

교차점 검출과 분류를 통한 카메라 문서영상에서의 테이블 구조 인식 알고리즘

서원교* 구형일** 이동혁† 김상호† 조남익*

*서울대학교 전기컴퓨터 공학부, ** 아주대학교 전자공학과, † 삼성전자

cusisi@spl.snu.ac.kr, hikoo@ajou.ac.kr, odyssey.lee@samsung.com, kimsangho@samsung.com, nicho@snu.ac.kr

Table recognition algorithm for camera-captured document images based on junction detection and labeling

*Seo, Won Kyo **Koo, Hyung Il † Lee, DongHyuk † Kim, Sang Ho *Cho, Nam Ik

*Department of Electrical Engineering and Computer Science, Seoul National University

** Division of Electrical and Computer Engineering, Ajou University † Samsung Electronics

요약

표는 중요한 정보를 함축적으로 담고 있는 문서 요소로서 문서 영상에서 표의 내용과 구조를 분석하고 이해하려는 연구가 많이 진행되어 왔다. 이러한 표의 검출과 인식에 관한 기존의 연구들은 평판 스캐너로 취득한 문서 영상을 대상으로 이루어졌는데 최근에는 디지털 카메라와 스마트폰이 보급됨에 따라 평판 스캐너 대신 카메라를 이용한 표 인식의 필요성이 대두되고 있다. 따라서 본 논문에서는 카메라로 획득한 문서 영상에서 표 인식에 대한 알고리즘을 제안한다. 먼저 표가 선들의 집합으로 이루어져 있다는 가정 아래 문서 이미지에 존재하는 선을 이진화와 강인한 곡선 맞춤 알고리즘을 사용하여 검출한다. 검출된 선들의 교차점은 표의 요소일 수도 있으며 오검출의 결과일 수도 있는데 교차점 주변의 관찰 결과와 교차점 사이의 연관 관계를 에너지 식으로 표현하고 이 식을 최소화함으로써 각각의 교차점에 최적의 레이블을 할당한다. 얻어진 레이블은 표로 유일하게 변환되며 표의 구조를 셀 단위까지 추정할 수 있다. 다양한 표 영상에 대한 실험 결과를 통하여 제안한 방법이 문서영상의 기하학적인 왜곡에도 불구하고 영상에 존재하는 표를 성공적으로 인식함을 보여준다.

1. 서론

문서 영상 인식은 문서 영상의 문자를 판독함으로써 문서의 해독과 편집을 가능케 하는 작업이다. 최근에는 디지털 카메라와 스마트폰이 널리 보급되면서 일반 카메라로 획득한 문서 영상 인식과 관련한 많은 응용 프로그램들이 등장하고 있다. 이러한 카메라 문서 영상 인식은 스캐너와는 달리 보다 복잡한 전처리를 거친 후 일반적인 광학문자인식 방법을 통해 이루어지는데, 문서상에 등장하는 표는 판독하기에 어려움이 있다. 이를 해결하기 위해 표의 검출 및 구조 분석에 대한 연구가 많이 진행되어 왔지만 대부분 스캐너로 입력된 문서에서의 표를 대상으로 하고 있다. 따라서 카메라로 획득한 표 영상을 기존 알고리즘에 적용하기 위해서는 평활화 과정을 거쳐야 한다. 하지만 왜곡된 전체 영상을 평활화하는 것은 많은 계산량을 필요로 하는 일이다. 또한 많은 평활화 방법들이 충분한 텍스트 라인을 가정하고 있기 때문에 일반적인 표 영상에 적용할 수 없다는 단점이 있다. 마지막으로 대부분의 평활화 알고리즘은 OCR(Optical Character Recognition) 성능향상을 위해서 개발되었으며, 표를 구성하는 선들이 거의 이상적인 직선으로 나타나는 스캐너 영상을 기준으로 개발된 알고리즘을 적용할 수 있을지에 대해서는 추가적인 연구가 필요하다[1]. 따라서 본 논문에서는 일반 카메라로 획득한 기하학적인 왜곡과 조명 조건이 좋지 않은 문서영상에서 표를 직접적으로 인식하는 방법을 제안한다. 표의 구조를 파악함으로써 셀의 형태와 위치를 알아낼 수 있는데, 이 경우 문자 인식을 위

