

카메라 영상의 화질 분석에 의한 적응적 이진화

김인중

한동대학교 전산전자공학부

e-mail : ijkim@handong.edu

Adaptive Binarization for Camera Image by Image Quality Estimation

In-Jung Kim

School of Computer Science and Electronic Engineering

요 약

카메라로 촬영한 문서 영상의 분석 및 인식을 위한 적응적 이진화 방법을 제안한다. 먼저 카메라 영상의 화질을 분석하여 정량화하는 방법을 제안한다. 그 후 다양한 화질의 카메라 영상에 대하여 제안한 방법으로 측정된 화질과 그 화질에 적합한 이진화 파라미터간의 상관 관계를 통계적으로 분석한다. 그 후 그들간의 상관 관계를 이용하여 입력 영상의 화질에 적합한 이진화 파라미터를 자동으로 추정하도록 함으로써 화질 변화에 대한 적응력을 강화하였다. 실험을 통해 초점 및 잡영의 상태에 따라 적절한 파라미터를 추정함으로써 화질 변화에 적응하는 결과를 확인하였다.

1. 서론

지난 수 십 년 동안 문서 인식 기술은 많은 발전을 이루었으며, 이로 인해 다양한 분야에 성공적으로 적용되어 왔다. 이 같은 문서 인식 시스템의 주요 스캐너로 입력 받은 영상을 인식 대상으로 하고 있다.

수 년 전부터는 카메라 폰을 비롯한 디지털 카메라의 보급이 급격히 증가하면서 카메라로 촬영한 문서를 인식하려는 노력이 증가하고 있다. 카메라 영상은 컬러, 또는 명도 영상으로 입력되며, 해상도가 낮을 뿐 아니라 조명, 촬영 거리 및 각도 등 촬영 환경에 따라서 매우 많은 화질의 변이가 발생한다. 이러한 변이로 인해 카메라 영상의 인식은 스캐너 영상의 인식에 비하여 난이도가 훨씬 높은 것으로 알려져 있다.

카메라 영상에 대하여 높은 성능의 인식 성능을 얻기 위해서는 다양한 화질에 적응할 수 있는 영상처리 기술이 필수적이다. 특히, 문서 인식과정에서는 많은 분석 및 인식 단계가 이진 영상에 대하여 수행되기 때문에 화질 변화에 적응력이 강한 이진화 방법론은 매우 중요하다.

그 동안 명도 영상의 이진화에 대하여는 매우 많은 연구가 이루어졌다. 그러나, 아직도 카메라 영상의 다양한 화질에 대해서는 만족할 만한 적응력을 보이지

못하고 있다. 이러한 문제점을 극복하기 위해 본 논문에서는 입력 영상의 화질을 분석하여 수치화 한 후 그 결과에 따라 적절한 파라미터를 자동으로 추정함으로써 화질 변화에 적응할 수 있는 이진화 방법을 제안한다.

먼저 초점과 잡영의 정도를 수치화 하는 방법을 제안한 후, 이를 이용하여 다양한 다수의 영상에 대하여 적합한 이진화 파라미터와 화질 수치 사이의 상관 관계를 통계적으로 분석한다. 그 상관 관계를 이용해 입력 영상에 맞는 이진화 파라미터를 화질에 따라 적응적으로 추정하도록 함으로써 다양한 화질에 효과적으로 적응할 수 있도록 한다.

본 논문의 구성은 다음과 같다. 2 절과 3 절에서는 기존의 이진화 관련 연구와 제안하는 방법의 기반이 되는 Sauvola 의 방법을 분석하여 입력 영상의 화질이 이진화 파라미터에 미치는 영향을 설명한다. 4 절에서는 영상의 초점과 잡영의 정도를 측정하여 수치화 하는 방법을 소개하고 5 절에서는 초점 및 잡영의 측정치와 이진화 파라미터 간의 상관관계를 통계적으로 추정해 이진화에 이용하는 방법을 설명한다. 6 절에서는 실험 결과를 소개하고, 7 절에서는 결론을 맺는다.

2. 기존 연구

이진화는 컬러 영상이나 명도 영상에 대하여 각 화소의 명도가 특정 임계치보다 클 경우 흰색으로, 그렇지 않을 경우 검정으로 변환함으로써 이진 영상으로 변환하는 과정이다. 이 때 임계치의 설정 방식에 따라 전역적 이진화와 지역적 이진화로 나눌 수 있다. 전역적 이진화는 영상 전체에 대하여 동일한 임계치를 적용해 이진화를 수행하는 반면 지역적 이진화는 각 화소, 또는 영역에 따라 다른 임계치를 이용해 이진화를 수행한다.

