

컬러 영상의 조명성분 분석을 통한 문자인식 성능 향상

최미영*, 김계영**, 최형일***

Improved Text Recognition using Analysis of Illumination Component in Color Images

Miyoung Choi *, Gyeyoung Kim **, Hyungil Choi ***

요약

본 논문에서는 컬러영상에 존재하는 문자들을 효율적으로 추출하기 위한 새로운 접근 방법을 제안한다. 빛 또는 조명성분의 영향에 의해 획득된 영상 내에 존재하는 반사성분은 문자 또는 관심객체들의 경계가 모호해 지거나 관심객체와 배경이 서로 혼합 되었을 경우, 문자추출 및 인식을 함에 있어서 오류를 포함시킬 수 있다. 따라서 영상 내에 존재하는 반사성분을 제거하기 위해 먼저, 컬러영상으로부터 Red컬러 성분에 해당하는 히스토그램에서 두개의 pick점을 검출한다. 이후 검출된 두 개의 pick점들 간의 분포를 사용하여 노말 또는 편광 영상에 해당하는지를 판별한다. 노말 영상의 경우 부가적인 처리를 거치지 않고 문자에 해당하는 영역을 검출하며, 편광 영상에 해당하는 경우 반사성분을 제거하기 위해 호모모픽필터링 방법을 적용하여 반사성분에 해당하는 영역을 제거한다. 이후 문자영역을 검출하기 위해 최적전역임계화방식을 적용하여 전경과 배경을 분리하였으며 문자영역 추출 및 인식의 성능을 향상시켰다. 널리 사용되고 있는 문자 인식기를 사용하여 제안한 방식 적용 전과 후의 인식결과를 비교하였다. 편광영상에서 제안된 방법 적용 후, 문자인식을 한 경우 인식률이 향상되었다.

Abstract

This paper proposes a new approach to eliminate the reflectance component for the detection of text in color images. Color images, printed by color printing technology, normally have an illumination component as well as a reflectance component. It is well known that a reflectance component usually obstructs the task of detecting and recognizing objects like texts in the scene, since it blurs out an overall image. We have developed an approach that efficiently removes reflectance components while preserving illumination components. We decided whether an input image hits Normal or Polarized for determining the light environment, using the histogram which consisted

• 제1저자 : 최미영

• 접수일 : 2007.5.18, 심사일 : 2007.6.18, 심사완료일 : 2007. 7.23.

* 숭실대학교 컴퓨터학과

** 숭실대학교 컴퓨터학과 부교수

*** 숭실대학교 미디어학과 교수

of a red component. We were able to go ahead through the ability to extract by reducing the blur phenomenon of text by light because reflection component by an illumination change and removed it and extracted text. The experimental results have shown a superior performance even when an image has a complex background. Text detection and recognition performance is influenced by changing the illumination condition. Our method is robust to the images with different illumination conditions.

▶ Keyword : 문자추출(text localization), 문자인식(text recognition), 컬러영상(color image)

I. 서 론

최근에 일상생활에서 많이 사용하고 있는 핸드폰 카메라, 디지털 카메라 그리고 디지털 캠코더로 취득된 정지영상이나 비디오 프레임에 인위적으로 삽입되거나 자연적으로 포함된 텍스트들은 영상 내의 내용을 함축적이고 구체적으로 표현하는 중요한 정보들이다. 이러한 정보들을 실시간으로 추출하여 인식할 수 있다면 휴머노이드 로봇의 자동 주행 시스템, 시각 장애인 보행 지원 시스템, 지능형 (교통) 운전 시스템, 박물관 등의 정보 안내 시스템, 고용량 비디오 프레임의 자동 검색 및 색인 시스템, 텍스트 자동 번역 시스템 등과 같은 다양한 분야에 널리 활용 될 수 있다.

일반적인 환경에서 획득된 칼라 영상은 단색조 영상에 비해 인간의 시각을 크게 향상시키므로, 장면에 있어서 물체에 관한 많은 정보를 제공할 수 있다. 따라서, 칼라 영상은 그래픽 예술, 모션 사진, 멀티미디어, 통신, 컴퓨터 비전 등에서 자주 이용된다[1]. 그러나 칼라 영상처리 분야는 칼라의 복잡한 속성과 계산량 때문에 단색조 영상에 비해 많은 연구가 이루어지지 않았으나, 프로세서의 속도가 개선됨에 따라 최근에는 이에 대한 연구가 활성화 되고 있다. 또한 조명의 변화에 강건한 물체인식, 문자인식, 이동물체 검출 등에 대한 많은 연구가 진행 중이며 본 논문에서는 조명의 변화를 예측하여 반사성분을 제거하는 방법을 제안한다.

