 생성하고, 에지 집중도를 계산하여 문서영상의 기울기로 결정하였다. 또한 고속 알고리즘을 위해 부표본화를 수행하였고, 에지 집중도 평가에 3단 coarse-to-fine 고속 탐색 알고리즘을 적용하였다. 다양하게 기울어진 영상들에서 기울기 검출 실험이 수행되었다. 선명 화질 영상과 저화질 영상에서 기존 Ishitani 알고리즘은 각각 최대 검출 오차가 0.6°와 0.7°이고, 평균 검출 오차가 0.12°와 0.18°인 반면에 제안된 알고리즘은 최대 검출 오차가 0.2°와 0.3°이고, 평균 검출 오차가 0.06°와 0.07°이다. 직접 취득한 밝기가 불균일 영상에서는 Ishitani 알고리즘은 적절한 2진 영상을 얻을 수 없어 기울기 검출이 어려우나 제안된 알고리즘의 평균 검출 오차 0.1° 이하로 기울기를 검출하였다. 또한 제안된 고속 알고리즘의 기울기 검출 시간은 Ishitani 알고리즘과 비교해 1/4에 불과하다. 실험 결과를 통해 제안된 알고리즘은 일반 문서영상뿐만 아니라 저화질 문서영상에서 고속으로 높은 신뢰도의 문서 기울기를 검출하고 영상을 잘 보정하는 것을 보여주고 있다.

참고문헌

- [1] H. Yan, "Skew Correction of Document Images using Interline Cross Correlation," *CVGIP*, Vol. 55, No. 6, pp 538-543, 1993.
- [2] Avanindra and S. Chaudhuri, "Robust Detection of Skew in Document Images," *IEEE Trans. on Image Processing*, Vol. 6, No. 2, pp. 344-349, 1997.
- [3] B. Gatos, N. Papamarkos, and C. Chamzas, "Skew Detection and Text Line Position Determination in Digitized Documents," *Pattern Recognition*, Vol. 30, No. 9, pp. 1505-1519, 1997.
- [4] W. Postl, "Detection of Linear Oblique Structures and Skew Scan in Digitized Documents," *Proc. of 8th Int'l. Conf. on Pattern Recognition*, pp. 687-689, 1986.
- [5] A. Amin and S. Fischer, "A document skew detection method using the hough transform," *Pattern Anal. and Appl.*, vol. 3, no. 3, pp. 243 - 253, 2000.
- [6] Lu and C. L. Tan, "A nearest-neighbor chain based approach to skew estimation in document images," *Pattern Recognition Letter*, vol. 24, no. 14, pp. 2315 - 2323, 2003.
- [7] Y. Ishitani, "Document Skew Detection Based on Local Region Complexity," *ICDAR*, pp. 49-52, 1993.

저자소개



주재현(Jae-hyon Ju)

2010년 부경대학교 이미지시스템
공학 학사
2010년~현재 부경대학교 대학원
이미지시스템공학과
석사과정

※ 관심분야: 영상압축, 문서영상처리



오정수(Jeong-su Oh)

1992년 중앙대학교 대학원 전자
공학 석사
2001년 중앙대학교 첨단영상대학원
영상공학과 박사

2002년~현재 부경대학교 이미지시스템공학과
※ 관심분야: 영상압축, 영상검색, 문서영상처리