

논문 2004-41SP-4-10

Water flow model에 기반한 문서영상 이진화 방법의 속도 개선

(Speed-up of Document Image Binarization Method Based on Water Flow Model)

오 현 화*, 김 도 훈**, 이 재 용***, 김 두 식****, 임 길 택*****, 진 성 일*

(Hyun-Hwa Oh, Do-Hun Kim, Jae-Yong Lee, Doo-Sik Kim, Kil-Taek Lim, and Sung-Il Chien)

요 약

본 논문에서는 water flow model의 개념을 적용한 문서영상 이진화 방법의 속도를 개선하는 방법을 제안한다. 제안한 방법은 문서영상에서 문자 주위를 관심영역(region of interest: ROI)으로 추출하고 3차원 영상지형에서 물이 뿌려지는 영역을 관심영역 이내로 제한한다. 국부 계곡에 누적되는 물의 양은 계곡의 깊이와 경사를 이용하여 자동으로 결정된다. 그리고 계곡의 최저 지점뿐만 아니라 그 주위에도 가중치를 부여하여 물을 누적함으로써 관심영역에 해당하는 영상지형에 물을 붓는 과정을 한번만 수행하여 충분한 양의 물이 계곡에 채워지도록 한다. 계곡에 형성된 연못의 깊이는 배경과 문자의 밝기 차에 따라 다양하므로 연못의 깊이를 기준으로 문자 분리를 위한 임계치를 적응적으로 결정한다. 실제 문서영상에 대한 실험에서 제안한 방법의 수행속도가 water flow model에 기반한 이진화 방법과 비교하여 월등히 향상되었으며 이진화 품질도 매우 우수함을 보였다.

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. The proposed method accumulates weighted water not only on the locally lowest position but also on its neighbors. Therefore, a valley is filled enough with only one try of pouring water onto the terrain surface of the ROI. Finally, the depth of each pond is adaptively thresholded for robust character segmentation, because the depth of a pond formed at a valley varies widely according to the gray-level difference between characters and backgrounds. In our experiments on real document images, 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.

Keywords : 문서영상, 이진화, water flow model, 국부적응, 속도개선

I. 서 론

영상 세그멘테이션(image segmentation)은 영상을 몇 개의 부분으로 분할하거나 배경으로부터 개체(object)

를 분리하는 방법으로서 문서영상을 해석하고 문자를 인식하기 위하여 수행되어야 하는 매우 중요한 전처리 단계이다^[1-3]. 특히 문자인식 성능은 이진화된 문자의 품질에 큰 영향을 받으므로 효과적이고 안정적인 문서영상 이진화가 요구된다.

이상적인 문서영상의 그레이레벨 히스토그램은 문자와 배경에 해당하는 두 최고치(peak)가 나타나는 바이모달 형태이다. 전역 이진화 방법^[4-6]을 이용하면 두 최고치 사이의 계곡(valley)에서 하나의 임계치가 결정되어 만족스러운 품질의 이진 문서영상을 획득할 수 있다. 그러나 스캐너와 카메라 등으로 획득된 실제 문서영상에서는 조명과 입력장치의 특성 등 영상 획득 조건

* 정회원, ** 학생회원, 경북대학교 전자전기공학부 (School of Electronic and Electrical Engineering, Kyungpook National University)

*** 학생회원, 삼성 SDI PDP본부 (Samsung SDI PDP Division)

**** 정회원, 한국전자통신연구원우정기술연구센터 (Postal Technology Research Center, ETRI)

***** 정회원, 경주대학교 컴퓨터멀티미디어공학부

※ 본 연구는 BK21연구지원과 한국전자통신연구원 우정기술연구센터의 위탁연구과제지원으로 수행되었음.

접수일자: 2003년10월15일, 수정완료일: 2004년4월31일

에 따라서 다양한 배경과 잡영이 나타나므로 단일 임계치를 적용한 이진화 결과에서는 품질이 크게 저하되는 경향이 있다. 그러므로 잡영에 의하여 심하게 손상되거나 문자 주위에 다양한 배경이 존재하는 문서영상을 건설하게 이진화하기 위하여 영상의 국부적 특성에 기반한 이진화 방법이 제안되었다^{[3][7-9][11]}. 국부적 이진화 방법의 하나인 water flow model에 기반한 방법^[8]은 영상의 밝기 값을 3차원의 영상지형으로 간주하고 물이 낮은 곳으로 흐르는 개념을 적용하여 국부 계곡에 해당하는 문자를 효과적으로 분리한다. 그러나 이 방법에서는 국부 계곡에 누적되는 물의 양을 자동으로 결정하기 어려우며 수행시간이 긴 단점이 있다. 또한 계곡에 쌓인 물의 양을 기준으로 전역 이진화 방법을 적용하므로 얇은 연못에 해당하는 문자가 올바르게 분리되지 못하는 경우가 발생한다.

본 논문에서는 영상지형의 평탄한 영역을 사막(desert)영역으로 간주하고 이 영역을 제외한 문자 주위의 관심영역에만 물 붓는 과정을 수행함으로써 water flow model에 기반한 이진화 방법의 느린 수행속도를 개선한다. 또한 영상지형의 국부 계곡에 누적되는 물의 양은 계곡의 경사와 깊이를 이용하여 자동으로 결정한다. 그리고 계곡의 최저 지점뿐만 아니라 그 주위 지점에도 가중치를 부여하여 물을 누적함으로써 영상지형에 물을 붓는 과정을 한번만 수행하더라도 충분한 양의 물이 계곡에 누적되도록 한다. 계곡에 형성된 연못의 깊이는 문자 주위 배경의 밝기와 잡영의 존재 여부에 따라서 큰 차이를 나타낸다. 연못의 깊이를 기준으로 전역 이진화를 수행할 경우 얇은 연못의 문자가 유실될 가능성이 있으므로 제안한 방법에서는 연못마다의 특징을 이용하여 최종 임계치를 적응적으로 결정한다.

본 논문의 구성은 다음과 같다. 제 II장에서는 water flow model에 기반한 문서영상 이진화 방법을 간략히 소개하고 개선되어야 할 사항을 기술한다. 제 III장에서는 water flow model에 기반한 이진화 방법의 속도를 개선하는 방법과 안정적인 이진화 품질을 획득하기 위하여 제안한 방법을 자세히 기술한다. 제 IV장에서는 다양한 실제 문서영상에 제안한 방법을 적용한 이진화 결과와 기존 방법들의 결과를 비교하여 성능과 속도를 평가하며, 제 V장에서 결론을 맺도록 한다.

II. Water flow model을 이용한 문서영상 이진화

Water flow model은 물이 항상 낮은 지점으로 흐르

는 특성에 착안하여 제안되었다. 일반적으로 그레이레벨 문서영상에서 문자에 해당하는 밝기 값은 낮고 배경의 밝기 값은 높다. 그러므로 문서영상을 3차원 지형으로 간주하면 문자는 계곡에 해당하며 배경부분은 평지나 산, 고원에 해당한다. Water flow model에 기반한 문서영상 이진화 방법은 크게 enhancement 과정과 thresholding 과정으로 구성되어 있다^[8]. 먼저 enhance-ment 과정에서는 3차원 지형으로 간주한 영상 표면에 물을 떨어뜨린다. 물은 주위의 고도가 낮은 지점으로 흐르므로 문자에 해당하는 계곡으로 흘러 들어가서 누적된다. 따라서 국부 계곡마다 물이 누적되어 연못(pond)이 형성된다. Thresholding 과정에서는 연못에 누적된 물의 양을 기준으로 전역 이진화 방법의 하나인 Otsu 방법을 적용하여 최종적으로 이진 문서영상을 획득한다. Water flow model 기반의 이진화는 국부 적응 방법으로서 배경의 밝기 변화가 다양한 문서영상에서 효과적인 이진화 성능을 나타낸다. 그러나 영상의 크기가 $M \times N$ 이고 탐색 마스크의 평균 수행 횟수를 D , 마스크의 크기를 $m \times m$, 전체 영상지형에 물을 붓는 횟수를 w 라고 한다면 총 연산량은 $(\alpha MN(m^2D+w))$ 가 되므로 전역 이진화 방법과 비교하여 수행속도가 매우 느리다. 또한 국부 계곡에 채워진 물의 양에 전역 이진화 방법을 적용하므로 문자에 해당하는 계곡임에도 불구하고 연못의 깊이가 얇으면 thresholding 과정에서 제거되는 문제가 발생한다. 그러므로 모든 계곡에 충분한 양의 물이 채워질 수 있도록 enhancement 과정의 반복 횟수를 증가시켜야 하며 결과적으로 수행속도는 더욱 느려지게 된다. 이 과정에서 영상지형 표면에 붓는 물의 양은 입력영상마다 실험적으로 결정되므로 문자를 자동으로 인식하여야 하는 응용 시스템에 직접 적용하는 데에는 어려움이 있다.

