

# 모폴로지를 이용한 문서 영상내의 특징영역 추출

이상협, 이경무  
홍익대학교 전파공학과

Sang Hyup Lee Kyoung Mu Lee

Dept. of radio science & communication engineering, Honk-Ik Univ.

## Abatract

컴퓨터를 이용한 문서정보의 처리를 위해서는 기본적으로 문서영상내의 각 특징영역을 분리하는 것이 필수적이다. 본 논문에서는 노이즈가 존재하는 non-manhattan layout 이치 문서영상내의 halftone 이미지, 선 및 텍스트등의 중요한 특징영역들을 자동으로 구분 추출 하는 효과적인 알고리즘을 제안한다. 제안한 알고리즘의 기본적인 아이디어는 먼저 처리속도의 고속화를 위하여 원본 영상을 축소시키는 것이 필수적인 바, 축소시 노이즈의 제거와 동시에 축소된 영상내에서 원하는 영역의 특징들이 잘 나타나도록하는 임계치 축소기법을 제안 사용하여 축소영상을 만든 다음, 축소영상에 다양한 모폴로지 필터를 적용함으로써 각 특징영역들을 효과적으로 추출한다. 제안한 알고리즘의 성능을 다양한 노이즈 문서영상을 이용한 시뮬레이션을 통하여 보인다.

## 1 서 론

최근 멀티미디어의 발달로 인하여 컴퓨터를 이용한 문서정보의 처리, 저장 및 전송의 요구가 증대됨에 따라 문서의 자동 분석 및 인식의 필요성이 대두되고 있고, 이에 따라 현재 국내외에서 문서자동인식(OCR)에 관한 연구가 매우 활발히 진행되고 있다. 문서자동인식의 효과적이고 정확한

처리를 위해서는 일반적으로 노이즈제거, 기울기보정(skew correction), 문서영상의 축소, 영역분리(blocking) 및 분리된 영역의 분류(classification)등과 같은 전처리과정이 필수적으로 선행되어야 하는데, 특히 영역 분리 및 분류의 결과가 이후 인식 및 해석에 결정적인 영향을 미치게 되므로 이에대한 연구가 매우 활발한 실정이다. 현재 개발된 영역 분리 및 분류알고리즘은 크게 top-down 방식과 bottom-up 방식으로 나눌 수 있는데, top-down 방식의 대표적인 것으로는 문서의 수평, 수직 투영 run length의 계층적 분할을 통해 블럭화를 수행하는 방식으로서 주로 manhattan layout 즉, 문서내의 모든 영역들이 동일 방향의 직사각형으로 구성되어 있어 수평 수직 분할이 용이한 문서에 효과적으로 적용된다 [4,5][9]. 반면, bottom-up 방식은 연결요소 (connected component)분석 기법과 같이 화소단위의 상호관계로부터 점진적으로 특징영역을 확장해 나가는 방식으로서 일반적인 non-manhattan layout 문서의 블럭화에 적용가능하나 halftone 이미지 영역분리 및 분류가 용이하지 않고 계산량이 많은 단점이 있다 [6,7,8]. 기존의 대부분의 알고리즘들은 처리와 해석이 비교적 간단한 manhattan layout 문서의 영역 분리와 분류에 국한 되고 있으나 최근 모폴로지 필터를 이용하여 축소된 영상내에서 non-manhattan layout 문서의 특징영역들을 효과적으로 분리하는 새로운 기법들이 연구되고 있다 [1,2].

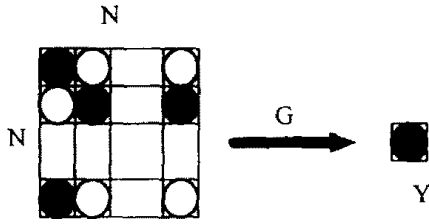


그림 1. 이치영상의 2-D N x N 축소 매핑.

이와 관련하여 본 논문에서는 원본 영상내에 불규칙 노이즈가 존재하거나 텍스트행간 간격이 매우 작은 경우 영상축소시 노이즈의 제거 및 aliasing을 최소화하는 기법과 축소영상내의 halftone 이미지, 수평 수직선 및 텍스트영역을 모폴로지 필터를 이용하여 효과적으로 추출하는 알고리즘을 제안한다.

