

Water flow model 을 이용한 문서영상 이진화의 속도 개선

오현화*, 이재용*, 김두식**, 장승익**, 임길택**, 진성일*
*경북대학교 전자전기공학부
**한국전자통신연구원 우정기술연구센터

A Speed-up method of document image binarization using water flow model

Hyun-Hwa Oh*, Jae-Yong Lee*, Doo-Sik Kim**, Seung-Ick Jang**,
Kil-Taek Lim**, Sung-Il Chien*

* School of Electronic and Electrical Engineering, Kyungpook National University,
**Postal Technology Research Center, ETRI
E-mail: ohh@palgong.knu.ac.kr

Abstract

This paper proposes a method to speed up the document image binarization using a water flow model. The proposed method extracts the region of interest (ROI) around characters from a document image and restricts pouring water onto a 3-dimensional terrain surface of an image only within the ROI. The amount of water to be filled into a local valley is determined automatically depending on its depth and slope. Then, the proposed method accumulates weighted water not only on the locally lowest position but also on its neighbors. Finally, the depth of each pond is adaptively thresholded for robust character segmentation. Experimental results on real document images shows that the proposed method has attained good binarization performance as well as remarkably reduced processing time compared with that of the existing method based on a water flow model.

I. 서론

영상 세그멘테이션(image segmentation)은 영상을 몇 개의 부분으로 분할하거나 배경으로부터 개체(object)를 분리하는 방법으로서 문서영상을 해석하고 문자를 인식하기 위하여 반드시 선행되어야 하는 매우 중요한 전처리 단계이다[1]. 이상적인 문서영상의 그레이레벨 히스토그램은 문자와 배경에 해당하는 두 최고치(peak)가 나타나는 바이모달 형태이다. 그러므로 전역 이진화 방법[2, 3]을 이용하여 결정된 하나의 문턱치를 적용함으로써 만족스러운 품질의 이진 문서영상을 획득할 수 있다. 그러나 실제 문서영상에서는 다양한 형태와 종류

의 잡영이 존재하므로 국부적인 특징에 기반한 이진화 방법[4, 5]이 적용되어야 한다.

Water flow model 을 이용한 이진화 방법은 영상의 밝기 값을 3 차원의 영상지형으로 간주하고 물이 낮은 곳으로 흐르는 개념을 적용하여 국부 계곡에 해당하는 문자를 분리한다[4]. 이 방법은 영상 전체에 물을 붓는 과정을 여러 번 반복하여야 하므로 수행속도가 매우 느리다. 본 논문에서는 영상지형에서 문자 주위의 관심영역(region of interest: ROI)을 추출하여 물을 붓는 영역을 제한한다. 그리고 계곡의 경사와 깊이를 이용하여 자동으로 결정된 일정량의 물을 계곡의 최저 지점과 그 주위 지점에 가중치를 부여하여 누적한다. 따라서 영상지형에 물을 붓는 과정을 한번만 수행하더라도 충분한 양의 물이 계곡에 누적된다. 최종적으로 다양한 잡영이 존재하는 문서영상에서 문자를 견실하게 추출하기 위하여 계곡에 형성된 연못의 특징을 이용하여 연못마다 임계치를 적응적으로 결정한다.

II. Water flow model 을 이용한 문서영상 이진화 방법

Water flow model 에 기반한 문서영상 이진화 방법은 3 차원으로 간주된 영상지형에 떨어진 물이 낮은 지점으로 흐르는 개념에 착안하여 제안되었다[4]. 이 방법은 크게 enhancement 과정과 thresholding 과정으로 구성되어 있다. Enhancement 과정에서는 영상지형에 떨어진 물이 주위의 고도가 낮은 계곡으로 흘러 들어가서 누적되어

연못(pond)을 형성한다. Thresholding 과정에서는 연못의 물 양을 기준으로 하여 전역 이진화 방법의 하나인 Otsu 방법을 적용함으로써 최종적으로 이진 문서영상을 획득한다. 이 방법은 국부 적응적 특징을 가지므로 배경의 밝기 변화가 다양한 문서영상에서 효과적인 이진화 성능을 나타낸다. 그러나 영상의 크기가 $N \times M$ 이고 탐색 마스크의 평균 수행 횟수를 D , 마스크의 크기를 $m \times m$, 물이 떨어지는 횟수를 w 라고 한다면 총 연산량은 $O(MN(m^2D+w))$ 가 되므로 수행속도가 매우 느리다.