II. 제안한 시스템 개요

일반적으로 영상에 존재하는 문자들은 많은 의미 있는 정보들을 포함하고 있다. 이러한 정보들을 실시간으로 추출하여 인식하기에는 많은 어려움이 있다. 그 중에 가장 큰 어려움은 조명의 변화이다. 대부분의 컴퓨터 비전 알고리즘들은 일정한 조명 조건하에 수행되며, 이런 가정 하에서는 거의 정확한 결과들을 보여준다. 그러나 실제 환경에서의

자연광 및 조명의 변화는 일정하지 않을 뿐만 아니라 변화 무쌍하다. 영상에 포함되어 있는 다양한 조명 변화를 인지하여 제거 할 수만 있다면, 실생활에서 많은 의미 있는 정보들을 쉽게 획득하여 인식 할 수 있을 것이다.

대부분의 영상에서 문자영역 추출은 조명이나 빛의 영향이 적은 흑백 문서 형태의 영상을 대상으로 적용되었다. 본 논문은 컬러 프린팅 기술에 의해 인쇄된 문서 및 물체 표면의 문자 영역에서 빛 또는 조명의 표면반사에 의해 발생하는 영역 추출의 문제점을 해결하고자 조명의 변화를 분석하여 제거하는 방법을 제안한다. 조명의 변화에 의해 발생 할 수 있는 영상은 일반적으로 노말(normal), 편광(polarized), 역광(backside) 세 가지로 구분 할 수 있다. 본 논문에서는 조명의 변화를 그림 1과 같이 노말과 편광 두 가지 상태로 구분하였다. 역광이란 피사체의 뒤쪽이 밝아서 상대적으로 피사체의 모습이 어둡게 되는 경우를 말한다. 조명이 문서의 뒤에서 발생하는 경우는 일반적으로 발생되지 않기 때문에 제외시켰다.



(a)노말(좌측),편광(우측)



(b)노말(좌측),편광(우측)

그림 1. 노말 영상과 편광 영상
Fig.1 Normal image and Polarized image

제안된 알고리즘은 그림 2와 같다. 먼저 입력영상의 red 성분에 해당하는 히스토그램을 분석하여 영상의 조명환경 상태를 판별한다. 영상의 조명 상태가 노말인 경우 HMMP(homomorphic filter)[2] 과정을 수행하지 않고 문자영역을 검출하게 되며, 편광

의 경우는 HMMP 과정을 수행하여 조명에 의한 반사성분으로 판단되는 계수 성분을 제거한 다음 문자영역을 검출한다.

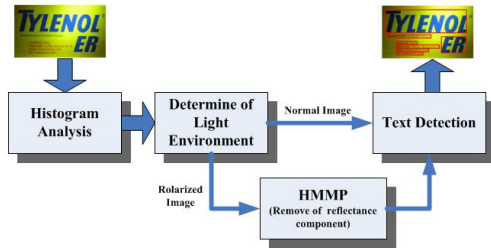


그림 2. 시스템 개요도
Fig.2 The flow of the proposed algorithm

III. 호모모픽 필터링을 이용한 반사성분제거

3.1 조명 환경의 정의 및 판단

입력된 영상의 조명 환경이 노말 상태인지 편광 상태인지를 판단하기 위해, R, G, B 성분 중 Red컬러 성분에 해당하는 히스토그램을 이용한다. Red컬러 성분은 조명의 변화에 가장 민감하게 반응하기 때문이다. 영상의 조명 상태를 판별하기 위해 본 논문에서 제안한 방법은 다음과 같다.

첫째, 그림 3과 같이 Red컬러 성분에 해당하는 히스토그램을 세 개의 그룹 S(Shadow), M(Midtone) and H(Highlight) 영역으로 구분하였다. Shadow 영역은 0~79, Midton 영역은 80~180, Highlight 영역은 181~255로 정의하였다.