## 2. 문서영상의 축소

일반적으로 문서영상처리에 사용되는 입력 영상은 그 크기가 매우 크기 때문에 실시간 처리 뿐 아니라 기억용량의 한계 때문에 축소된영상을 사용하여 처리하는 것이 바람직하다. 그러나 단순한 subsampling 기법을 사용하여 축소를 할 경우 특정 특징정보들이 사라지거나 변형되는 문제 (aliasing) 등이 발생 하게 된다. 따라서 축소영상을 사용하여 문서 처리를 할 경우 원본 문서영상내의 특징들이 축소시 그대로 보존되도록 하는 것이 필수 조건이 된다. 이와 더불어 원본 영상내에 노이즈가 존재할 경우 영상축소시 일반적으로 SNR이 감소하게 되어 특징 정보의 손상이 커지는 문제가 발생하게 되므로 이에 대한 대책 또한 강구되어야 만 한다. 본 장에서는 불규칙 노이즈 및 텍스트행간 간격의 분리가 충분하지 않은 문서영상의 효과적인 축소방법을 제시한다.

### 2.1 모폴로지를 이용한 임계치 축소

이진 영상에서 선, 그림, 문자를 이루는 화소 (black)을 ON, 배경화소 (white)을 OFF 라고 정의한다. 이진 영상을 N x N의

비율로 축소하는 것은 원본 영상을 먼저 N x N 크기의 블럭들로 나누고, 각 블럭들을

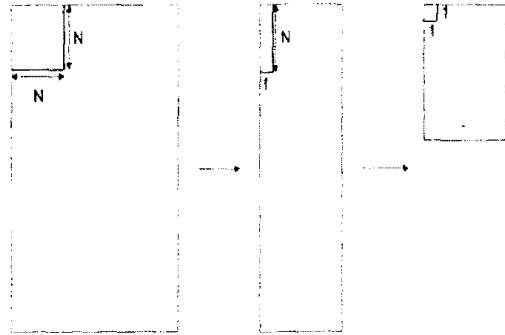


그림 2. 단계적 임계치 축소 알고리즘.

내부의 ON, OFF 분포에 따라 ON 또는 OFF 한 화소값으로 매핑시키는 것과 동일하다 (그림 1 참조). 블럭내부의 모든화소가 ON일 경우만 ON으로 매핑하는 경우 모폴로지 opening (AND)에 해당하는 반면 블럭내부의 화소중 최소 하나 이상 ON으로 매핑하는 경우는 모폴로지 closing (OR)에 해당한다. 또한 임계치를 이용하여 매핑을 일반화 시킬 경우 아래와 같은 확장된 개념의 임계치 AND 및 OR 연산을 정의할 수 있다.

임계치 AND:

$$Y = A(m) = \begin{cases} \text{ON} & \text{if } m = (N^2 - t) \\ \text{OFF} & \text{otherwise} \end{cases}$$

임계치 OR:

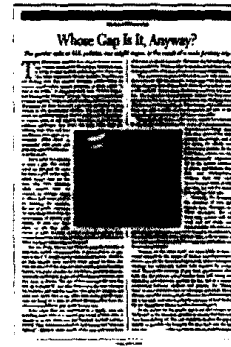
$$Y = O(m) = \begin{cases} \text{ON} & \text{if } m > t \\ \text{OFF} & \text{otherwise} \end{cases}$$

여기서 m과 t는 각각 N x N 블럭내의 ON 화소의 개수와 설정한 임계치를 나타낸다. 이러한 매핑을 이용하여 다양한 방법으로 영상을 축소할 수 있는 데 이 것을 임계치 축소 (threshold reduction)기법이라고 한다 [1]. 일반적으로 문서영상은 수평 수직 방향으로 서로 다른 텍스처 특성을 가지고 있으므로 2차원 직접축소 대신 수평 수직 방향으로 서로 다른 1차원 매핑을 순차적으로 행하여 축소하는 방식이 더욱 효과적이다. 그림 2는 원본 영상을 수평,

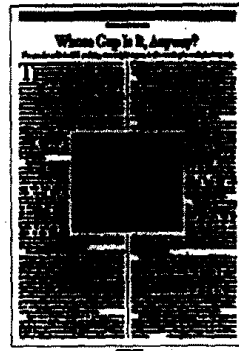
수직으로 축소하여 최종  $N \times N$  배 축소된 영상을 얻는 과정의 예를 보여준다. 본문에서는 halftone 이미지, 선 그리고 텍스트 영역을 정확히 추출 하는 것이 목적이므로 축소된 영상내에서 원하는 특징들이 잘 나타나도록 하기 위해서는 매핑과정에서 다음과 같은 사항이 고려되어야 한다.