III. 제안한 방법

본 절에서는 전술한 water flow model 기반 이진화 방법의 속도를 개선하기 위하여 제안한 방법을 기술한다. 그림 1의 개요도에서 회색 블럭 부분이 본 논문에서 개선하거나 추가한 과정에 해당된다.

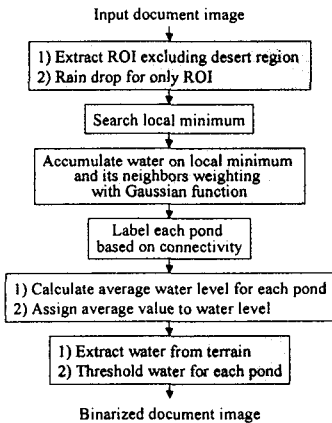


그림 1 제안한 방법에 대한 개요도

3.1 관심영역 추출

일반적인 문서영상에서 문자가 차지하는 흑화소 비는 약 30%미만이다. 본 논문에서는 영상지형에 물을 붓는 과정을 문자 주위영역으로 제한하여 수행속도를 향상시키는 방법을 제안한다. 그레이레벨의 문서영상에서 그래디언트 크기는 문자와 배경의 경계부분에서 크며 배경부분에서는 상대적으로 작은 특징을 나타낸다. 그러므로 그래디언트 크기가 큰 영역을 ROI 로 간주하고 그래디언트 크기가 작은 배경영역을 물이 뿌려질 가능성이 매우 희박한 사막(desert)으로 정의한다. 그래디언트 크기가 식 (1)의 임계치보다 작은 화소는 사막의 일

부로 간주하여 물을 붓는 과정을 수행하지 않고, 임계치 이상인 화소에만 물을 뿌리는 과정을 수행한다.

$$Th_{gradient} = 10 \log_{10} (\|\nabla I\|_{avg}) \quad (1)$$

여기서 $\|\nabla I\|_{avg}$ 는 입력 영상의 평균 그래디언트 크기이다. ROI 에 떨어진 물은 $m \times m$ 의 탐색창을 이용한 국부 최저지점 탐색과정을 거쳐 계곡으로 흘러 들어가서 누적된다.

3.2 계곡에 누적되는 물의 양 결정

전술한 바와 같이 ROI 내에만 물을 붓는 과정을 수행하면 계곡에 누적되는 물의 양이 전술한 water flow model 기반의 방법보다 상대적으로 적게 된다. 본 논문에서는 계곡의 높이와 경사를 이용하여 계곡에 누적되는 물의 양을 자동으로 결정한다. 즉, ROI 내의 한 지점에 물이 떨어지면 그 지점으로부터 탐색된 국부 최저지점까지의 거리 d_i 와 높이 차 h_i 을 이용하여 계곡의 경사 S_i 을 계산한다. 계곡에 누적되는 물의 양을 식 (2) 를 이용하여 산출한다.