한 전처리 단계로 평활화를 수행하더라도 상대적으로 작은 영역이며 둘레를 알기 때문에 경계 보간법을 적용할 수 있다[2,3]. 보다 구체적인, 제안하는 알고리즘에서는 표가 선들의 집합으로 이루어져 있다는 가정 하에 문서상에 존재하는 선을 검출하고 이를 바탕으로 교차점을 구성한 후, 각 교차점에 레이블을 할당하는 방법으로 표의 구조를 파악한다. 이를 위하여 이진화와 세션화 알고리즘을 적용하여 리지(ridge) 맵을 얻고 가로와 세로 방향 각각에 대해 곡선 맞춤을 수행함으로써 선을 검출한다. 마지막으로 이러한 선들의 교차점에 적절한 레이블을 할당할 수 있도록 하는 에너지 식을 설계하고, 이 식을 최소화함으로써 표의 구조를 파악한다.

본 논문의 구성은 다음과 같다. 2절에서는 제안하는 표 구조 인식 방법에 대해 설명하고 3절에서는 제안하는 방법을 통해 얻은 실험결과를 분석한다. 마지막으로 4장의 결론으로 본 논문을 마무리 한다.

2. 제안하는 알고리즘

본 절에서는 일반 카메라로 취득한 표 영상에서 가로 방향과 세로 방향 선들의 교차점에 적절한 레이블을 할당할 수 있는 에너지 식을 세우고 이 식을 최소화함으로써 표의 구조를 파악하는 방법을 제안한다.

2.1 교차점 찾기

레이넥스 이론을 이용한 영상 이진화. 레이넥스 이론에서는 영상의 크기 $I(x,y)$ 를 반사율(reflectance) $R(x,y)$ 와 조도(illumination) $L(x,y)$ 의 곱으로 표현될 수 있다고 가정한다. 이를 정리하면

$$R(x,y) = \frac{I(x,y)}{L(x,y)} \quad (1)$$

이므로 실제 조도 $L(x,y)$ 를 추정할 수 있다면 반사율 추정 값을 얻을 수 있다.

레이넥스 이론을 이용한 이진화 알고리즘에서는 조도 성분이 전체적으로 느리게 변화한다는 가정 아래 적당히 큰 가우시안 커널을 원영상에 필터링 시킴으로써 $L(x,y)$ 를 추정한다[4]. 적당한 문턱 값을 선택한다면 결과적으로 고른 이진화 영상을 얻을 수 있다. (그림 1(b))

연결요소 필터링과 세선화. 레이넥스 이론을 이용하여 얻은 이진화 영상에서 표의 선 성분이 아닌 글자들이나 잡음의 존재는 가로 방향과 세로 방향의 선을 추출하는데 있어서 방해가 된다. 따라서 일차적으로 선이 아닐 것이라고 판단되는 객체를 제거해야 할 필요가 있는데 이를 위해 앞서 구한 이진화 영상에서 연결 요소를 추출하고 그 가운데 선이 아닌 성분을 필터링한다.

일반적으로 표의 둘레를 구성하는 선 성분은 가늘고 긴 형태를 갖기 때문에 이러한 특성을 바탕으로 선에 대응되는 성분만을 취하기 위해 먼저 각각의 연결 요소에 대해서 공분산(covariance matrix)을 구한다. 이 공분산을 아래와 같이 고유값 분해(eigenvalue decomposition)를 함으로써 각각의 연결 요소를 타원 형태로 모델링 할 수 있다.