- Halftone 이미지 정보: Halftone 영역은 축소되었을 때 가능한 한 내부가 ON 화소로 채워져 견고화 (solidate)되는 것이 바람직하므로 최소한 수평 또는 수직 한 방향으로의 OR 연산이 필요하다.
- 선 정보: 원본 영상의 선 정보가 축소 영상내에 투영되는 것은 OR 연산을 통해 가능하다. 따라서 수평선과 수직선 축소를 위해 각각 수평방향, 수직방향 모두 OR 연산이 요구된다.
- 텍스트 영역 정보: 텍스트 영역은 축소영상내에서 halftone 이미지와 선을 제외한 나머지로 구별 가능하나, 이후 텍스트인식을 위한 텍스트영역의 블러화를 위해서는 텍스트선은 수평 방향으로 연속되는 반면, 행간 간격이 축소된 영상내에 잘 분리되어 나타나 는 것이 바람직 하다. 따라서 이를 위하여 수평 방향으로는 OR 그리고 수직 방향으로 는 AND 연산이 요구된다.

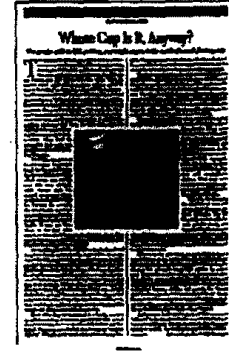
위의 조건들을 종합하여 보면 수평 수직 방향 모두 OR 연산을 적용한 축소 영상으로부터는 선 정보를, 그리고 수평 방향으로 OR, 수직 방향으로는 AND 연산을 이용하여 얻은 축소 영상으로부터는 halftone 이미지 및 텍스트 정보를 추출 할 수 있다는 것을 알 수 있다. 그림 3은 이러한 방식으로 얻은 두 축소영상 HoVo와 HoVa의 예를 보여 주는 데, 여기서 H, V는 각각 수평, 수직 그리고 첨자 o, a는 각각 OR와 AND 연산을 표시하며 H, V의 순서는 축소과정의 순서를 의미한다.



a) 원본 영상



b) HoVo 영상



c) HoVa 영상

그림 3. 축소 영상의 예: a) 원본 영상 b) 수평 방향으로 OR 수직 방향으로 AND로 축소된 영상 c) 수평 방향으로 OR, 수직 방향으로 AND된 축소영상.

## 2.2 노이즈가 존재하는 경우의 영상 축소

문서영상을 입력할 때 일반적으로 불규칙 노이즈가 발생하는 경우가 많은데 이러한 노이즈는 문서영상의 처리과정에서 많은 문제점들을 발생시킨다. 그림 4는 불규칙 노이즈가 존재할 경우 축소된 영상의 예를 보여 주는 데 일반적으로 OR 연산에 의해 연결 확장 되어 문서 특징들을 변형 왜곡시키는 것을 볼 수 있다. 이러한 불규칙 노이즈는 대부분 격리된 작은 ON 화소들로 구성되어 있으므로 AND 연산을 통하여 쉽게 제거될 수 있다. 따라서, 축소과정에서 수평 또는 수직 방향으로의 AND 연산을 수행함으로써 이러한 노이즈를 제거하는

것이 바람직한 데, 본 논문에서는 이러한 불규칙 노이즈의 제거와 아울러 halftone 이미지, 선, 텍스트 정보들이 잘 투영된 축소영상을 얻기 위하여 VoHa 와 HoVa 두 축소영상을 OR 연산 하여 얻은 영상을 제안 사용한다. 그림 5에서 보듯이 VoHa 및 HoVa 축소영상에서는 노이즈가 AND 연산에 의하여 대부분 제거 되었음을 알 수