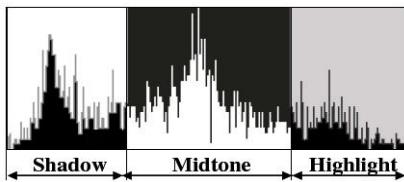


그림 3. RED컬러 컴포넌트 히스토그램
Fig.3 Histogram of Red component

둘째, Red컬러 히스토그램에서 그림 4와 같이 2개의 pick를 검출한다. Pick를 검출하기 위한 방법은 Ping-Sung Liao multi-threshold algorithm(3)을 사용하였다.

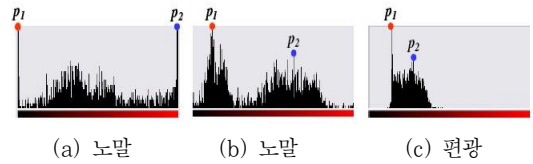


그림 4. 두 가지(노말, 편광) 조명 상태에서의 red 컬러 히스토그램 분포 예시

Fig.4 Typical distribution of the red color histograms obtained under two lighting environments

셋째, 영상의 조명환경을 판별은 검출된 두 pick가 수식 1의 L(LightEnvironment)에 의해 Red컬러 히스토그램의 세 그룹(S, M, H)중 어디에 포함되는지와 두 pick간의 거리(d)에 의해 결정된다. 그림 5는 조명환경에 따른 영상의 히스토그램 분포 및 pick의 위치를 보여준다.

$$L = \begin{cases} Normal, & 70 \leq d \leq 130, & \begin{cases} P_i, P_j \in S \\ \text{or } P_i, P_j \in H \end{cases} \\ Polarized, & 140 \leq d \leq 210, & \begin{cases} P_i \in S, P_j \in H \\ \text{or } P_i \in S, P_j \in M \\ \text{or } P_i \in M, P_j \in H \end{cases} \end{cases} \dots\dots (1)$$

$$d = |P_i - P_j|$$

P_i 와 P_j 는 히스토그램에서 2개의 pick를 나타낸다.



그림 5. 편광 영상의 Red 히스토그램 분포 및 2개의 pick
Fig.5 Histogram distribution of polarized image and position of picks

영상의 조명 환경이 노말에 해당하는 경우 빛이 전체적으로 고르게 분포한다고 판단하여 호모모픽 필터링을 수행하지 않고 문자영역 추출과정을 수행하며, 편광에 해당하는 경우는 빛이 어느 한쪽 방향에서 강하게 비취짐에 따라 물체에서 반사되었다고 판단하고 호모모픽 필터링을 적용하여 반사성분에 해당하는 계수성분을 제거한 후, 문자영역 추출과정을 수행한다.

3.2 호모모픽 필터링에 의한 조명 성분 제거

호모모픽 필터링은 푸리에 변환을 기반으로 주파수 영역에서 조명에 의해 만들어진 저주파를 감소시키고 고주파를 증폭시킴으로써 대비도를 높여 영상을 향상시킨다. 균질하지 않은 조명의 반사성분을 제거하기 위해 호모모픽 필터링을 사용했다. 실제 환경에서 조명성분은 대체로 느리게 변화하는 반면, 반사성분은 매우 빠르게 변화하는 특성을 가지고 있다[4].

영상은 식 2와 같이 두 가지 성분으로 구분할 수 있다. 장면에 입사된 조명 $I(x,y)$ 과 장면에 있는 물체에서 반사된 조명 $R(x,y)$ 의 곱으로 표현하며, 조명(Illumination)성분과 반사(Reflectance)성분이라고 한다. 그림 6의 호모모픽 필터는 식 3에 의해 영상을 주파수 영역으로 적용할 수 있다. 입력 영상이 다양한 조명과 반사성분을 가진다고 가정하면,

$$\ln [F(x,y)] = \ln [I(x,y)] + \ln [R(x,y)] \dots\dots\dots (2)$$

$$H(u,v) = (\gamma_H - \gamma_L) [1 - e^{-c(D^2(u,v)/D_0^2)}] + \gamma_L \dots\dots\dots (3)$$

주파수 영역에서 조명성분은 저주파에 해당하며 반사성분은 고주파에 해당한다. 일반적으로 조명성분은 변화가 느리고 연계적인 반면, 반사성분은 특히 경계부분에서 민감한 반응을 보인다.