III. Water flow model기반 이진화 방법의 속도 및 성능 개선

본 장에서는 전술한 water flow model기반 이진화 방법에서 나타나는 단점을 개선하기 위하여 제안한 방법에 대하여 서술한다. 그림 1에 전체적인 개요도를 나타내었다. 회색 블록으로 표현된 과정이 속도 개선, 계곡에 누적되는 물의 양 자동 결정 및 연못 깊이를 기준으로 전역 이진화 방법을 적용함으로써 발생하는 문자 유실 문제를 보완하기 위하여 개선하거나 추가한 부분에 해당한다.

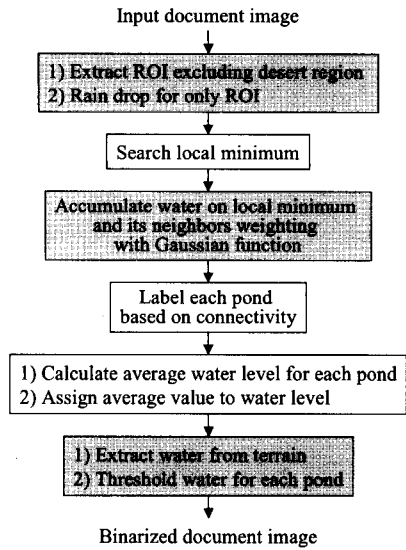


그림 1. Water flow model기반의 문서영상 이진화 방법의 속도 개선을 위하여 제안된 방법에 대한 개요도

Fig. 1. Overall flow chart of proposed method to speed up document binarization based on water flow model.

1. 문자 주위의 관심영역 추출

제 II장에서 서술한 water flow model기반의 문서영상 이진화 방법에서는 영상지형의 전체 표면에 물을 붓는 과정을 여러 번 반복 수행하므로 속도가 느린 단점이 있다. 만약 영상지형의 비교적 평탄하고 넓은 고원에 물이 떨어지면 국부 최저 지점으로 흘러 들어가는 경로가 길기 때문에 수행 시간이 더욱 길어지게 된다. 그러므로 평탄한 지형에 해당하는 배경영역은 물 붓는 과정에서 배제시킴으로써 수행시간을 단축시킬 수 있다. 일반적인 그레이레벨 문서영상에서 문자가 차지하는 흑화소 비는 약 30%미만^[10]이며 그레디언트 크기는 문자와 배경의 경계부분에서 크고 배경부분에서는 상대적으로 작은 특징을 나타낸다. 본 논문에서는 그레디언트 크기가 큰 영역을 문자가 존재하는 관심영역으로 간주하고 그레디언트 크기가 작은 배경영역을 물이 뿌려질 가능성이 매우 희박한 사막으로 정의한다. 그리고 그레디언트 크기가 식 (1)의 임계치보다 작은 화소는 사막의 일부로 간주하여 물 붓는 과정을 수행하지 않고, 그레디언트 크기가 임계치 이상인 화소에만 물 붓는 과정을 수행한다.

$$Th_{gradient} = 10 \log_{10} (\nabla I_{avg}) \quad (1)$$

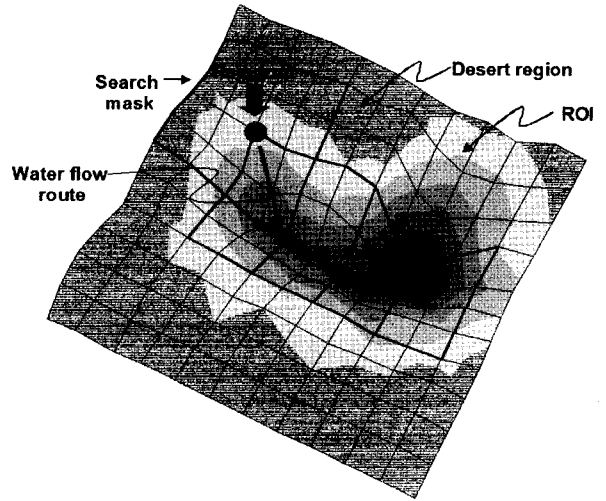


그림 2. 관심영역과 사막영역 및 탐색창을 이용한 국부 최저지점 탐색 과정

Fig. 2. ROI, desert region, and searching procedure of local minimum position using search mask.

여기서 ∇I_{avg} 는 입력 영상의 평균 그레디언트 크기를 나타낸다. 그림 2에 나타난 3차원 영상지형에서 빗금 친 영역이 사막영역에 해당된다. 사막을 제외한 관심영역에 뿌려진 물은 $m \times m$ 의 탐색창을 이용한 국부 최저지점 탐색과정을 거쳐 계곡으로 흘러 들어간다. 이 과정에서 국부 최저지점이 탐색창의 중심점과 일치하면 탐색을 멈추고 물이 바닥에 고이게 된다. 이와 같이 물이 뿌려질 영역을 관심영역 이내로 제한함으로써 수행시간이 일차적으로 단축된다. 또한 계곡 주위의 영역이 관심영역으로 추출되므로 국부 최저지점을 검출하기 위한 탐색 경로의 길이가 현저하게 짧아지므로 수행 속도가 더욱 향상되는 효과를 획득할 수 있다.

2. 계곡에 누적되는 물의 양 자동 결정

기존의 water flow model기반의 방법에서는 영상지형의 모든 지점에 물을 뿌리고 주위 고도가 최저인 지점에 물을 1 만큼씩 누적하였다. 본 논문에서는 관심영역에만 물 붓는 과정을 수행하므로 기존의 방법과 비교하여 계곡으로 흘러 들어가는 물의 양이 상대적으로 매우 적다. 그러므로 계곡에 누적되는 물의 양을 1로 제한하게 되면 계곡에 충분한 양의 물이 누적되도록 하기 위하여 물 붓는 과정의 반복횟수를 크게 증가시켜야 한다. 따라서 1절에서 기술한 관심영역을 이용한 속도향상의 효과를 기대하기 어렵다. 만약 1보다 훨씬 큰 일정한 양의 물이 누적되도록 한다면 얇은 계곡의 경우 물이 크게 넘쳐서 배경과 잡영 부분까지 많은 양의 물

이 누적되는 문제가 발생한다.

실제 지면에 떨어진 물이 주위의 저고도 지점으로 흘러가는 과정에서 지면의 경사가 급하면 현재 지점에 존재하는 물 양의 대부분이 고도가 낮은 지점으로 흘러간다. 반면 지면의 경사가 완만하다면 주위의 저고도 지점으로 흘러가는 물의 양은 상대적으로 적다. 본 논문에서는 water flow model에 위의 현상을 도입하여 영상지형의 계곡에 채워지는 물의 양을 자동으로 조정한다.