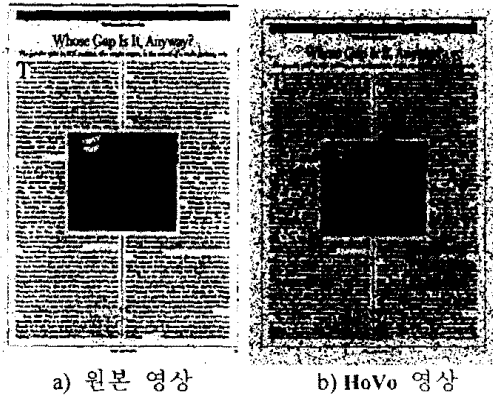


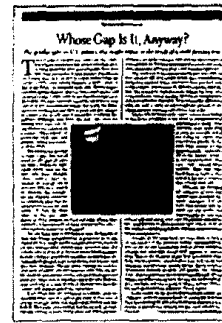
그림 4. 노이즈가 있는 영상의 축소 예

있다. 그러나 AND 연산의 결과에 따라 VoHa 에서는 수직선과 텍스트들이, HoVa 에서는 수평선이 사라지는 문제가 나타난다. 그러나 그림 5. C)에서 알 수 있듯이 이 두 영상의 OR 연산된 결과 영상에서는 수평 수직선 그리고 텍스트들이 잘 드러날 뿐만 아니라 halftone 이미지 영역 또한 더욱 견고해진다.

불규칙 노이즈 문제와 더불어 문서 영상 축소시 나타나는 중요한 문제로 원본 영상내의 텍스트행간 간격이 매우 작거나 연결되어 있는 경우이다. 이러한 영상을 수직 방향으로 AND 또는 OR 연산을 통해 축소할 경우 축소된 영상 내에서 텍스트행간 간격이 분리되지않게 되어 이후 텍스트 영역의 세분화 및 인식에 커다란 장애가 된다. 이러한 문제를 해결을 위해 VoHa 및 HoVa 축소과정시 OR 연산에 1 또는 2 화소 정도의 임계치를 설정하여 사용하면 매우 좋은 결과를 얻는 데, 그림 6의 예가 이것을 보여 준다. 그림 6 a)는 임계치를



a) VoHa 영상      b) HoVa 영상



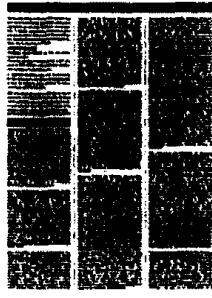
c) VoHa  $\circledR$  HoVa

그림 5. 노이즈 제거를 위한 영상 축소

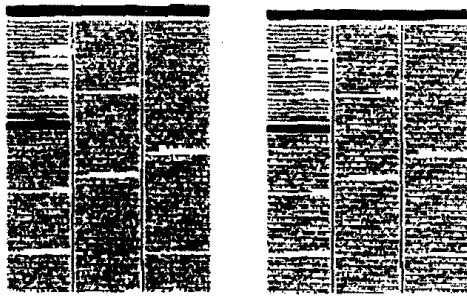
0으로 OR 연산을 사용한 결과이고, b)는 임계치 2의 OR연산을 이용한 결과인 데, 텍스트 행간 간격이 분리되는 효과를 확인할 수 있다.

### 3. 축소 문서영상에서의 특징영역 추출

2 장에서 설명한 바와 같이 VoHa 및 HoVa 영상의 OR연산에 의해 얻어진 축소영상은 노이즈에 별 영향없이 원본 영상이 갖고 있는 halftone 이미지, 선 및 텍스트 정보를 포함 한다. 따라서, 이 축소 영상으로부터 각각의 특징 영역들을 추출 하는 것이 필요한 데, 이는 그림 7에 도시된 과정을 통하여 구현할 수 있다.



a) 원본 영상



b) VoHa (OR) HoVa c) VoHa (OR) HoVa

그림 6. A) 텍스트 간격이 작은 문서 영상의 축소 예: b), c) AND연산은 고정, OR연산의 임계치는 각각 0, 2. 로 설정한 결과 영상.

즉 모폴로지 필터를 이용하여 축소영상으로부터 먼저 halftone 이미지 그리고 수평, 수직선을 추출한 다음 축소영상에서 이들 성분들을 XOR (Exclusive OR) 연산을 통해 빼줌으로써 텍스트 영역을 추출 한다. 각특징들을 추출하기 위한 구체적인 방법들은 아래와 같다.