$$\Sigma = \sigma_1 v_1 v_1^T + \sigma_2 v_2 v_2^T \quad (2)$$

여기서 σ_1, σ_2 는 고유값(eigenvalue) 이고, v_1, v_2 는 각각의 고유값에 해당하는 고유벡터(eigenvector)이다. $\sigma_1 > \sigma_2$ 일 때, $\sqrt{\sigma_1} \sqrt{\sigma_2}$ 는 각각 타원의 장축과 단축 길이에 해당하기 때문에 적당한 T_1 에 대해서 $\frac{\sqrt{\sigma_1}}{\sqrt{\sigma_2}} < T_1$ 에 해당하는 연결 요소를 걸러냄으로써 글자의 상당부분을 제거할 수 있다. 하지만 표의 경우 표의 선 성분 전체가 하나의 연결요소로 묶이게 되면 가로 혹은 세로 방향이 다른 쪽 방향에 비해 길다는 특성이 없어지기 때문에 그 역시 $\frac{\sqrt{\sigma_1}}{\sqrt{\sigma_2}} < T_1$ 의 조건에 해당할 수 있다. 이를 해결하기 위해 타원 안에 존재하는 픽셀의 밀도를 고려한다. 글자 하나하나의 경우 타원 안에 존재하는 픽셀의 밀도가 높지만 표 전체가 하나의 연결 요소로 묶일 경우 타원 안에 존재하는 에지 픽셀의 밀도가 낮다. 다시 말해 하나의 연결 요소를 구성하는 에지 픽셀의 개수를 N , 그 연결 요소에 해당하는 타원의 넓이를 M 이라 할 때, 적당한 문턱 값 T_2 에 대해서 $\frac{N}{M} > T_2$ 에 해당하는 연결요소는 글자 또는 잡음, 반대의 경우 표의 일부로 판단할 수 있다. 교차점을 찾기 위해서는 표를 구성하는 선에 대해서 가로 방향과 세로 방향으로 곡선 맞춤 알고리즘을 수행해야 한다. 하지만 연결 요소 필터링을 통해 남은 각각의 연결 요소들이 두꺼운 너비를 가질 가능성이 높기 때문에 정확한 곡선 맞춤 결과를 보장할 수 없다. 따라서 세선화(thinning)를 통해 각 연결 요소들의 너비를 얇게 만들어준다. (그림 1(c))

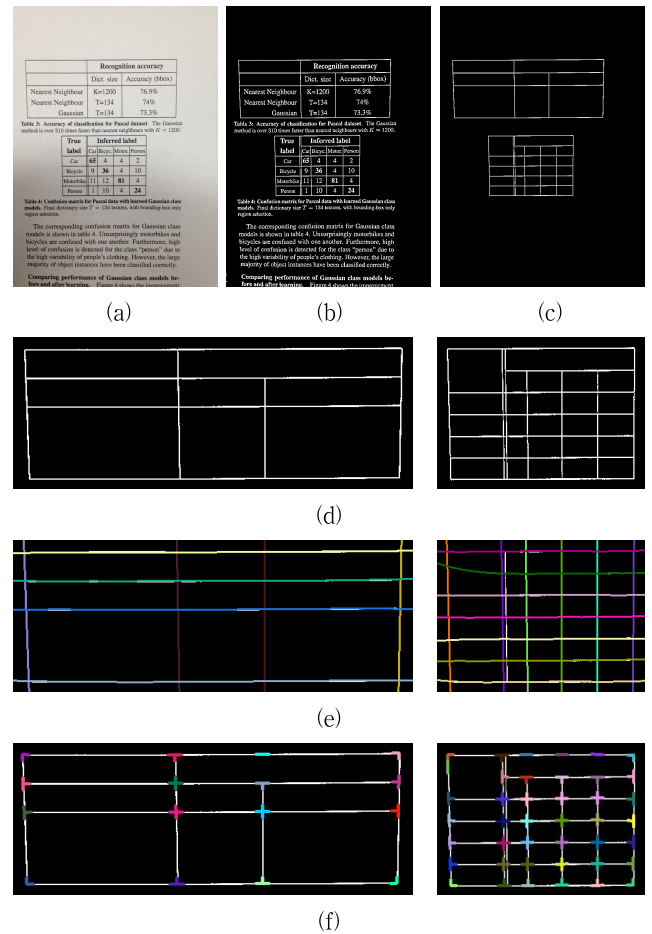


그림 1. 제안하는 알고리즘 진행과정 (a) 카메라로 취득한 표 영상 (b) 이진화 영상 (c) 연결요소 필터링 및 세선화 알고리즘 적용 영상 (d) 각각의 표 후보군에 대한 분할된 영상 (e) 가로, 세로 방향으로 곡선 맞춤 적용 결과 (f) 레이블링 결과 영상