먼저 그림 3(a)의 방법에서는 영상지형 표면에 물이 떨어지면 탐색창을 이용하여 국부 최저지점을 탐색해가는 각 단계마다 물의 양이 순차적으로 계산된다. 즉 계곡 주변의 한 지점에 물이 떨어지면 국부 최저 지점으로 흘러 들어가서 누적되는 물의 양은 식 (2)에서 산출된 지면의 경사 S_{ij} 에 비례하도록 결정한다.

$$S_{ij} = h_{ij} / d_{ij} \tag{2}$$

여기서 d_{ij} 와 h_{ij} 는 i 번째 계곡으로 물이 흘러 들어가는 과정에서 j 번째 지점과 그 이전지점 간의 거리와 높이 차를 나타낸다. 이 방법을 이용하면 물의 양은 자동으로 결정할 수는 있으나 최저지점 탐색과정이 길어질수록 물의 양을 계산하는 횟수가 증가되므로 수행시간이 길어질 가능성이 있다. 그러므로 본 논문에서는 그림 3(b)와 같은 방식으로 물의 양을 결정하는 방법을 최종적으로 제안한다. 즉 관심영역 내의 한 지점에 물이 떨어지면 그 지점으로부터 탐색과정을 거쳐 최종 결정된 국부 최저지점과의 거리 d'_i 와 높이 차 h'_i 를 이용하여 두 지점 간의 경사 s'_i 를 계산한다. 그리고 계곡에 누적되는 물의 양을 식 (3)을 이용하여 한번에 산출한다.

$$S'_i = h'_i / d'_i$$

$$w' = 0.2h'_i \times \log_{10}(1 + aS'_i) / \log_{10}(256a) \tag{3}$$

여기서 a 는 누적되는 물의 양을 조절하는 상수이며 실험에서는 2로 고정하였다. i 번째 계곡의 최저지점에서 한번에 누적될 수 있는 물의 최대 양을 h'_i 의 20%로 설정하고 계곡의 경사 s'_i 에 의하여 그 양이 조절되도록 한다. 이 과정에서 계곡의 최저 지점에만 물이 집중적으로 누적되는 현상을 방지하기 위하여 최저지점의 주위 7×7 영역에 $\sigma = 4$ 인 Gaussian 형태의 가중치를 부여하여 계곡의 바닥면에 물이 고르게 누적되도록 한

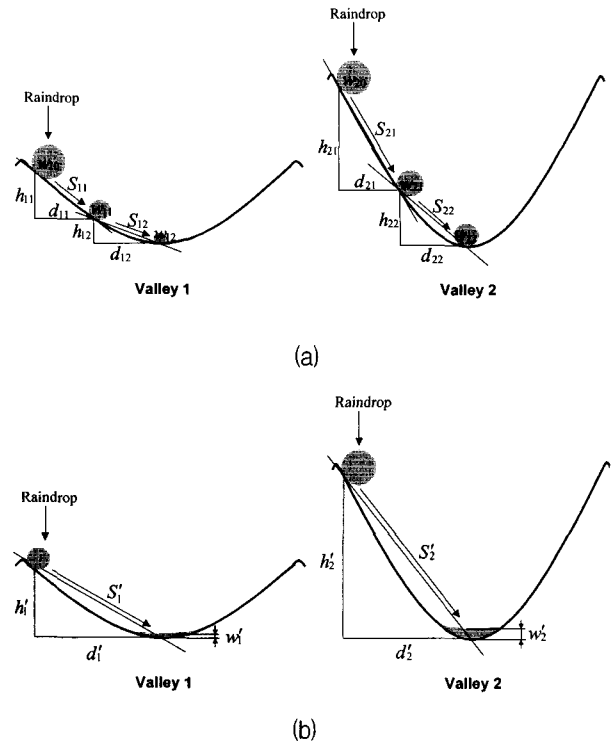


그림 3. 계곡의 깊이와 경사에 따라 누적되는 물의 양 (a) 탐색 경로 상의 모든 지점에서 물의 양을 순차적으로 계산하는 방법, (b) 물이 떨어진 지점과 국부 최저지점과의 경사를 이용하여 누적되는 물의 양을 계산하는 방법

Fig. 3. Amount of water filled into valley according to its depth and slope; (a) Calculation method which determines amount of filled water on every step of searching procedure, (b) Calculation method which determines amount of filled water by using depth and slope between rain dropped position and local minimum position.

다. 이때 주위 7×7 영역 내에서도 물이 누적될 수 있는 지점을 국부 최저지점과의 높이 차이가 $0.25w'_i$ 이 하인 곳으로 제한하여 잡음에 해당하는 주위의 작은 계곡에 물이 채워지는 현상을 방지한다. 결과적으로 제안한 방법에서는 관심영역 내의 모든 화소에 물 붓는 과정을 한번만 수행하더라도 충분한 양의 물이 계곡에 누적되게 된다. 크기가 $N \times M$ 인 영상에서 사막을 제외한 관심영역 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 되므로 water flow model을 이용한 기존의 이진화 방법과 비교하여 수행시간이 크게 단축된다.

3. 국부 임계치 결정

제 II장에서 전술한 바와 같이 water flow model을 이용한 이진화 방법에서는 계곡에 누적된 물의 양을 기준으로 Otsu 방법을 적용하여 최종적인 thresholding 결과를 획득한다. 문자의 밝기 값이 일정하더라도 밝은 배경에 존재하는 문자에 해당하는 계곡은 깊이가 깊고 경사가 급하므로 물이 많이 채워지게 된다. 반면 어두운 배경에 존재하는 문자의 경우는 계곡의 깊이가 얇고 경사가 완만하므로 적은 양의 물이 채워진다. 그림 4(a) 문서영상의 30번째 수평라인에 대하여 지형 단면과 제안한 방법으로 계곡에 물을 채운 결과를 그림 4(b)에 나타내었다. 그림 4(c)는 계곡에 형성된 연못의 수면에서 바닥면까지의 깊이를 나타낸 것으로 연못 바닥면의 지형에 해당한다. 깊은 계곡에는 많은 양의 물이 채워지므로 연못의 깊이도 깊지만 얇은 계곡에 형성된 연못의 깊이는 상대적으로 매우 얇음을 확인할 수 있다. 그러므로 Otsu 방법으로 결정된 단일 임계치를 이용하여 그림 4(c)를 thresholding하면 어두운 배경의 문자가 유실되는 문제가 발생한다.

본 논문에서는 이러한 문제를 해결하기 위하여 계곡에 형성된 연못의 특징을 이용하여 연못마다의 임계치를 적응적으로 결정하는 방법을 제안한다. 먼저 물이 채워진 지점의 연결성(connectivity)을 이용하여 연못을 레이블링한다. 깊이가 매우 얇은 연못은 잡영에 해당하므로 레이블링된 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 나머지 연못에 대하여 임계치 Th_i 를 식 (4)와 같이 결정한다.

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

$$K_i = D_i \exp\left(-\frac{\mu_i}{\mu_g}\right) \times \exp\left(-\frac{\sigma_i}{(D_i/2)(R/c)^2}\right) \quad (4)$$

여기서 D_i 와 μ_i , σ_i 는 i 번째 연못의 최고 깊이와 평균 깊이, 표준편차를 각각 나타낸다.