- Halftone 이미지 추출:  
Halftone 이미지 영역을 추출하기 위해서는 먼저 큰 사각형 SE(Structure Element)를 이용한 모폴로지 erosion 연산을 통하여 halftone seed를 추출한 다음, seed의 ON 화소를 중심으로 작은 사각형 SE를 설정 축소영상의 해당 halftone 이미지와의 화소값 비교를 통해 원 이미지 영역을 복원한다.
- 수평 수직선 추출:  
수평 수직선 성분들은 먼저 축소영상에 작은 크기의 1차원 SE를 설정하여 모폴로지 closing을 수행한 다음 다시 최대 글자 크기 이상이고 최소 선분 길이 이하 크기의 1차원 SE를 이용한 모폴로지 opening을 수행하여 얻을 수 있다. 일반적으로 이러한 방법으로 얻은 영상에는 선 이외에 halftone 이미지 영역이 포함되기 때문에 미리 구한 halftone 이미지를 XOR 연산을 이용하여 제거함으로써 최종적으로 선 정보만을 추출 한다.
- 텍스트 영역추출:  
텍스트영역 정보는 축소영상에서 추출된 halftone 영상과 선 성분을 제거하면 얻을 수 있는데, 이것은 XOR 연산을 통하여 쉽게 구현할 수 있다.

그림 8은 제안한 방식을 이용하여 그림 4. c) 의 축소영상으로부터 추출된 halftone 이미지, 중간단계의 수평 수직 성분 그리고 최종 수평 수직선 그리고 텍스트 영역 결과를 보여 주고 있다.

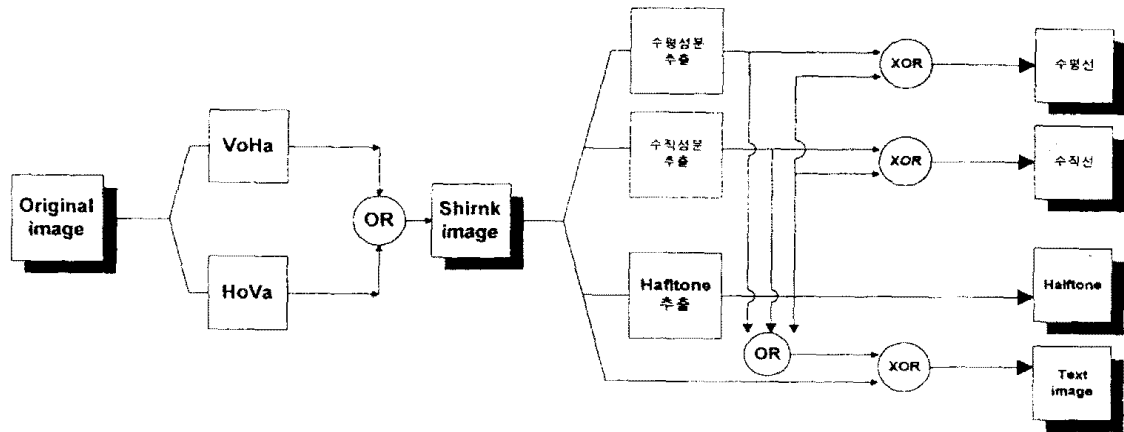
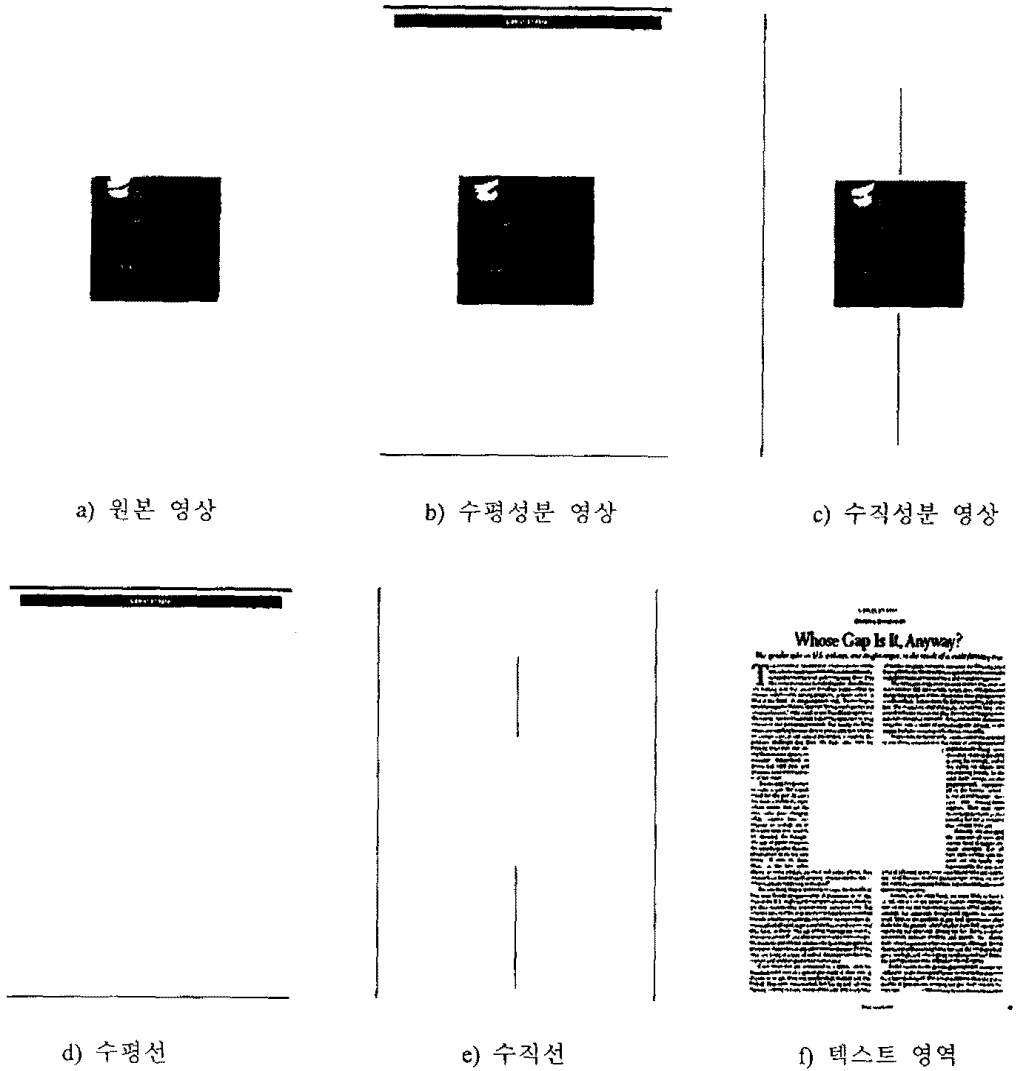