영역 분할. 교차점을 찾기 위해서는 가로 방향과 세로 방향으로 가능한 정확하게 근사된 곡선을 찾을 필요가 있다. 하지만 하나의 문서 영상에 2개 이상의 표가 존재하거나 표를 구성하는 선 성분들을 제외한 나머지 선 성분들이 존재한다면 부정확한 근사 곡선을 얻을 가능성이 존재한다. 이를 해결하기 위해 각각의 표 후보들이 하나의 영역에 포함되도록 영역을 분할할 필요가 있다. Recursive X-Y cut은 일반적으로 평판 스캐너로 취득한 문서 영상에서 X방향과 Y방향으로 픽셀 값을 투영시킴으로써 영역을 분할하는 알고리즘이다[5]. 하지만 카메라로 취득한 문서 영상은 기하학적 왜곡을 가지고 있기 때문에 Recursive X-Y cut 알고리즘을 바로 적용하기에는 어려움이 있다. 따라서 제안하는 알고리즘에서는 단순히 X방향과 Y방향으로의 픽셀 값을 투영하는 대신 X방향과 Y방향의 minimal path를 이용한다. 먼저 Y방향으로의 minimal path는 dynamic programming을 이용하여 다음과 같이 누적된 path cost map d 를 구함으로써 얻을 수 있다.

$$d(i,j) = \min(d(i-1,j), d(i-1,j-1) + \lambda, d(i-1,j+1) + \lambda) + g^* I_{\alpha}(i,j) \quad (3)$$

여기서 $g^* I_{\alpha}(i,j)$ 는 연결요소 필터링 및 세선화 알고리즘 적용 후 가우시안 블러링 한 영상에서 (i,j) 의 픽셀 값이고, path cost map의 첫 번째 행은 이미지의 첫 번째 행과 같은 값을 갖는다. d 의 마지막 행이

각 path의 cost가 되고, 그 가운데 적당한 문턱 값보다 작으면서 지역적으로 작은 cost를 가지는 path를 역추적 함으로써 minimal path를 구할 수 있다. λ 는 작은 양의 상수로서 휘어지는 path의 cost를 높게 할당하게 되고 따라서 표를 둘러싸는 Y방향 path들의 cost가 표의 중심 방향으로 갈수록 커지게 만든다. X방향으로도 같은 작업을 수행하고, X와 Y 방향 cost를 비교하면서 recursive하게 알고리즘을 수행함으로써 효과적으로 영역을 분할할 수 있다. (그림 1(d))

곡선 맞춤. 교차점은 앞에서 언급한 바와 같이 표를 구성하는 선에 대해서 가로 방향과 세로 방향으로 곡선 맞춤을 수행함으로써 얻어질 수 있다. 가로 방향과 세로 방향 선들은 고차 다항 함수로 모델링 할 수 있지만, 일반적으로 표를 취득한 영상은 펼쳐진 책처럼 원통형 표면을 갖는 경우나 평면에 놓여있어서 평평한 표면을 갖는 경우가 대부분이기 때문에 세로 방향 선들은 일차 다항 함수로 모델링이 가능하다. 이를 통해 얻어진 각각의 교차점은 MRF(Markov Random Field)를 구성하는 그래프의 노드가 된다. (그림 1(e))

2.2 MRF 모델링

본 논문에서는 유효한 셀을 찾는 문제를 가로 방향 선들과 세로 방향 선들로 만들어지는 각각의 교차점에 최적의 레이블을 할당하는 방식으로 해결하고자 한다. 이를 위해 MRF 모델을 바탕으로 한 에너지 식을 세운다.