Th_i 는 i 번째 연못 평균 깊이의 1/2을 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. 배경과 문자의 밝기 값의 차가 크다면 형성된 연못은 그림 5의 Pond 1과 같이 깊이가 깊으므로 μ_i 값이 크다. 이 경우는 임계치가 너

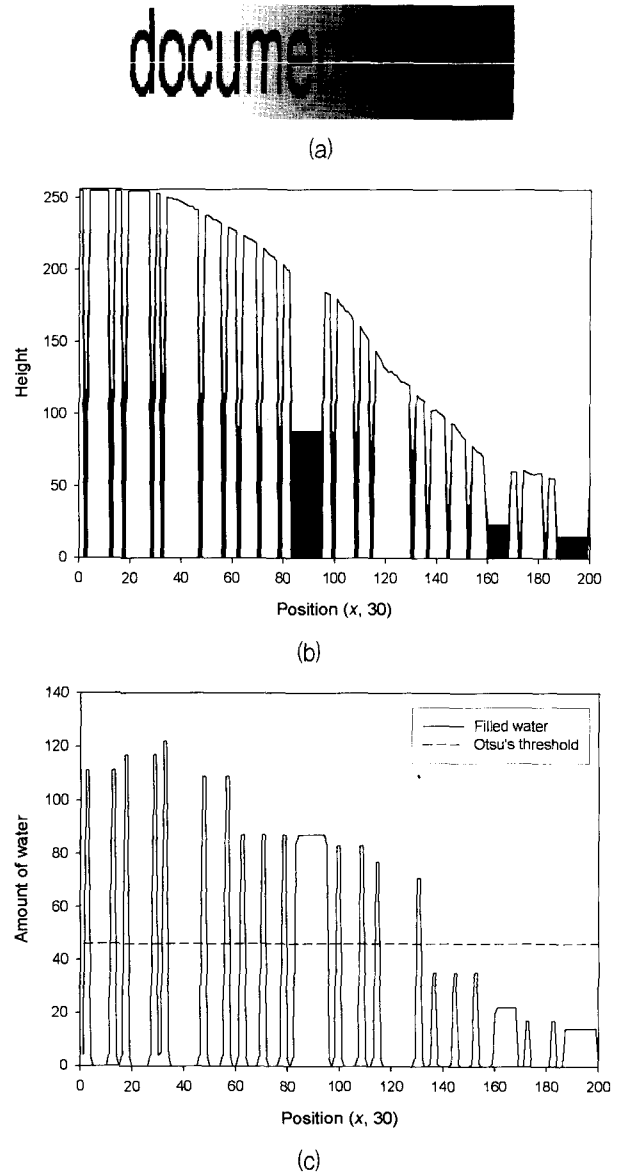


그림 4. 계곡에 채워진 물 양에 대해 Otsu 방법을 적용하여 결정된 임계치 (a) 입력영상, (b) 30번째 수평라인에 대한 단면과 물이 채워진 단면, (c) 계곡에 형성된 연못의 물 양과 Otsu 방법으로 결정된 임계치

Fig. 4. Threshold value determined by applying Otsu's method to filled water; (a) Input image, (b) Profile of original terrain and water filled terrain of 30th horizontal line of input image, (c) Amount of water which is filled into pond and Otsu's threshold.

무 높게 설정되면 이진화된 문자의 품질이 저하될 가능성이 있으므로 기준 임계치에서 상향 조정되는 정도를 상대적으로 적게 한다. 만약 문자 주위에 잡영이 심하게 존재한다면 문자에 해당하는 계곡의 깊이가 얇으므로 물이 넘쳐서 주위 잡영 부분까지 물이 채워지게 된다. 따라서 그림 5의 Pond 3과 같이 문자와 잡영이 공

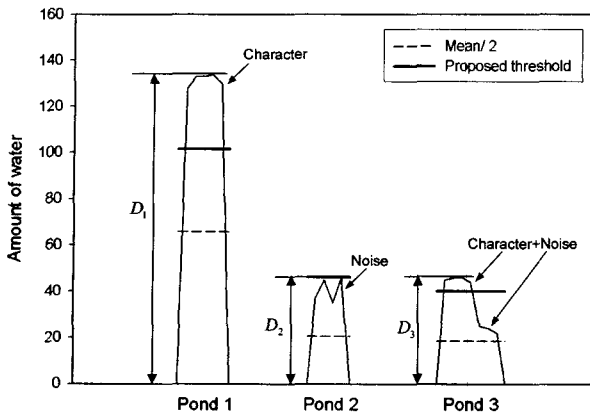


그림 5. 제안한 방법을 이용하여 다양한 연못에 적응적으로 임계치 결정

Fig. 5. Determination of adaptive threshold to various ponds by using proposed method.

존하는 연못이 형성된다. 다양한 잡영이 존재하는 영상 지형에서는 잡영에 해당하는 연못의 깊이가 문자와 잡영이 공존하는 연못의 깊이와 비슷한 경우가 발생할 수 있다. 그림 5의 Pond 2가 이 경우에 해당한다. Pond 2와 Pond 3은 μ_i 가 작으므로 K_i 의 첫번째 항에 의하여 상향 조정 정도가 크다. 그러므로 Pond 2가 제거되는 동시에 Pond 3도 제거될 가능성이 높다. 제안한 방법에서는 잡영에 해당하는 Pond 2는 제거하고 Pond 3의 잡영과 문자는 분리되도록 하기 위하여 연못의 표준편차를 이용한다. 그림 5에서 Pond 2의 표준편차가 Pond 3의 표준편차보다 적다. 식(5)에서 K_i 의 두 번째 항 $e^{-\sigma_i^2/(D_i/2)(R/\sigma)^2}$ 은 σ_i 가 클수록 첫번째 항에 의한 임계치 상향 조정 폭을 작게 재조정한다. 따라서 잡영 연못의 임계치는 연못이 제거될 수 있을 정도로 높게 결정되며 문자와 잡영이 공존하는 연못에 대해서는 잡영만 제거될 수 있도록 임계치가 적응적으로 결정된다. R 은 영상지형에 형성된 연못들 넓이의 평균에 대한 최고 넓이의 비를 나타낸다. 문자 주위에 존재하는 잡영의 면적이 넓을수록 넓은 면적의 연못이 형성될 가능성이 높으므로 R 값이 크다. 그러므로 잡영이 심한 영상에서 임계치가 비교적 높게 결정되도록 하는 역할을 한다. c 는 상수로서 실험에서는 200으로 고정하였다.

IV. 실험결과 및 고찰

본 논문에서 제안한 문서영상 이진화 방법의 성능을 평가하기 위하여 Pentium IV 1.6GHz PC, Windows 2000 OS에서 실험을 수행하였다. 실험에서는 상용 스

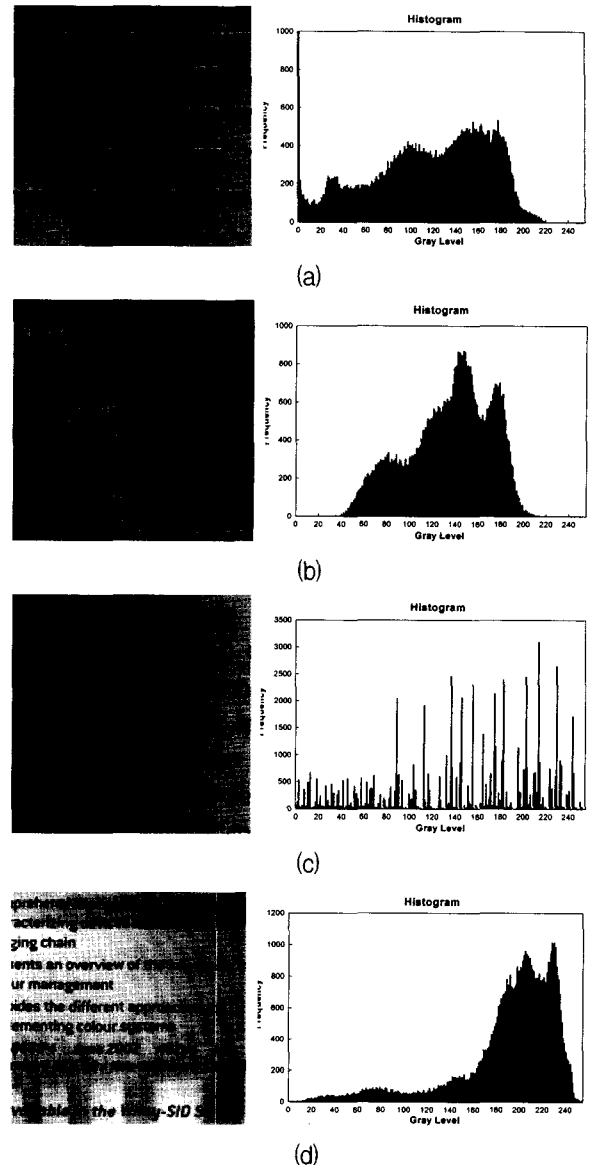


그림 6. 실험에 사용된 문서영상의 예와 히스토그램 (a) 실험영상1, (b) 실험영상2, (c) 실험영상3, (d) 실험영상4

Fig. 6. Real test images for binarization and their histograms; (a) Test image 1, (b) Test image 2, (c) Test image 3, (d) Test image 4.