그림 7. 문서특징 영역 추출 알고리즘의 개략도.



a) 원본 영상

b) 수평성분 영상

c) 수직성분 영상

d) 수평선

e) 수직선

f) 텍스트 영역

그림 8. 그림 4 a) 의 영상으로부터 추출된 문서 특징.



a) 원본 영상



b) HoVa **OR** VoHa 영상



c) Half-tone 영상



d) 수평선

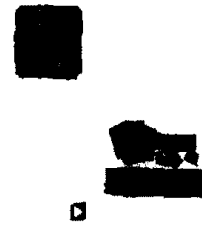
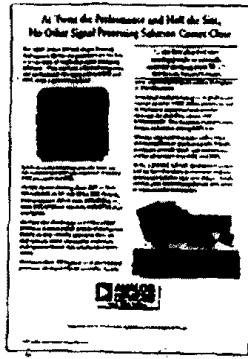
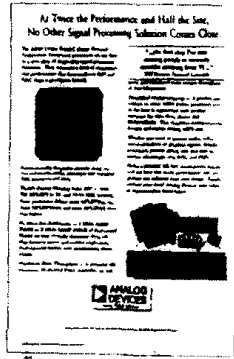


e) 수직선



f) 텍스트 영역

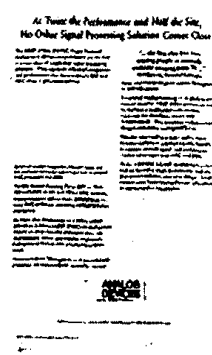
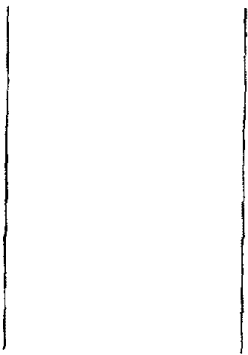
그림 9 문서특징 추출 알고리즘의 테스트 결과 1



a) 원본 영상

b) HoVa **OR** VoHa 영상

c) Halfone 영상



d) 수평선

e) 수직선

f) 텍스트 영역

그림 10. 문서특징 추출 알고리즘의 테스트 결과 2

#### 4. 시뮬레이션 및 검토

제안한 문서영상 축소 및 특징영역 분리 알고리즘을 스캐너로 입력된 다수의 이치 영상에 대하여 시뮬레이션을 수행하였고 그결과를 그림 9, 10에 도시하였다. 본 시뮬레이션에 사용된 원본 문서영상은 3500