제안하는 방법에서 그래프 G 는 $G = \{V, E\}$ 로 구성되며, $V = \{V_1, V_2, \dots, V_N\}$ 는 N 개의 교차점이고, E 는 이웃한 노드들을 연결한 에지를 나타낸다. 이 그래프에 해당하는 에너지 식은 다음과 같다.

$$l^* = \min_p [\sum_p D_p(l_p) + \sum_{p,q \in N(p)} V_{p,q}(l_p, l_q)] \quad (3)$$

여기서 p 는 현재 사이트를 나타내며 $N(p)$ 는 p 의 이웃(neighbor)을 나타낸다. $D_p(\cdot)$ 는 데이터 향으로서 레이블 l_p 에 대응되는 관찰 값을 반영한 에너지를 나타내고, $V_{p,q}(\cdot)$ 는 둘로 이루어진 향으로서 이웃한 사이트에 레이블 l_q 가 위치했을 경우 l_p 와 l_q 사이의 관계에 해당하는 에너지를 반영한다. 본 논문에서는 이와 같은 에너지 식을 신뢰 확산(Belief Propagation) 알고리즘을 이용해 최소화함으로써 각 사이트에 최적의 레이블을 추정한다. (그림 1(f))

레이블. 제안하는 알고리즘에서는 표 1에서와 같이 각각의 교차점들이 가질 수 있는 레이블을 12가지로 분류한다. 예를 들어 레이블 {1}의 경우 교차점을 중심으로 위, 아래, 오른쪽 방향으로 유효한 선분을 갖는 레이블이고, 레이블 {11}의 경우 위, 아래, 왼쪽, 오른쪽 방향 모두 유효한 선분을 갖는 레이블이다. 레이블 {12}는 네 방향 모두 유효한 선분이 존재하지 않는 레이블이다. 간단한 표기를 위해 각 레이블을 4개의 이진수를 사용하여 나타낼 수 있다. 예를 들어 레이블 {1}, {2}, {3}, {4}는 각각 (1,1,1,0), (1,0,1,1), (0,1,1,1), (1,1,0,1)로 표기할 수 있는데, 이 표기법은 위쪽 방향에서 시계 방향으로 선분이 존재하는 지를 이진수로 표시한 것이다.

표 1. 각 교차점이 가질 수 있는 레이블

레이블 1	레이블 2	레이블 3	레이블 4	레이블 5	레이블 6
레이블 7	레이블 8	레이블 9	레이블 10	레이블 11	레이블 12

데이터 향. 각 교차점에 어떠한 레이블이 가장 적합한지는 각각의 레이블에 대한 공산(likelihood)으로부터 알 수 있고, 이 때 우리가 관찰할 수 있는 것은 그 교차점을 기준으로 네 방향에 존재하는 에지 픽셀의 수이다. 구체적으로 벡터 레이블 $l_p = (x_{u,p}, x_{r,p}, x_{d,p}, x_{l,p})$, $x_{u,p}, x_{r,p}, x_{d,p}, x_{l,p} \in \{0,1\}$ 에 대한 공산은 네 방향이 모두 독립이라는 가정 아래 다음과 같다.

$$f_p(l_p) = f_{u,p}(x_{u,p})f_{r,p}(x_{r,p})f_{d,p}(x_{d,p})f_{l,p}(x_{l,p}) \quad (4)$$

여기서 $f_{u,p}(x_{u,p} = 1)$ 는 p 기준으로 위쪽 방향에 선분이 존재할 때의 공산이며, $f_{u,p}(x_{u,p} = 0)$ 는 선분이 존재하지 않을 공산이다. 나머지는 다른 세 개 방향에 대해서 비슷하게 정의된다. 구체적으로 위쪽 방향으로 곡선을 따라 n 개의 픽셀을 관찰하였을 때, k 개가 에지 픽셀이었다면 선분이 존재할 공산은

$$f_{u,p}(x_{u,p} = 1) = p_c^k(1-p_c)^{n-k}$$

으로 주어지고, 반대의 경우

$$f_{u,p}(x_{u,p} = 0) = p_f^k(1-p_f)^{n-k}$$

으로 주어진다. p_c 는 선분이 실제 존재할 때, 그 선분 위에 픽셀이 에지로 검출될 확률이며, p_f 는 선분 위에 있지 않은 점이 에지 픽셀로 검출될 확률이다. 데이터 향의 값은 공산에 음의 로그를 취함으로써 구할 수 있다.