캐너를 이용하여 200dpi로 획득된 256 그레이레벨 문서영상과 water flow model을 이용한 이진화 방법의 실험에서 사용된 영상^[8]을 이용하였다. 실험에 사용된 256 × 256 크기의 문서영상 예와 히스토그램을 그림 6에 나타내었다.

그림 6(a)의 실험영상1은 배경의 밝기 값 변화가 매우 크므로 히스토그램에서 밝기 분포가 넓게 나타난다. 그리고 좌측 상단부의 배경에 잡영이 심하게 존재함으로 인하여 문자를 배경으로부터 효과적으로 분리하는데 어려움이 있다. 실험영상2는 문자주위의 배경의 밝기

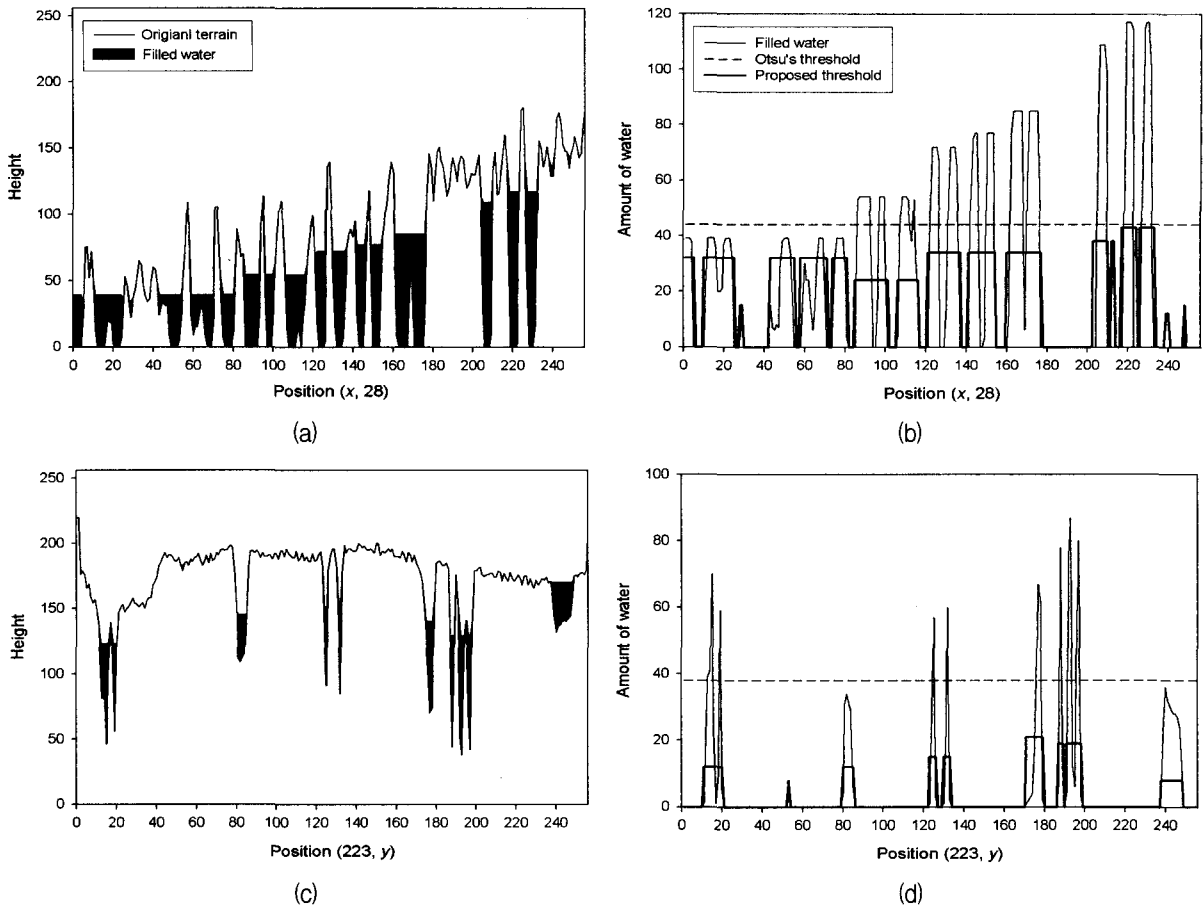


그림 7. 제안한 방법으로 계곡에 채워진 물의 양과 임계치 (a) 실험영상1의 28번째 수평라인에 대한 영상지형 단면과 물이 채워진 단면, (b) (a)의 계곡에 채워진 물에 대한 임계치, (c) 실험영상4의 223번째 수직라인에 대한 영상지형 단면과 물이 채워진 단면, (d) (c)의 계곡에 채워진 물에 대한 임계치

Fig. 7. Amount of filled water on valleys and thresholds by using proposed method: (a) Profile of original terrain and water filled terrain of 28th horizontal line of test image 1, (b) Threshold values for water filled ponds of (a), (c) Profile of original terrain and water filled terrain of 223 vertical line of test image 4, (d) Threshold values for water filled ponds of (c).

값이 교대로 변화하며 히스토그램에서 문자와 배경의 밝기가 넓게 분포한다. 그림 6(c)의 실험영상3은 스캔 과정에서 조명이 불균일하여 좌측의 배경이 어둡고 문자가 흐린 특징이 있다. 히스토그램 분포는 넓고 이산적인 특징을 나타낸다. 그림 6(d)의 실험영상4는 배경의 밝기 값이 다양할 뿐만 아니라 문자의 밝기 값과 크기가 다양한 특징을 나타낸다. 따라서 히스토그램의 분포가 매우 넓은을 확인할 수 있다.

제안한 방법으로 연못마다 적응적으로 결정된 임계치를 평가하기 위하여 물이 채워진 영상지형의 단면을 조사하였다. 물을 채우는 과정에서 국부 최저지점을 탐색하기 위한 탐색창의 크기는 7×7 로 고정하였다. 그림 7(a)는 실험영상1의 28번째 수평라인에 대한 영상지형 단면과 국부 계곡에 물이 채워진 결과를 나타내며 그림 7(b)는 계곡에 형성된 연못의 깊이에 대하여 결정

된 임계치를 나타낸다. 제안한 방법으로 물을 채운 결과 한번의 누적과정만으로도 계곡에 충분한 양의 물이 채워졌음을 알 수 있다. 그러나 문자 주위의 배경 밝기에 따라서 계곡에 형성된 연못의 깊이가 차이가 크다. 그림 7(b)와 같이 Otsu 방법으로 결정된 단일 임계치를 이용하여 연못 깊이를 thresholding하면 왼쪽 부분의 얇은 연못이 유실되는 문제가 발생한다. 반면 제안한 방법에서는 연못마다의 특성에 따라서 임계치가 독립적으로 결정되므로 얇은 연못에 대해서도 이진화 임계치가 적절히 결정됨을 확인할 수 있다. 그림 7(c)는 배경뿐만 아니라 문자의 밝기가 다양한 실험영상4의 223번째 수직라인에 대한 단면과 채워진 물을 나타낸다. 배경과 문자의 밝기 차이에 따라서 계곡에 형성된 연못의 깊이가 매우 크므로 Otsu 방법으로 결정된 임계치를 적용하면 문자 유실이 발생한다. 그러나 제안한 방법을