전역적 이진화는 속도가 빠르고 조명의 변화가 크지 않은 영상에 대하여는 매우 효과적이다. 특히, Otsu의 임계치 추정 방법은 성능이 우수하여 널리 사용되고 있다[1]. 그러나, 전역적 이진화는 하나의 임계치를 영상의 모든 화소에 적용하기 때문에 명도가 불균일한 영상에 카메라 영상에는 적용하기 어렵다.

반면 지역적 이진화는 화소, 또는 영역에 따라 임계치를 각각 설정하기 때문에 명도가 불균일한 영상에 대한 적응력이 우수하다. 지금까지 많은 지역적 이진화 방법들이 제안되어 왔다. Tier 는 여러 가지 이진화 방법에 대하여 비교 평가를 수행했는데, 실험 결과 Niblack 의 방법이 가장 우수한 것으로 보고하였다.[2,3] 그 후 Sauvola 는 Niblack 의 이진화 방법을 좀더 개선한 방법을 제안하였다[4].

Sauvola 의 방법은 각 화소에 대하여 그 화소를 중심으로 하는 window 를 설정하고 그 window 안 화소들의 평균 명도와 표준편차를 계산한다. 그 후 그 값을 이용하여 다음 식과 같이 지역 임계치 $T(x,y)$ 를 계산한 다음, 해당 화소의 명도가 $T(x,y)$ 보다 크면 흰색, 작으면 검정색으로 변환한다.

$$T(x, y) = m(x, y) \left[1 + k \left(1 - \frac{s(x, y)}{R} \right) \right]$$

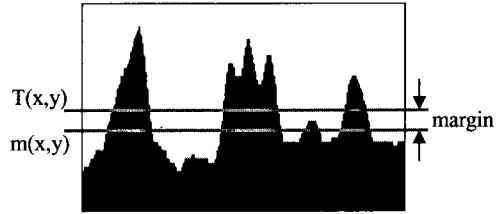
위의 식에서 $m(x, y)$ 와 $s(x,y)$ 는 각각 화소 (x,y) 를 중심으로 하는 window 내의 화소들의 명도에 대한 지역 평균과 표준 편차를 나타낸다. 또한 k 와 R 은 각각 이진화 지수(ratio)와 표준 편차의 범위(dynamic range)이다. 입력 영상이 256 레벨의 명도 영상일 경우 표준 편차의 최대값은 128 이므로 R 은 128 의 값으로 고정된다.

이 방법은 성능이 우수한 것으로 알려져 있어 널리 사용되나, 두 가지 파라미터의 영향을 받는다. 첫째는 window 의 크기이고 둘째는 이진화 지수 k 이다. 이 중 전자에 대해서는 기존 연구에서 이미 분석된 바 있다[5]. 따라서 본 논문에서는 이진화 지수 k 가 이진화에 미치는 영향에 대하여 분석하도록 하겠다.

3. Sauvola 의 방법에서 이진화 지수의 영향

Sauvola 의 방법에서 이진화 지수가 미치는 영향을 분석하기 위해서는 Sauvola 의 공식을 이해하여야 한다. 그림 1 은 한 명도 영상의 단면이다. Sauvola 의 방법은 그림 1 에서와 같이 $T(x,y)$ 를 지역 평균 $m(x,y)$ 에 따라서 정하되 $(1+\alpha)$ 의 값을 곱함으로써 상향으로 조

정한다. 이는 지역 평균을 직접 지역 임계치로 설정하는 경우에 비하여 미세한 명도 변화에 둔감하게 하는 효과가 있다. 따라서 α 가 크면 명도의 변화에 둔감해 지는 반면 α 가 작으면 $T(x,y)$ 가 $m(x,y)$ 의 값에 근접하게 되어 작은 명도의 변화에도 민감하게 된다.



(그림 1) 명도 영상의 단면

Sauvola 의 방법에서 α 는 다음과 같다.

$$\alpha = k \left(1 - \frac{s(x, y)}{R} \right)$$

이 중 $(1-s(x,y)/R)$ 은 표준 편차의 영향을 반영한 것으로서 표준 편차에 반비례하도록 하고 있다. 여기에 곱해진 이진화 지수 k 는 결국 표준 편차의 영향력을 조절하는 역할을 한다. 즉 k 가 작으면 $s(x,y)$ 의 영향이 작아져서 $T(x,y)$ 가 $m(x,y)$ 에 근접하게 되는 반면 k 가 커지면 margin 이 커지기 때문에 작은 변화는 무시되고 흰색으로 판정된다. 결론적으로 이진화 지수 k 는 전체적으로 $T(x,y)$ 가 결정되는 과정에서 명도 변화의 민감도를 결정한다.