둘로 이루어진 향. 표를 구성하는 교차점의 경우 서로 이웃한 레이블끼리 가질 수 있는 레이블 쌍과 가질 수 없는 레이블 쌍이 존재한다. 예를 들어 임의의 교차점에 레이블 {1}이 위치한다면 오른쪽에 놓인 교차점에는 레이블 {2, 3, 4, 6, 8, 9, 11}이 올 수 있지만 레이블 {1, 5, 7, 10, 12}가 올 수는 없다. 따라서 이를 둘로 이루어진 향에 조건으로 설정할 수 있다. 둘로 이루어진 향에서는 가능한 레이블 쌍일 경우 0의 값을, 불가능한 레이블 쌍일 경우 ∞ 값을 준다.

2.3 셀의 구성

일반적으로 표의 가장 왼쪽과 가장 오른쪽 셀이 세로 선으로 닫혀 있는 경우도 있지만 세로 선이 없이 열려 있는 경우도 존재한다. 최종적으로 셀을 구성하기 위해 닫힌 셀을 구성하는 방법과 열린 셀을 구성하는 방법을 설명한다.

닫힌 셀 구성. 하나의 닫힌 셀이 구성되기 위해서는 셀의 좌상단 꼭짓점 레이블 {1, 3, 5, 11}, 우상단 꼭짓점 레이블 {2, 3, 6, 11}, 우하단 꼭짓점 레이블 {2, 4, 8, 1}, 좌하단 꼭짓점 레이블 {1, 4, 7, 11}을 만족해야 한다. 또한 각각의 셀 모서리에 레이블이 존재한다면 상단 모서리

레이블 {4, 9}, 우측 모서리 레이블 {1, 10}, 하단 모서리 레이블 {3, 9}, 좌측 모서리 레이블 {2, 10} 만이 올 수 있다.

열린 셀 구성. 왼쪽이 열려 있는 셀의 경우 각각의 세로 줄을 위에서부터 따라가면서 우상단 꼭짓점 후보인 레이블 {2, 3, 6, 11}을 찾는다. 만약 후보가 존재한다면 다시 그 세로 줄을 따라가면서 우하단 꼭짓점 후보 레이블 {2, 4, 8, 11}을 찾는다. 우하단 꼭짓점 후보 레이블을 찾은 과정에서 다른 레이블이 존재한다면 가능한 레이블은 {1, 10} 뿐이다. 마찬가지로 오른쪽이 열려 있는 셀의 경우 각각의 세로 줄을 위에서부터 따라가면서 좌상단 꼭짓점 후보 레이블 {1, 3, 5, 11}을 찾고 그 후보가 존재한다면 다시 그 줄을 따라 좌하단 꼭짓점 후보 레이블 {1, 4, 7, 11}을 찾는다. 이 과정에서 다른 레이블이 존재한다면 가능한 레이블은 {2, 10} 뿐이다.

표가 존재하는 대상을 원통형이나 평면 모델로 가정했기 때문에 세로 선들은 소실점을 가질 것이다. 왼쪽으로 열린 모든 셀들의 우상단, 우하단 꼭짓점에서 열린 방향으로 선을 따라가면서 가장 긴 선을 찾고 그 끝점을 소실점과 이음으로써 열린 셀을 효과적으로 구성할 수 있다. 오른쪽으로 열린 셀도 같은 방법으로 구성이 가능하다. 표를 정면에서 촬영하여 얻은 영상의 경우 세로 줄이 평행할 수 있기 때문에 이 경우 나머지 세로 줄들과 평행한 직선을 이동시키면서 열린 셀을 구성한다.

3. 실험 결과

본 절에서는 제안하는 알고리즘을 이용하여 일반 카메라로 취득한 표 영상에서 표의 구조 분석 및 인식에 대한 실험을 진행하였다. 모든 실험에서는 데이터 항을 구하기 위해 사용되어지는 관찰할 픽셀 개수를 위, 아래, 왼쪽, 오른쪽 모든 방향에 대해서 $n = 10$ 으로 고정시켰다. 또한 실제 선분 위의 한 점이 에지로 관찰될 확률은 $p_e = 0.9$ 로, 선분 위에 있지 않은 점이 에지로 관찰될 확률은 $p_f = 0.1$ 로 하였다. 전체 실험에서는 29개의 표 영상을 사용하였고 29개의 표 영상 안에는 49개의 표가 존재하였으며, 모두 1435개의 셀을 포함하였다. 표 2에서 볼 수 있듯이 셀 단위 검출 성공률은 92.4%, 표 단위 검출 성공률은 87.7%, 영상 단위 검출 성공률은 82.7%로 모두 높은 성공률을 달성함을 확인할 수 있었다. 일반 카메라로 획득한 문서 영상의 경우 종종 블러에 의해 영상의 질이 떨어지는데 검출 성공률을 떨어뜨리는 가장 큰 요인으로 분석된다.