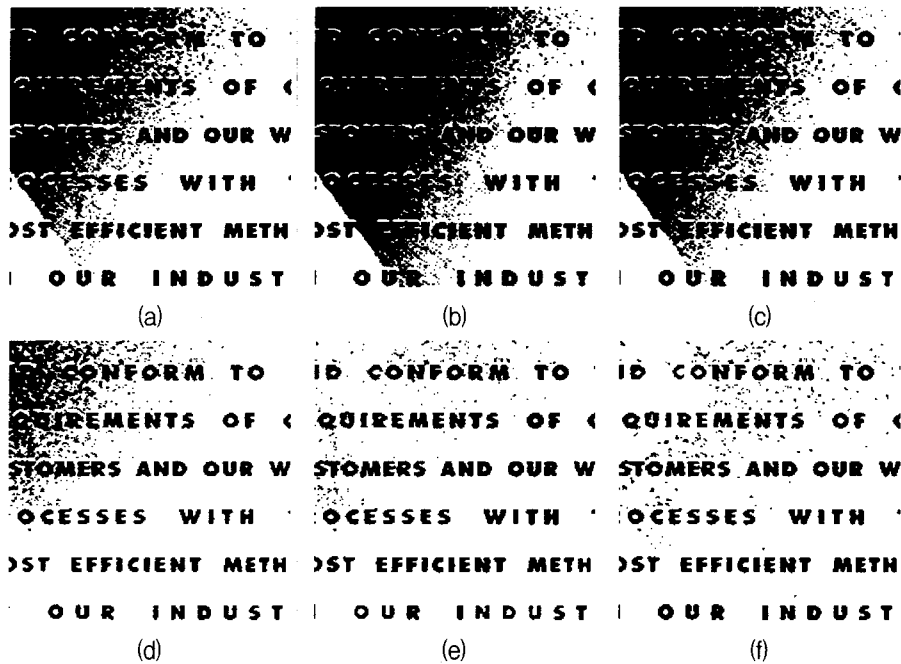


그림 8. 실험영상1에 대한 이진화 결과 (a) Otsu 방법, (b) Kapur 방법, (c) Weszka와 Rosenfeld 방법, (d) Liu와 Srihari 방법, (e) water flow model 기반 방법, (f) 제안한 방법

Fig. 8. Binarization results of test image 1; (a) Otsu's method, (b) Kapur's method, (c) Weszka and Rosenfeld's method, (d) Liu and Srihari's method, (e) Water flow model based method, (f) Proposed method.

적용하여 연못마다 독립적으로 결정된 임계치는 국부 적응적이므로 다양한 깊이의 연못을 안정적으로 thresholding할 수 있다.

제안한 방법의 이진화 성능을 평가하기 위하여 5가지의 다른 이진화 방법을 이용하였다. 먼저 기존의 water flow model기반의 이진화 방법을 모든 실험영상에 적용하였다. 실험에서 최적의 이진화 성능을 획득하기 위하여 water flow model에 기반한 이진화 방법의 탐색 마스크의 크기는 7×7 로 고정하였으며, 영상지형에 물을 뿌리는 횟수는 실험영상1과 실험영상 2에 대하여 15번, 실험영상3은 20번, 실험영상4는 30번으로 하였다. 전역 이진화 방법으로는 히스토그램의 분포 특징을 이용하여 하나의 임계치를 결정하는 Otsu 방법^[4]과 Kapur 방법^[5]을 실험영상에 적용하였다. 그리고 문서영상에서 에지 정보를 이용하여 히스토그램을 수정(modification)하고 임계치를 결정하는 Weszka와 Rosenfeld 방법^[12]을 실험영상에 적용하였다. 지역 이진화 방법으로는 문자의 특징과 이진 영상의 run-length 히스토그램으로부터 임계치를 결정하는 Liu와 Srihari 방법^[11]을 수행하였다.

그림 8은 실험영상1에 대한 이진화 결과를 나타낸다. 그림 8(a)의 Otsu 방법을 적용한 결과에서는 좌측 상단부에 위치한 문자들이 배경으로부터 올바르게 분리되지

못하였다. 이와 같은 문제는 그림 8(b)의 Kapur 방법과 그림 8(c)의 Weszka와 Rosenfeld 방법의 결과에서 더욱 심각하게 발생하였다. 그림 8(d)의 Liu와 Srihari 방법에서는 앞의 결과보다는 개선된 성능을 나타내지만 좌측 상단부에 위치한 다수의 문자가 배경과 분리되지 못하는 문제는 여전히 나타났다. 그림 8(e)에서 보는 바와 같이 water flow model 기반의 이진화 방법을 적용한 결과에서는 좌측 상단 부분의 문자가 배경으로부터 매우 효과적으로 분리되어 우수한 이진화 성능을 나타내었다. 그림 8(f)의 제안한 방법을 적용한 결과에서도 그림 8(e)의 결과와 비슷한 이진화 성능은 나타냄을 확인할 수 있다. 제안한 방법의 결과에서 잡영이 조금 더 많이 발생하였으나 문자의 품질을 크게 저하시키지는 않았다. 그러므로 잡영 제거를 위한 후처리 과정을 적용함으로써 간단하게 제거될 수 있을 것으로 판단된다. 본 논문에서는 구현된 이진화 방법들의 객관적인 평가를 위하여 잡영 제거를 위한 후처리 과정을 적용하지 않은 결과를 서로 비교하였다.

그림 9는 실험영상2에 대한 이진화 실험 결과를 나타낸다. 반복되는 배경의 밝기 값과 문자의 밝기 값이 서로 비슷하므로 전역 이진화 방법인 Otsu 방법, Kapur 방법, Weszka와 Rosenfeld 방법을 적용한 결과에서 어두운 배경의 문자가 배경으로부터 분리되지 못함을 확

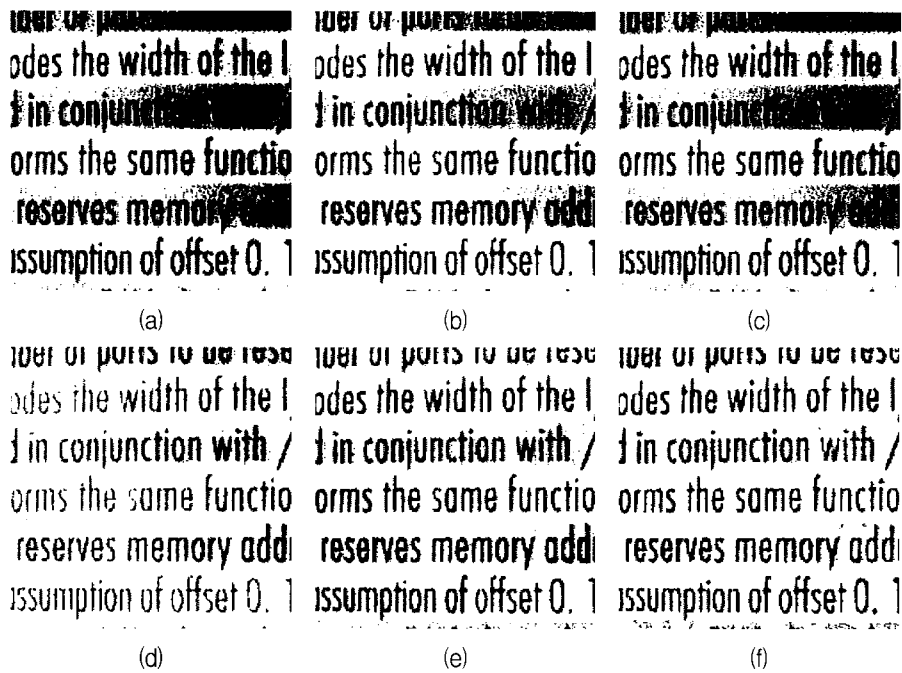


그림 9. 실험영상2에 대한 이진화 결과 (a) Otsu 방법, (b) Kapur 방법, (c) Weszka와 Rosenfeld 방법, (d) Liu와 Srihari 방법, (e) water flow model 기반 방법, (f) 제안한 방법

Fig. 9. Binarization results of test image 2; (a) Otsu's method, (b) Kapur's method, (c) Weszka and Rosenfeld's method, (d) Liu and Srihari's method, (e) Water flow model based method, (f) Proposed method.