$$S_i = h_i / d_i \quad (2)$$

$$w_i = 0.2h_i \times \log_{10}(1 + aS_i) / \log_{10}(256a)$$

여기서 a 는 누적되는 물의 양을 조절하는 상수이며 실험에서는 2 로 고정하였다. 이 과정에서 국부 최저지점의 주위 7×7 영역에 $\sigma = 4$ 인 Gaussian 형태의 가중치를 부여하여 물을 누적한다. 이때 주위 영역 내에서도 물이 누적될 수 있는 지점을 국부 최저지점과의 높이 차이가 $0.25w_i$ 이하인 곳으로 제한하여 주위의 spot noise 에 해당하는 계곡에 물이 채워지는 현상을 방지한다. 최종적으로 ROI 내의 모든 화소에 물을 붓는 과정을 한번만 수행하더라도 충분한 양의 물이 계곡에 누적되게 된다. 크기가 $N \times M$ 인 영상에서 사막을 제외한 ROI 내의 화소 수가 P 개이고 탐색 마스크의 평균 수행 횟수를 D' , 마스크 크기를 $m \times m$ 이라고 한다면 제안한 방법의 총 연산량은 $O(Pm^2D')$ 가 된다.

3.3 임계치의 적응적 결정

문자와 배경의 밝기 값의 차이에 따라서 연못의 깊이가 크게 다르므로 Otsu 방법으로 단일 임계치를 결정하여 thresholding 하면 어두운 배경에 형성된 얇은 연못의 문자는 제거되는 문제가 발생한다. 본 논문에서는 이러한 문제를 해결하기 위하여 계곡에 형성된 연못의 특징을 이용하여 연못마다의 임계치를 적응적으로 결정

한다. 먼저 깊이가 매우 얇은 연못은 잡영에 해당하므로 j 번째 연못의 평균 깊이 μ_j 와 영상지형에 형성된 모든 연못의 평균 깊이 μ_g 를 비교하여 $\mu_j \leq 0.2\mu_g$ 인 연못을 일차적으로 제거한다. 그리고 남은 i 번째 연못의 임계치 Th_i 를 식 (3)와 같이 결정한다.

$$Th_i = \frac{\mu_i}{2} + K_i$$

$$K_i = D_i \exp\left(-\frac{\mu_i}{\mu_g}\right) \times \exp\left(-\frac{\sigma_i}{(D_i/2)(R/c)^2}\right) \quad (3)$$

여기서 D_i 와 μ_i , σ_i 는 i 번째 연못의 최고깊이와 평균 깊이, 표준편차를 각각 나타낸다. Th_i 는 i 번째 연못 평균 깊이의 1/2 을 기준으로 하여 K_i 에 의하여 상향 조정된다. K_i 의 첫번째 항 $D_i e^{-\mu_i/\mu_g}$ 은 Th_i 가 기준 임계치 $\mu_i/2$ 에서 μ_i 값에 의하여 상향 조정되도록 하는 역할을 한다. 그리고 K_i 의 두 번째 항 $e^{-\sigma_i/(D_i/2)(R/c)^2}$ 은 σ_i 가 클수록 첫번째 항에 의한 임계치 상향조정 폭을 작게 재조정하는 역할을 한다. R 은 영상지형에 형성된 연못들의 넓이 평균에 대한 최고 넓이의 비를 나타낸다. 문자 주위에 존재하는 잡영의 면적이 넓을수록 넓은 면적의 연못이 형성되므로 R 값이 커지게 된다. c 는 상수로서 실험에서는 200으로 고정하였다.

IV. 실험 및 고찰

본 논문에서 제안한 문서영상 이진화 방법의 성능을 평가하기 위하여 Pentium IV 1.6GHz PC, Windows 2000 OS에서 실험을 수행하였다. 실험에서는 상용 스캐너를 이용하여 200dpi로 획득된 256 그레이레벨의 문서영상과 water flow model을 이용하여 제안된 이진화 방법의 실험에서 사용된 영상[5]을 이용하였다. 그림 2에 사용된 256x256 크기의 문서영상 예를 나타내었다.