그림 2에는 카메라로 취득한 표 영상과 제안하는 알고리즘을 적용한 결과 영상 샘플을 나타내었다. 한 개의 이상의 표가 존재하는 경우와 열린 셀을 포함하는 표에 대해서 모두 정확히 셀을 찾아내는 것을 확인할 수 있다.

4. 결론

본 논문에서는 디지털 카메라나 스마트폰과 같은 일반 카메라를 이용하여 획득한 표 영상에 대해서 표의 구조를 파악하는 방법을 제시하였다. 제안하는 알고리즘에서는 이러한 문제를 교차점에 레이블을 할당하는 문제로 구성하였고, 이를 풀기 위하여 MRF 기반의 에너지 식을 정의하고 최소화하였다. 각 교차점이 가질 수 있는 레이블의 형태를 12가지로 표현하였고, 이웃된 레이블 간에 위치할 수 없는 레이블

표 2. 셀, 표, 영상 단위 검출 성공률

	Total	Correct	Incorrect	Detection Rate
Cell	1435	1326	109	92.4%
Table	49	43	6	87.7%
Image	29	24	5	82.7%

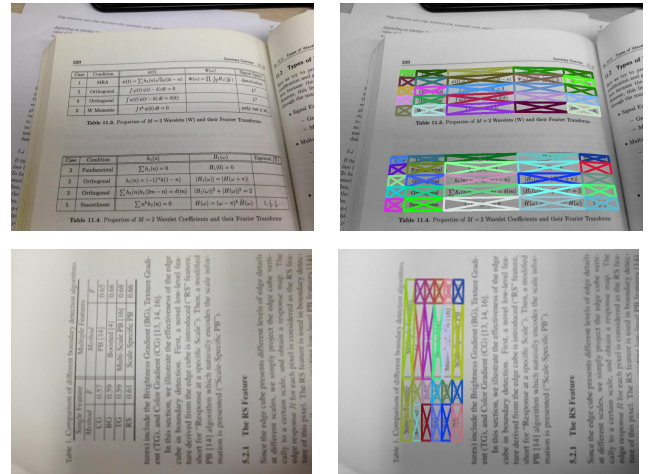


그림 2. 카메라로 취득한 표 영상과 제안한 알고리즘을 이용한 결과

쌍을 이용해 둘로 이루어진 항을 구성하였다. 카메라로 취득한 영상의 처리를 위해서는 이진화 과정을 거쳐야 하는데, 둘로 이루어진 항들을 추가함으로써 이진화 과정에서 소실되는 에지정보가 있더라도 표의 구성이 가능하도록 하였다. 실험을 통하여 제안하는 알고리즘을 적용했을 때 문서 영상에서 하나의 표가 있을 때뿐만 아니라 2개 이상의 표가 존재하는 경우에도 표의 구조를 잘 파악하는 것을 확인하였다. 또한 표의 양 측면이 단혀 있지 않은 경우에도 셀을 잘 검출함을 볼 수 있었다.

참고 문헌

- [1] Koo, H.I., Kim, J., Cho, N.I.: Composition of a dewarped and enhanced document image from two view images. *IEEE Trans. Image Process.* 18, 1551 - 1562 (2009)
- [2] M. Brown and Y.-C. Tsai, "Geometric and shading correction for images of printed materials using boundary," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1544-1554, June 2006.
- [3] N. Stamatopoulos, B. Gatos, I. Pratikakis, S.J. Perantonis, "Goal oriented Rectification of Camera-Based Document Images", *IEEE Trans. on Image Processing*, 2010.
- [4] M. Pilu and S. Pollard, "A light-weight text image processing method for handheld embedded cameras", *bmvc*, Mar. 2002.
- [5] G. Nagy and S. Seth, "Hierarchical representation of optically scanned documents." *Proc. of ICPP*, (1984) 347-349