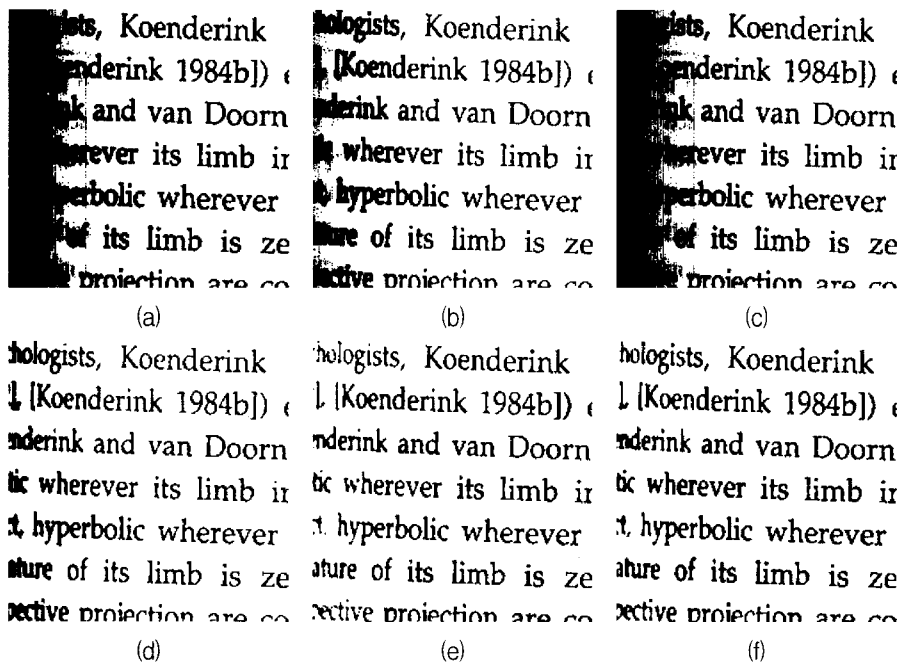


그림 10. 실험영상3에 대한 이진화 결과 (a) Otsu 방법, (b) Kapur 방법, (c) Weszka와 Rosenfeld 방법, (d) Liu와 Srihari 방법, (e) water flow model 기반 방법, (f) 제안한 방법

Fig. 10. Binarization results of test image 3; (a) Otsu's method, (b) Kapur's method, (c) Weszka and Rosenfeld's method, (d) Liu and Srihari's method, (e) Water flow model based method, (f) Proposed method.

인할 수 있다. 그림 9(d)의 Liu와 Srihari 방법을 적용한 결과에서는 배경과 문자가 분리되는 되었으나 문자 획의 굵기가 균일하지 못한 저품질의 이진영상이 획득되었

다. 그림 9(e)의 water flow model 기반의 이진화 결과에서는 문자 분리 성능은 우수하지만 문자 획의 굵기가 배경의 밝기에 따라서 균일하지 않았다. 반면 제안한

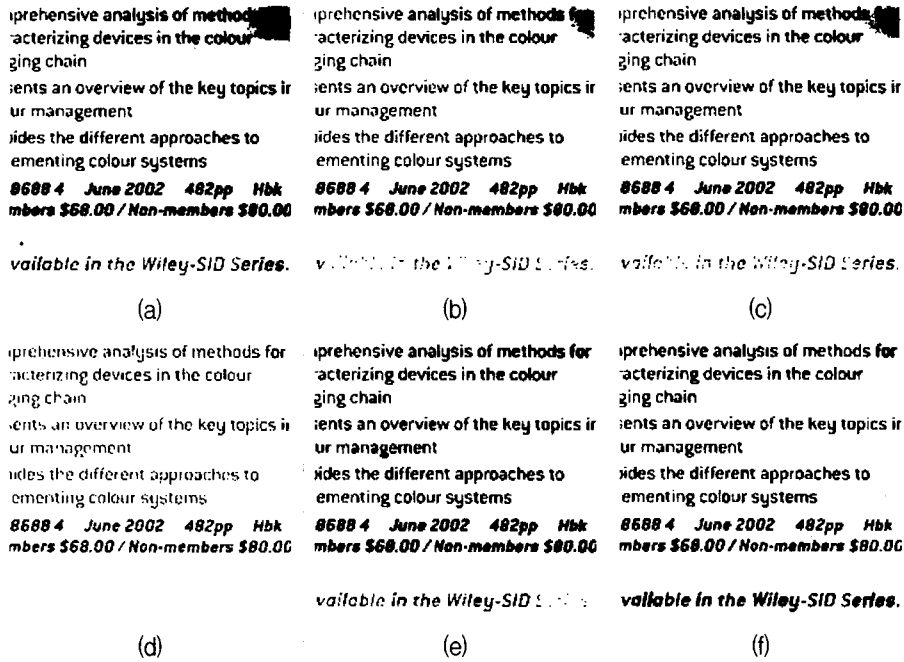


그림 11. 실험영상4에 대한 이진화 결과 (a) Otsu 방법, (b) Kapur 방법, (c) Weszka와 Rosenfeld 방법, (d) Liu와 Srihari 방법, (e) water flow model 기반 방법, (f) 제안한 방법

Fig. 11. Binarization results of test image 4; (a) Otsu's method, (b) Kapur's method, (c) Weszka and Rosenfeld's method, (d) Liu and Srihari's method, (e) Water flow model based method, (f) Proposed method.

방법의 결과에서는 문자 분리 성능이 우수할 뿐만 아니라 문자 획의 굵기가 균일한 이진영상이 획득됨을 그림 9(f)에서 확인할 수 있다.

실험영상3은 좌측의 배경 밝기가 어둡고 문자가 다소 흐릿하므로 Otsu 방법과 Kapur 방법, Liu와 Srihari 방법, Weszka와 Rosenfeld 방법을 적용하면 이진화 품질이 크게 저하됨을 그림 10(a)에서 (d)의 결과로부터 확인할 수 있다. 그림 10(e)의 water flow model기반의 방법을 적용한 결과에서는 우수한 이진화 성능을 나타내었으나 문자 획의 일부가 유실되는 경우가 발생함을 볼 수 있다. 제안한 방법을 적용한 그림 10(f)의 결과를 살펴보면, 문자 획의 유실이 거의 없으며 문자 획 굵기의 균일성이 기존의 water flow model 기반 방법의 결과와 비교하여 우수함을 확인할 수 있다.

그림 11은 배경의 밝기 값이 다양하고 문자의 밝기 값과 크기가 다양한 실험영상4에 대한 이진화 결과이다. 그림 11(a)의 Otsu 방법을 적용한 결과에서는 우측 상단부분의 문자가 배경으로부터 분리되지 못하고 병합된 결과를 나타내었다. Kapur 방법과 Liu와 Srihari 방법을 적용한 결과는 그림 11(b)와 그림 11(c)에 나타내었다. 두 결과에서는 우측 상단의 문자가 분리되지 못하였으며 영상의 하단에 위치한 밝기 값이 높은 문자에서 획의 유실이 많이 발생함을 확인할 수 있다. 그림

11(d)의 Weszka와 Rosenfeld 방법을 이용한 결과에서는 하단의 문자가 모두 유실되는 문제가 발생하였다. 그림 11(e)의 water flow model에 기반한 방법을 이용한 결과에서도 하단부의 문자가 일부 유실되는 문제가 발생하였다. 이 결과는 계곡에 채워진 물의 양에 전역 이진화를 적용함으로써 인하여 발생한 문제이다. 제안한 이진화 방법을 적용한 그림 11(f)의 결과에서는 연못마다 적용적으로 결정된 임계치에 의하여 문자의 유실 없이 가장 우수한 품질의 이진 문서영상이 획득됨을 확인할 수 있다.