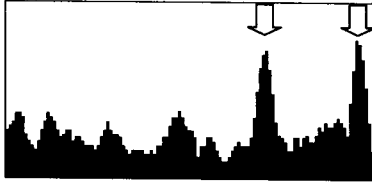
Sauvola 는 이진화 지수를 특정 값으로 고정시켰으나, 화질의 변화가 다양한 카메라 영상에 대하여 특정 값으로 고정할 경우 충분한 적응력을 얻을 수 없다. 잡영이 많은 영상의 경우 큰 값을 지정하여 명도의 작은 변화는 무시하는 것이 바람직하며, 초점이 흐린 영상에 대하여는 작은 값으로 지정하여 흐린 확대 섬세하게 추출하는 것이 효과적이다. 표준 편차 $s(x,y)$ 가 이 같은 적응을 돕지만, 화질 편차가 크고 특히 초점이 부정확할 수 있는 카메라 영상에 대해서는 지역 표준 편차 만으로는 충분한 적응력을 얻을 수 없다. 따라서 카메라 영상에 대해서는 이진화 지수의 값이 화질에 따라 적응적으로 결정되어야 한다.

이진화 지수로 지나치게 작은 값을 지정할 경우 명도 변화에 지나치게 민감해져서 잡영이 많이 발생하고, 반대로 지나치게 큰 값을 지정할 경우 초점이 흐린 영상에 대하여 문자 획들이 소실되는 문제가 발생한다. 따라서 영상마다 적절한 이진화 지수를 결정하기 위하여는 입력 영상의 초점과 잡영의 정도를 측정하는 방법이 요구된다.

4. 화질 분석

본 연구에서는 이진화 지수와 관련이 깊은 초점과 잡영의 측정 방법에 중점을 두었다. 초점은 배경 영역과 전경 영역의 명도 차이가 얼마나 뚜렷한 지를 반영한다. 그림 2 에서 화살표로 표시된 봉우리들은 문자 획의 일부를 나타낸다. 따라서 전경 영역으로

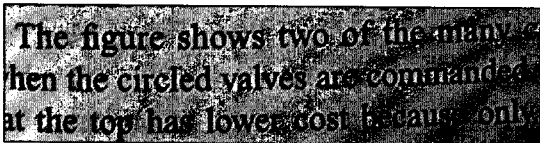
분리되어야 하며 이들과 인접한 배경 영역과의 명도 차이는 초점의 정확도와 비례한다. 반면 좌측에 나타나는 명도의 변화는 전경과 배경의 구분과 무관하며 단순히 배경에 나타난 불필요한 명도 변화로서 이진화 시 잡영의 원인이 된다.



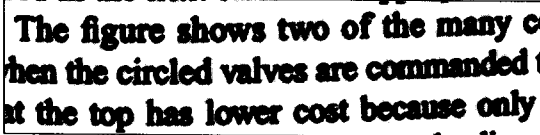
(그림 2) 초점과 잡영

이러한 관점에서 볼 때 초점과 잡영의 정도는 모두 명도의 변화와 밀접한 관계가 있으나, 초점은 배경과 전경의 경계에서의 명도 변화, 잡영은 배경, 또는 전경 내에서의 명도 변화를 각각 의미한다고 할 수 있다. 따라서, 영상에서의 명도 변화를 측정할 후 그것을 각각 전경과 배경의 경계선과 그렇지 않은 부분에 대하여 구분한다면 초점과 잡영을 측정할 수 있다.

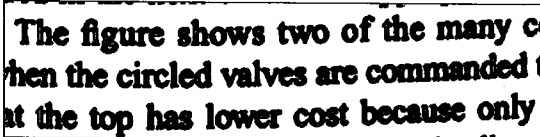
본 연구에서는 Sobel 의 edge operator 를 이용하여 명도의 변화량을 측정하였다. 예를 들어 그림 3(a)의 영상에 Sobel operator 를 적용할 경우 그림 3(b)과 같은 결과를 얻을 수 있다. 여기에서 각 화소의 명도는 edge 값의 크기를 의미한다. 그림 3에서 볼 때 전경과 배경의 경계는 edge 값이 큰 값을 갖는 반면, 배경 또는 전경 내에서의 명도 변화는 상대적으로 작은 값을 갖는다. 따라서, edge 값이 큰 화소는 전/배경의 경계 화소로, 그렇지 않은 화소는 비경계화소로 구분할 수 있다. 따라서 그림 3(b)와 같은 edge 영상에 이진화를 수행하면 그림 3(c)와 같이 경계화소를 추출할 수 있다. 본 연구에서는 이를 위해 Otsu 의 방법을 적용하였다.