x 2500 크기의 300dpi 이치 영상이고 수평 수직방향으로 각각 8:1 축소를 하였다. Halfone seed 추출을 위해서는 13 x 13 크기의 사각형 SE를 이용하여 erosion을 하였고, halfone 이미지 복원을 위해서는 7 x 7 크기의 윈도우를 사용하였다. 선 추출을 위해서는 먼저 수평 수직 모두 3 화소 크기의 closing SE를 설정 하였고 300dpi의 영상을 8 배 축소할 경우 약 40dpi가 되므로 이를 기준으로 수평



opening SE의 크기는 60 화소 (약 1.5 inch) 그리고 수직 opening SE의 크기는 약 40 화소 (약 1.0 inch)로 설정하여 모폴로지 필터링을 수행하였다. 그림 8 c)의 halftone 영역에 수평선 성분이 포함된 것은 원본 영상에서 두 성분이 분리되어 있지 않은 결과이다. 또한 e)의 텍스트 영상의 halftone 영역 부근에 노이즈가 남아 있는 것은 축소영상의 halftone 영역이 충분히 견고하지 않아 이의 복원이 완벽하지 않기 때문이다. 그림 10의 예는 대체로 만족할 만한 추출 결과를 보여주고 있다. 실제 알고리즘 구현시 정확한 특징 영역 추출을 위해서는 축소 비율에 맞는 모폴로지 필터의 크기를 설정하는 것이 중요한 데, 텍스트 부분이 추출되지 않도록 모폴로지 필터의 크기를 설정해야 함은 물론 선 또는 halftone 영역 추출 후 주위 화소를 조사하여 텍스트 영상내에 노이즈가 남지 않도록 하는 후처리과정도 필요하다.

## 5. 결론

본 논문에서는 자동문서인식 처리에 필요한 문서내의 halftone 이미지, 선 및 텍스트 영역들을 추출하는 효과적인 알고리즘을 제안 하였다. 제안한 알고리즘은 실시간 처리를 위하여 원본 문서를 축소하여 사용하는 데, 축소시 노이즈를 제거하고 aliasing을 방지하는 동시에 원하는 문서정보의 특징들의 손상을 최소화하기 위하여 임계치 축소기법을 이용하였고, 축소된 영상을 모폴로지 필터링함으로써 텍스트, 선 그리고 halftone 이미지 정보등을 정확히 추출할 수 있다. 노이즈가 있는 다수의 문서영상에 대하여 알고리즘의 성능을 테스트한 바, 정확한 특징영역 추출이 가능함을 보였다. 향후 보다 정확한 분리를 위한 후처리과정에 대한 연구와 사선추출 에 대한 연구가 필요하다.

## 참고문헌

[1] D.S. Bloomberg, " Image analysis using

- threshold reduction," SPIE Conf. on Image Algebra and Morphological Image Processing II, Vol.1568, San diego, CA, July 1991, pp.38-52.
- [2] D.S. Bloomberg, " Textured reduction for document image analysis", SPIE Conf.2660, Document Recognition III, San Jose, California January 1996, pp.160-174.
- [3] Edward R. Dougherty, " An Introduction to Morphological Image Processing ", SPIE OPTICAL ENGINEERING PRESS, TUTORIAL TEXTS in optical engineering .Volume TT9.
- [4] J. L. Fisher, S. C. Hinds, and D. P. D'Amato, "A rule-based system for document image segmentation", Proc. 10<sup>th</sup> ICPR, Los Alamitos, CA, 1990, pp. 567-572.
- [5] G. Nagy and S. C. Seth; "Hierarchical representation of optically scanned document", 7<sup>th</sup> ICPR, Montreal, 1984, pp. 347-349.
- [6] L. O'Gorman, "The document spectrum for layout analysis", IEEE trans. on PAMI, vol. 15, No. 11, Nov. 1993, pp. 1162-1173
- [7] T. Pavlidis and J. Zhou, "Page segmentation and classification", CVGIP: Graphical Models and Image Processing, vol. 54, No. 6, Nov. 1992, pp. 484-496.
- [8] J. Sauvola and M. Pietikainen, "Page segmentation and classification using fast feature extraction and connectivity analysis",
- [9] F. M. Wahl, K. Y. Wong and R. G. Casey, "Block segmentation and text extraction in mixed text/image document", Computer Graphics and Image Processing, vol. 20, 1982, pp. 375-390.