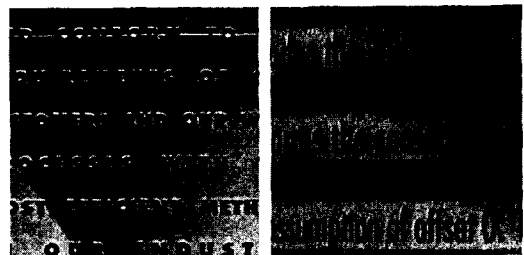
제안한 방법으로 결정된 이진화 임계치를 평가하기 위하여 물이 채워진 영상지형의 단면과 임계치를 조사하였다. 그림 3(a)는 실험영상 1의 28 번째 수평라인에 대한 영상지형 단면과 국부 계곡에 물이 채워진 결과를 나타낸다. 제안한 방법으로 물을 채운 결과 한번의 누적과정만으로도 계곡에 충분한 양의 물이 채워졌음을 확인할 수 있다. 그림 3(b)에서는 제안한 방법을 적용함으로써 연못마다 임계치가 적응적으로 결정됨을 확인할 수 있다.

제안한 방법의 이진화 성능을 평가하기 위하여 Otsu 방법, Liu와 Srihari 방법, water flow model 기반 방법의

결과와 비교하였다. 그림 4은 배경의 밝기 값의 변화가 매우 큰 실험영상 1에 대한 이진화 결과를 나타낸다. Otsu 방법과 Liu와 Srihari 방법을 적용한 결과에서 좌측 상단부에 위치한 문자들이 배경으로부터 올바르게 분리되지 못하였다. 그러나 Water flow model 기반의 이진화 방법과 제안한 방법을 적용한 결과에서는 우수한 이진화 성능을 나타내었다.

그림 5는 배경의 밝기 값과 문자의 밝기 값이 서로 비슷한 실험영상 2에 대한 이진화 실험 결과를 나타낸다. Otsu 방법을 적용한 결과에서 어두운 배경의 문자가 배경으로부터 분리되지 못함을 확인할 수 있다. Liu와 Srihari 방법을 적용한 결과에서는 배경과 문자가 분리되는 되었으나 문자 획의 굵기가 균일하지 못한 결과를 나타내었다. Water flow model 기반의 이진화 결과에서도 문자 획의 굵기가 배경의 밝기에 따라서 균일하지 않은 결과를 나타내었다. 제안한 방법의 결과에서는 문자 획의 굵기가 균일한 우수한 이진화 성능을 나타냄을 그림 5(d)에서 확인할 수 있다.

제안한 방법의 수행시간은 water flow model에 기반한 이진화 방법과 비교하여 월등히 향상되었음을 표 1에서 확인할 수 있다.



(a) 실험영상 1 (b) 실험영상 2

그림 2. 실험에 사용된 실제 문서영상의 예

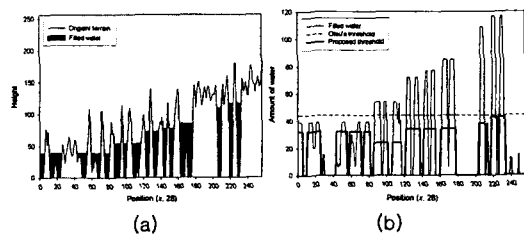
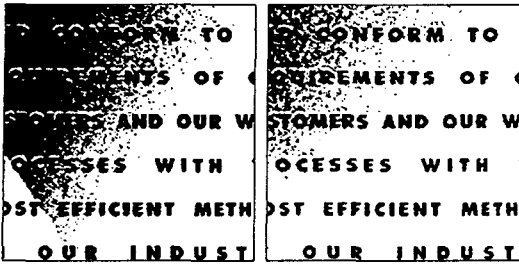
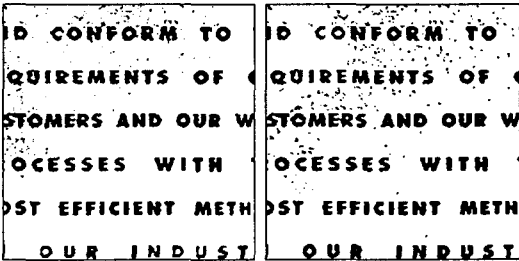