(a) 명도 영상



(b) Edge operator 적용 결과



(c) 그림 (b)에 대한 전역적 이진화 결과

(그림 3) 경계화소의 추출

그림 3(c) 와 같이 경계화소와 비경계화소를 구분한 후에는 각각의 평균 edge 값을 계산할 수 있다. 경계화소의 평균 edge 값은 배경과 전경의 경계에서의 명도 변화를 의미하므로 초점 값으로 사용될 수 있고, 비경계화소의 평균 edge 값은 잡영의 정도로 사용될 수 있다.

5. 화질 분석에 의한 적응적 이진화

입력 영상으로부터 초점과 잡영의 정도를 측정할 후에는 그 결과와 각 영상에 적합한 이진화 지수 사이의 상관 관계를 통계적으로 분석할 수 있다. 각 영상에 적합한 이진화 지수를 수학적으로 찾는 것은 어려운 일이다. 그러나, 사람이 다수의 영상에 대하여 최적의 이진화 결과를 내는 이진화 지수를 결정하도록 한다면 각 영상에 대한 최적 이진화 지수를 측정할 수 있다. 그 결과를 4 절에서 설명한 방법으로 측정된 영상의 초점 및 잡영 값과의 연관성을 통계적으로 분석하면 화질과 그에 따른 최적 이진화 지수간의 관계를 예측할 수 있다.

먼저 각각의 카메라 영상에 대하여 이진화 지수를 변화시키며 이진화를 수행하여 그 중 최적의 결과를 내는 이진화 지수를 기록한다. 이 때 각 영상의 초점과 잡영 측정치는 고려하지 않아야 한다. 그 후 각 영상의 초점과 잡영을 측정하여 그들간의 상관 관계를 통계적으로 분석한다.

본 연구에서는 최적 이진화 지수가 초점 및 잡영 수치와 통계적 관계를 분석하기 위해 선형 회귀 분석 (Linear Regression)을 이용하였다. 즉 초점과 이진화 지수, 그리고 잡영과 이진화 지수간의 상관 관계를 통계적으로 조사한 후 이진화 지수를 다음과 같이 초점 (focus) 및 잡영(noise)의 선형 조합에 의해 표현하였다. 여기에서 가중치 a, b, c 는 모두 세 변수간의 다변량 회귀 분석에 의해 통계적으로 계산하였다.

$$k = a * focus + b * noise + c \quad (1)$$

새로운 카메라 영상을 이진화하기 위해서는 먼저 초점과 잡영을 측정할 후 위의 식에 의해 적정한 이진화 지수를 추정하고 그것을 이용하여 Sauvola 의 방법을 이용해 이진화를 수행한다.

6. 실험

6.1. 초점 및 잡영과 이진화 지수의 상관 관계

본 연구에서는 초점과 잡영의 정도가 다양한 85 개의 카메라 영상을 사용하여 실험하였다. 그 중 일부의 데이터는 표 1 과 같다.

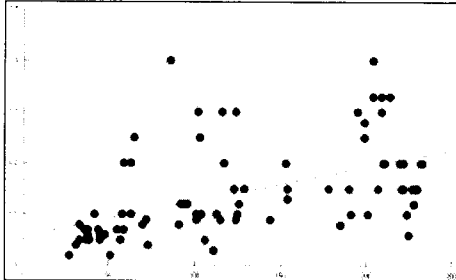
영상 번호	초점	잡영	최적 k
1	94	9	0.12
2	123	11	0.15
3	72	11	0.09
4	154	14	0.13

(표 1) 최적 이진화 지수 및 초점, 잡영 측정 에

통계 분석을 위해서는 통계 소프트웨어인 dbStat 를 이용하였다[6]. 먼저 초점과 최적 이진화 지수의 상관 관계를 분석하기 위하여 다음과 같은 수식을 이용하여 회귀 분석을 수행하였다.

$$k = a_f * focus + b_f$$

85 개의 영상에 대하여 분석한 결과 초점과 최적 이진화 지수의 관계는 그림 4 와 같이 $\alpha=0.05$ 일 때 유의한 상관관계가 있는 것으로 나타났으며 이 때 $a_f = 0.00067, b_f = 0.058$ 로 분석되었다.