호모모픽 필터링에 의한 조명성분 제거과정을 수행함으로써 영상의 비균질한 조명성분을 균질하게 하며 또한 대비도를 향상시킨다.



그림 6. 호모모픽 필터링 수행 과정
Fig.6 Homomorphic filtering approach

그림 7과 같이 실험에 의해 반사성분이 제거되었다. 여러 실험영상을 대상으로 실험한 결과 0.75~1.25에 해당하는 계수 값을 반사성분으로 간주하고 제거하였으며, 제거한 부분의 값은 주변에 인접영역의 픽셀 값을 기반으로 보간한다.

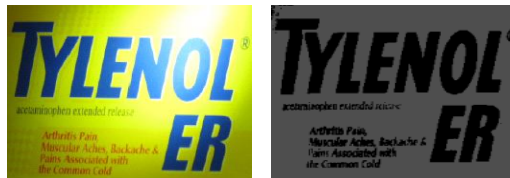


그림 7. 입력영상과 반사성분이 제거된 영상
Fig.7 Original image(left) and removed image of reflectance component(right)

3.3 컬러영상에서 문자영역 추출

문자영역 추출 과정은 크게 두 가지 단계로 구성되어 있다. 첫째, threshold 단계는 전역 최적 임계화 방법(global optimal thresholding method) [5][6]을 적용하여 배경으로부터 전경을 분리한다. 둘째, localization 단계는 Run Length Smoothing 방법(RLS) [7][8]을 적용하여 분리된 전경에서 문자 영역을 추출한다.

해상도가 낮은 영상일수록 히스토그램의 피크가 겹쳐서 나타나는 경향이 있어 배경으로부터 전경을 분리하는 것은 매우 어려운 일이다. 그래서 최소 평균 분할 에러를 산출하여 임계값을 예측하는 방법을 사용하였다. 초기 T값은 전체 영상의 평균값(μ)으로 한다.

만약 두 분산이 같으면, $\sigma^2 = \sigma_1^2 = \sigma_2^2$, 단일 임계값이면 충분하다.

$$T = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2}{\mu_1 + \mu_2} \ln \left(\frac{P_2}{P_1} \right) \dots\dots\dots (4)$$

μ_1 과 σ_1^2 는 한 화소 집합(object)의 가우스 밀도의 평균과 분산이고, μ_2 과 σ_2^2 는 다른 집합의 평균과 분산이다.

$P_1 = P_2$ 이면, 최적의 임계값은 평균값들의 평균

$$T = \frac{\mu_1 + \mu_2}{2} \text{ 이 된다.}$$

$$f(x,y) = \begin{cases} 1 & \text{if } f(x,y) \geq T \\ 0 & \text{otherwise} \end{cases} \dots\dots\dots (5)$$

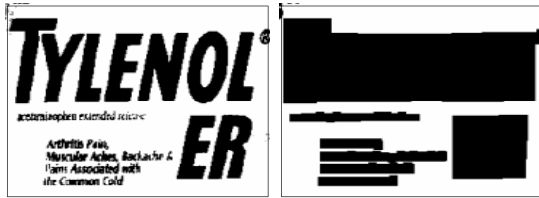


그림 8. 전역 최적 임계화 적용한 결과와 전역 최적 임계화 적용 후 Run Length Smoothing 적용한 결과
 Fig.8 The result of applying the Global Optimal Thresholding(left) and the run length smoothing algorithm(right)

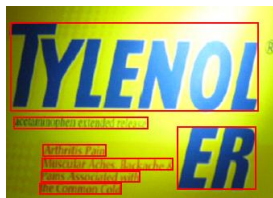


그림 9. 문자영역 추출
 Fig.9 Detected text region

RLS는 주로 연결요소를 분석하여 영상을 분리하는 방법으로 사용된다. RLS는 수평방향, 수직방향으로 존재하는 흰색 띠를 조사한다. 영상에서 주어진 임계값 보다 작은 모든 흰색 띠를 선택하여 각각의 방향에 대해 수평방향 스무딩과 수직방향 스무딩을 수행한다. 수평방향 스무딩 결과 영상과 수직방향 스무딩 결과 영상을 AND연산하여 문자 영역을 추출 할 수 있다.