이진화 성능을 비교하기 위하여 구현된 Otsu 방법, Kapur 방법, Weszka와 Rosenfeld 방법, Liu와 Srihari 방법을 4개의 실험영상에 적용한 경우 평균 수행시간은 0.01초 미만으로 매우 짧다. 제안한 방법에서 소요되는 평균 수행시간은 0.54초이지만 water flow model에 기반한 방법의 9.33sec와 비교하여 현저하게 단축된 결과이다. 그림 12에 제안한 방법과 water flow model에 기반한 방법의 수행시간을 영상마다 비교하여 나타내었다. 제안한 방법을 적용한 결과에서 실험영상 모두에 대하여 수행시간이 현저하게 단축되었음을 확인할 수 있다. 특히 실험영상4에 water flow model 기반 방법을 적용한 경우는 하단의 문자가 유실되는 문제를 방지하고자 물 붓는 과정을 30번 반복하였음에도 불구하고 그

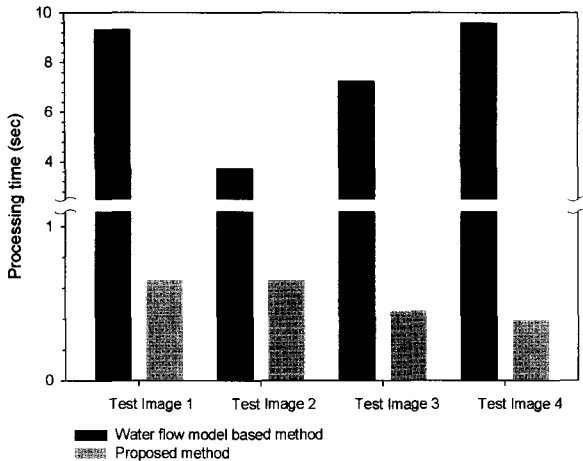


그림 12. 실험영상에 대한 제안한 방법과 water flow model 기반 방법의 수행시간 비교

Fig. 12. Processing time comparison between proposed method and water flow model based method with respect to test images.

림 11(e)에서 보인 바와 같이 문자 유실은 여전히 발생하였으며 수행시간은 9.59초로 매우 길었다.

V. 결 론

Water flow model에 기반한 문서영상 이진화 방법은 영상의 밝기 값을 3차원 영상지형으로 간주하고 물이 흘러가는 개념을 적용한 국부 적응 이진화 방법이다. 이 방법에서는 영상지형 전체에 물 붓는 과정을 여러 번 반복하여야 하므로 수행속도가 매우 느리다. 또한 계곡에 누적되는 물의 양을 자동으로 결정하기 어려우며, thresholding과정에서 전역 이진화를 적용함으로써 문자가 유실되는 문제가 발생한다. 본 논문에서는 water flow model에 기반한 이진화 방법의 속도를 향상시키기 위한 방법을 제안하였다. 먼저 제안한 방법에서는 그래디언트 크기를 이용하여 문자 주위 영역을 관심영역으로 추출하였다. 그리고 관심영역 내에만 물 붓는 과정을 수행함으로써 물이 뿌려지는 영역을 감소시키고 국부 계곡으로 흘러 들어가는 경로를 단축시키는 효과를 획득하였다. 국부 계곡에 누적되는 물의 양은 계곡의 깊이와 경사를 이용하여 자동으로 결정되도록 하였다. 그리고 국부 최저 지점과 그 주위 지점에 가중치를 부여하여 물을 누적함으로써 한번의 수행으로 충분한 양의 물이 계곡에 채워지도록 하였다. 계곡에 형성된 연못의 깊이는 배경과 문자의 밝기 값 차에 따라 차이가 크므로 연못마다의 특징을 이용하여 thresholding을 위한 임계치를 독립적으로 결정하는 방법을 제

안하였다. 실험결과에서 제안한 방법은 water flow model에 기반한 이진화 방법과 비교하여 속도가 월등하게 향상되었음을 확인하였다. 또한 잡음에 의하여 심하게 손상된 문서영상이나 문자 주위에 다양한 배경이 존재하는 국부적 특징이 강한 문서영상에 대하여 다른 이진화 방법의 결과와 비교하여 우수한 이진화 성능을 나타내었다. 특히 이진화된 문자 획 굵기의 균일성은 water flow model을 이용한 방법의 결과보다도 우수한 품질을 나타내었다. 제안한 방법은 실제 문서영상에서 문자를 분리하고 자동으로 인식하는 응용에서 효과적인 전처리 과정으로써 사용될 수 있을 것으로 사료된다.

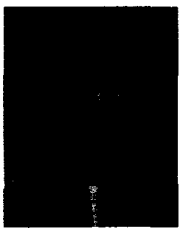
참 고 문 헌

- [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] G. Leedham, S. Varma, A. Patankar, and V. Govindaraju, "Separating Text and Background in Degraded Document Images-A Comparison of Global Thresholding Techniques for Multi-Stage Thresholding," in *Proc. of the Eighth Int'l Workshop on Frontiers in Handwritten Recognition*, pp. 244-249, 2002.
- [3] J. Sauvola and M. Pietikainen, "Adaptive Document Image Binarization," *Pattern Recognition*, vol. 33, pp. 225-236, 2000.
- [4] N. Otsu, "A Threshold Selection Method from Gray-Scale Histogram," *IEEE Trans. Systems, Man, and Cybernetics*, vol. SMC-8, pp. 62-66, 1978.
- [5] J.N. Kapur, P.K. Sahoo, and A.K.C. Wang, "A New Method for Gray-Level Picture Thresholding Using Entropy of the Histogram," *Graphical Models and Image Processing*, vol. 29, pp. 273-285, 1985.
- [6] J. Kitter and J. Illingworth, "On Threshold Selection Using Clustering Criteria," *IEEE Trans. Systems, Man, and Cybernetics*, vol. SMC-15, pp. 652-655, 1985.
- [7] L. O'Gorman, "Binarization and Multithresholding of Document Image Using Connectivity," *Graphical Models and Image Processing*, vol. 56, no. 6, pp. 494-506, 1994.
- [8] 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.

- [9] S. Rodtook and Y. Rangsanteri, "Adaptive Thresholding of Document Images Based on Laplacian Sign," in *Proc. of Int'l Conf. on Information Technology: Coding and Computing*, pp. 501-505, 2001.
- [10] 박성호, 장대근, 황찬식, "1차원 메디안 필터 기반 문서영상 영역해석," *전자공학회논문지*, 제40권 SP편, 제3호, 10-18쪽, 2003년 5월.
- [11] Y. Liu and S.N. Srihari, "Document Image Binarization Based on Texture Features," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 540-544, 1997.
- [12] 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.

저 자 소 개



오 현 화(정회원)

1998년 2월 경북대학교
전자공학과 (공학사),
2000년 2월 경북대학교 전자공학과
대학원(공학석사)
2000년 3월~현재 경북대학교
전자공학과 대학원 박사과정

<주관심분야: 컴퓨터비전, 패턴인식, 영상처리 등>



김 도 훈(학생회원)

2003년 2월 창원대학교 정보통신
공학과 (공학사),
2003년 3월~현재 경북대학교
전자공학과 대학원
석사과정,

<주관심분야: 영상처리, 컴퓨터비전, 패턴인식>



이 재 응(학생회원)

2002년 2월 경북대학교 전자
공학과 (공학사).
2004년 2월 경북대학교 전자
공학과 대학원(공학석사),
2004년 2월~현재 삼성 SDI PDP
본부

<주관심분야: 영상처리, 패턴인식, 컴퓨터비전>



김 두 식(정회원)

1995년 고려대학교 전산과학과
(이학사),
1997년 고려대학교 전산과학과
(이학석사),
1998년 한국IBM 소프트웨어
연구소 연구원,

1999년~현재 한국전자통신연구원 우정기술연구
센터 선임연구원,

<주관심분야: 영상처리, 패턴인식 등>



임 길 택(정회원)

1993년 2월 경북대학교
전자공학과 (공학사),
1995년 2월 경북대학교
전자공학과 (공학석사),
1999년 2월 경북대학교
전자공학과 (공학박사),

1999년 3월~2004년 2월 한국전자통신연구원
우정기술연구센터 선임연구원,

2004년 3월~현재 경주대학교 컴퓨터멀티미디어
공학부전임강사

<주관심분야: 패턴인식, 컴퓨터비전, 멀티미디어
신호 처리, 신경망 등>

진 성 일(정회원)

대한전자공학회 논문지 제 39권 SP편 4호 참조