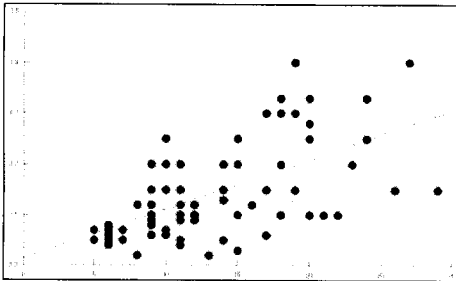


(그림 4) 초점과 최적 이진화 지수의 상관 관계

다음으로는 잡영과 최적 이진화 지수의 상관 관계를 분석하였다. 이번에도 역시 다음과 수식을 이용하여 회귀 분석을 수행하였다.

$$k = a_n * noise + b_n$$

이번에도 그림 5 와 같이 $\alpha=0.05$ 일 때 유의한 상관 관계가 있는 것으로 나타났다. $a_n = 0.0098, b_n = 0.014$ 으 로 분석되었다.



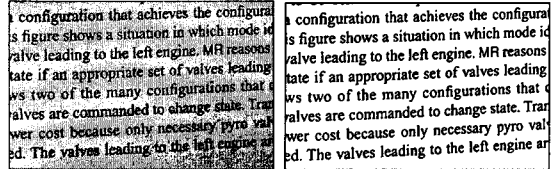
(그림 5) 잡영과 최적 이진화 지수의 상관 관계

초점과 잡영이 각각 이진화 지수와 양의 상관 관계를 갖는다는 것을 확인한 후 초점과 잡영 모두에 대한 이진화 지수를 추정하기 위하여 식 (1)에 대한 다변량 회귀분석을 실시하였다. 그 결과 $a = 0.00006, b = 0.0056, c = 0.0067$ 의 값으로 결정되었다.

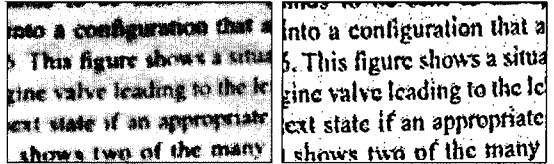
6.2. 초점 및 잡영 수치를 반영한 이진화 결과

초점이 다른 다수의 영상에 테스트해 본 결과 초점 및 잡영의 상태에 따라 이진화 지수가 적절히 결정되는 것을 확인할 수 있었다. 예를 들어 그림 6(a)과 같이 초점이 정확한 영상에 대해서는 비교적 큰 이진화

지수(0.15)가 적용되었으나, 그림 6(b)와 같이 초점이 흐린 영상에 대해서는 작은 값(0.07)이 적용되어 획성분을 섬세하게 추출함으로써 화질 변화에 대한 적응력을 가지고 있음을 확인할 수 있었다.



(a) 초점이 정확한 영상 (k=0.15)



(b) 초점이 흐린 영상 (k=0.07)

(그림 6) 초점이 다른 영상에 대한 적응 예

7. 결론

본 논문에서는 화질의 변이가 다양한 카메라 영상의 문서 인식을 위한 적응적 이진화 방법을 제안하였다. 먼저, 입력 영상의 초점과 잡영의 정도를 측정하여 수치화 하는 방법을 제안한 후 입력 영상의 화질과 그에 적합한 이진화 파라미터의 상관 관계를 통계적으로 분석하였다. 그 결과를 이용하여 입력 영상의 화질에 따라서 이진화 파라미터를 자동으로 추정함으로써 다양한 화질 변화에 효과적으로 적응할 수 있는 이진화 방법을 제안하였다.

실험을 통해 제안한 화질 측정 수치와 적절한 이진화 파라미터 간에 통계적 상관 관계가 있음을 보였으며 그 결과를 이진화에 적용시켰을 때 화질 변화에 대한 적응력이 강화됨을 확인하였다.

참고문헌

- [1] N. Otsu, A threshold selection method from grey level histogram, IEEE Trans. SMC-9, pp. 62-66, 1979
- [2] W. Niblack, An Introduction to Image Processing, Prentice-Hall, Englewood Cliffs, NJ, 1986, pp.115-116
- [3] O. D. Trier, Goal-directed evaluation of binarization methods, IEEE PAMI Vol. 17, No. 12, 1995.
- [4] J. Sauvola and M. Pietikainen, Adaptive document image binarization, Pattern Recognition, Vol. 33, pp. 225-236, 2000.
- [5] I. J. Kim, "Multi-Window Binarization of Camera Image for Document Recognition," Proceedings of 9th International Workshop on Frontiers in Handwriting Recognition, pp. 323-327, Tokyo, Japan, Oct. 26-29, 2004.
- [6] dbSTAT, <http://www.dbstat.com>