그림 3. (a) 실험영상 1의 28 번째 수평라인에 대한 단면과 물이 채워진 단면, (b) 연못의 물 양에 대한 임계치

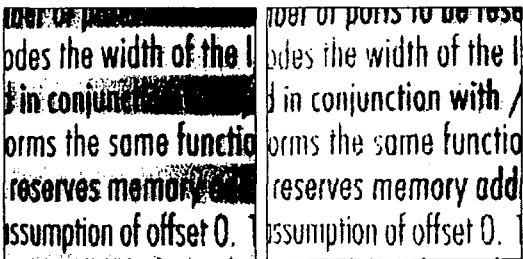


(a) Otsu의 방법 (b) Liu와 Srihari 방법

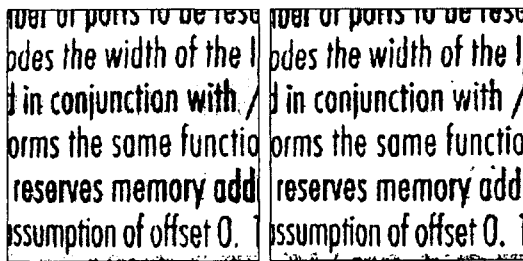


(c) water flow 기반 방법 (d) 제안한 방법

그림 4. 실험영상 1에 대한 이진화 결과



(a) Otsu의 방법 (b) Liu와 Srihari 방법



(c) water flow 기반 방법 (d) 제안한 방법

그림 5. 실험영상 2에 대한 이진화 결과

V. 결론

본 논문에서는 water flow model에 기반한 이진화 방법의 속도를 향상시키기 위하여 문자 주위의 ROI를 추출하고 자동으로 결정된 일정량의 물을 국부 계곡에 누적

표 1. 실험영상과 이진화 방법에 따른 수행시간 비교

이진화 방법	수행시간 (sec)	
	실험영상 1	실험영상 2
Otsu의 방법	0.001	0.001
Liu와 Srihari 방법	0.016	0.001
Water flow model에 기반한 방법	9.33	3.75
제안한 방법	0.656	0.652

하는 방법을 제안하였다. 실험결과에서 기존의 water flow model에 기반한 방법과 비교하여 수행속도가 월등히 향상되었을 뿐만 아니라 이진화 품질도 우수함을 확인하였다. 그러므로 제안한 방법은 문자를 자동으로 인식하여야 하는 실제 응용에서 효과적인 전처리 과정으로써 사용될 수 있을 것으로 사료된다.

참고문헌

- [1] H. Cheng and Z. Fan, "Background Identification Based Segmentation and Multilayer Three Representation of Document Images," in *Proc. of Int'l Conf. on Image Processing*, vol. 3, pp. 1005-1008, 2002.
- [2] N. Otsu, "A Threshold Selection Method from Gray-Scale Histogram," *IEEE Trans. Systems, Man, and Cybernetics*, vol. SMC-8, pp. 62-66, 1978.
- [3] J.S. Weszka and A. Rosenfeld, "Histogram Modification for Threshold Selection," *IEEE Trans. Systems, Man, and Cybernetics*, vol. SMC-9, no. 1, pp. 38-52, 1979.
- [4] I.-K. Kim, D.-W. Jung, and R.-H. Park, "Document Image Binarization Based on Topographic Analysis Using a Water Flow Model," *Pattern Recognition*, vol. 35, pp. 265-277, 2002.
- [5] Y. Liu and S.N. Srihari, "Document Image Binarization Based on Texture features," *IEEE Trans. on Pattern Recognition and Machine Intelligence*, vol. 19, no. 5, pp. 540-544, 1997.
- [6] 박성호, 장대근, 황찬식, "1차원 메디안 필터 기반 문서영상 영역해석," *전자공학회논문지*, 제 40권 SP편 제 3호, pp. 10-18, 2003.