IV. 실험결과

일상생활에서 많이 사용하고 있는 핸드폰 카메라, 디지털 카메라로 취득된 컬러 영상에서 사람이 조명의 변화를 인지하는 것과 같이 조명에 의한 반사성분을 제거하는 방법을 제안했다. 실험은 펜티엄IV CPU 3.0GHz, 1GB RAM, Windows XP 환경에서 Visual C++을 사용하였으며, 640*480의 해상도의 편광영상 350장, 노말영상 150장으로 총 500장의 컬러영상을 실험영상으로 획득하여 사용하였다.



(a) 결과영상 (b) 그레이 영상 (c) 반사성분 제거영상 (d) 이진영상

그림 10. 실험 결과영상
 Fig.10 Experiment result images

(a) Result image (b) Gray image (c) Removed reflectance component of image (d) Binary image

V. 결론

본 논문에서는 컬러영상에 존재하는 문자들을 효율적으로 추출하기 위한 새로운 접근 방법을 제안했다. 카메라로부터 획득된 컬러영상의 조명환경을 판별하기 위해 첫째, Red컬러 성분 히스토그램에서 두 개의 pick점을 검출하고, 검출된 두 pick점들 간의 분포를 사용하여 노말 또는 편광영상에 해당하는지를 판별한다. 둘째, 호모모픽 필터링으로 조명에 의한 반사성분을 제거하였다. 셋째, 전역 최적 임계화 방법으로 전경과 배경을 분리하여 문자영역 추출 및 인식의 성능을 향상시켰다. 넷째, RLS를 적용하여 분리된 전경에서 문자영역을 추출했다.

획득된 컬러영상에서 바로 문자영역 추출 후 문자 인식을 사용하여 문자인식 했을 때의 인식률과 제안한 방법 적용 후 문자인식기 ABBYY FinReader 6[9]과 아르미 6.5[10]를 사용하여 문자인식 한 결과를 비교하였다. 편광영상에서 제안된 방법 적용 후, 문자인식을 한 경우 인식률이 향상되었음을 알 수 있다.

표1. 노말영상과 편광영상에서 문자인식기만 사용한 문자인식률과 제안된 방법 후 문자인식률 비교

Table 1. OCR recognition rate with and without the incorporation of the proposed method

	노말 영상	편광 영상
FineReader6 or 아르미6.5	72.1 %	61.3%
Proposed method + FineReader6 or 아르미 6.5	-	70.4%

참고문헌

[1] 엄경배, 한서원, 이준환, "혼합된 칼라 잡음하에서 칼라 영상 향상을 위한 조직적인 퍼지 클러스터 필터", 한국정보처리학회 논문지, 제6권, 제12호, 1999.
 [2] K. Delac, M. Grgic, T. Kos, "Sub-Image Homomorphic Filtering Technique for Improving Facial Identification", International Conference on Systems, Signals and Image

Processing, pp.95-98, 2006.
 [3] Ping-Sung Liao, Tse-Sheng Chen, Pau-Choo Chung, "A Fast Algorithm for Multilevel Thresholding", Information Science and Engineering, Vol. 17, 2001.
 [4] J.R. Parker, "Algorithms for Image Processing and Computer Vision", Wiley computer publishing, 1997.
 [5] S.E. Umbaugh, "Digital Image Analysis and Processing", CRC Press, Florida, 2005.
 [6] R.C. Gonzalez, R.E. Woods, "Digital Image Processing", Prentice Hall, Upper Saddle River, NJ, 2002.
 [7] N.Papamarkos, J. Tzortzakis and B. Gatos, "Determination of run-length smoothing values for document segmentation", ICECS, 1996.
 [8] Qing. Hong Liu, "Newspaper Headlines Extraction from Microfilm Images", IEEE, 2002.
 [9] ABBYY : www.finereader.com, 2005.
 [10] Aremi : www.perceptcom.com, 2006.

저자 소개

최 미 영

승실대학교 컴퓨터학과

김 계 영

승실대학교 컴퓨터학과 부교수

최 형 일

승실대학교 미디어학과 교수