

블록의 속성과 질감특징을 이용한 문서영상의 블록분류

장영내[†], 김종수^{**}, 이철희^{***}

요 약

본 논문에서는 블록의 속성과 질감특징을 이용하여 효과적인 블록 분류 방법을 제안하였다. 제안한 방법에서는 먼저 명암도 문서영상을 이진화한 후, 평활화 기법을 적용하여 블록의 위치정보와 본 논문에서 사용할 특징 중에 하나인 각 블록의 내부에 있는 작은 블록들의 최대 높이 값을 구하였다. 이 위치정보들을 이용하여 문서영상을 각 블록으로 분할한다. 이 블록의 명암도 블록영상에서 문서의 속성이 잘 반영된 (0,1) 방향의 공간 명암도 의존 행렬을 구하여 7가지 질감특징을 구하였다. 먼저 블록의 속성을 최소거리 규칙(Nearest Neighbor Rule)에 입력하여 문자와 비문자 영역으로, 상세분류를 위하여 7가지 질감특징을 이용하여 큰 문자, 작은 문자, 표, 그래픽 및 사진 등으로 구분함으로써 문서인식을 위한 구조 해석뿐만 아니라 다양한 응용 분야에 효과적으로 이용될 수 있도록 하였다.

Block Classification of Document Images by Block Attributes and Texture Features

Young-Nae Jang[†], Joong-Soo Kim^{**}, Cheol-Hee Lee^{***}

ABSTRACT

We propose an effective method for block classification in a document image. The gray level document image is converted to the binary image for a block segmentation. This binary image would be smoothed to find the locations and sizes of each block. And especially during this smoothing, the inner block heights of each block are obtained. The gray level image is divided to several blocks by these location informations. The SGLDM(spatial gray level dependence matrices) are made using the each gray-level document block and the seven second-order statistical texture features are extracted from the (0,1) direction's SGLDM which include the document attributes. Document image blocks are classified to two groups, text and non-text group, by the inner block height of the block at the nearest neighbor rule. The seven texture features(that were extracted from the SGLDM) are used for the five detail categories of small font, large font, table, graphic and photo blocks. These document blocks are available not only for structure analysis of document recognition but also the various applied area.

Key words: Binary Image(이진 영상), Gray Level Image(명암도 영상), Texture Image(질감 영상), Block Segmentation(블록 분할), Block Classification(블록분류)

※ 교신저자(Corresponding Author): 장영내, 주소: 서울특별시 영등포구 문래3가 82-29 우리벤처타운2 11층 2호 (150-700), 전화: 02)2164-0500, FAX: 02)2164-0600, E-mail: acebrain@greenbell.com

접수일: 2006년 12월 19일, 완료일: 2007년 5월 31일

[†] (주)그린벨시스템즈 정보기술연구소 IETM 개발팀

^{**} 정회원, 안동대학교 컴퓨터공학과 (E-mail: kimjs@andong.ac.kr)

^{***} 정회원, 산업자원부 기술표준원 ISO/TC-130 전문위원 (E-mail: chlee@andong.ac.kr)

※ 본 연구는 안동대학교 특별 연구지원으로 수행되었음

1. 서 론

최근 저장 매체의 혁신과 컴퓨터 시스템의 발전에 힘입어 기존의 출판문서에 대한 전자문서화가 가속되고 있다. 과거 환경에 비하여 정보 전달의 역할을 맡고 있던 문서들이 모두 전자화 되어 간다고 할 수 있다. 따라서 많은 양의 문서정보를 저장하여 빠르게 검색하고 처리할 수 있는 시스템의 구축과 이에 따른 컴퓨터 시스템에 의한 문서자동인식이 불가피하게 되었다. 이러한 흐름에 따라 문서영상의 구조를 분석하여 문자, 도표, 사진 등을 자동 검출하고 인식하는 문서자동인식에 대한 많은 연구가 주로 전개되어 왔다[1-4].

문서자동인식 시스템에서는 스캐너나 카메라를 이용하여 여러 유형의 블록으로 복합 구성된 문서영상을 입력하여 동질의 블록별로 분할하고, 분할된 블록으로 특징을 추출하여 각 블록을 유형별로 분류한 후, 블록의 유형과 위치정보를 이용하여 문서를 인식하거나 문서구조를 분석하게 된다[5].

지금까지의 문서자동인식 소프트웨어들의 대부분은 문서영상에서 문자 블록과 비문자 블록 등으로 포괄적으로 분류하였다. 그러나 문서영상의 블록을 상세하게 분류하여야 각 블록 유형별로 처리를 할 수 있고 또한, 문서 데이터베이스시스템의 경우에는 특정 블록별로 검색하여 수정, 편집, 재구성, 저장 및 출력하는 것이 가능해야 하며[6], 문서 전송에서도 여러 가지 유형의 블록들에 대하여 유형에 맞는 최적의 알고리즘을 적용하여 블록을 인식 및 압축하여 전송하는 것이 효과적이므로 문서 구조 분석에 더 많은 정보를 줄 수 있도록 유형을 가급적 상세히 분류하는 것이 요구된다. 따라서 문서를 구성하는 여러 유형의 블록들을 잘 구별해 낼 수 있는 효과적인 특징추출과 유형 분류에 관한 연구가 필요하다[4,5].

블록분류를 위한 특징은 이진영상(Binary Image)에서 추출하거나 명암도 영상(Gray Level Image)에서 추출된다. 이진영상에서 특징을 추출하는 경우에는 블록의 높이, 블록 부분의 화소의 비율 및 평균 흑화소의 길이 등, 블록의 속성을 특징으로 추출하는 방법[3,7,8]과 질감(Texture)을 특징으로 추출하는 방법[9] 등이 있다. 블록을 단순히 문자와 비문자, 기사, 부표제, 표제, 사진 등 포괄적으로 분류하는 경우에는 블록의 속성 특징을 많이 이용하고, 블록을 큰

문자, 작은 문자, 그래픽, 사진 등 구체적으로 세분할 경우에는 질감특징을 많이 이용하고 있다. 그러나 명암도 영상을 배경으로 가지거나, 영상의 질감이 포함된 문서영상을 이진화를 하면 문서 영상 내에 포함된 일부 정보들이 손실될 우려가 있다. 따라서 명암도 영상에서 직접 특징을 추출하여 문서의 구조를 해석하거나 블록분류 및 문서인식을 수행하는 연구들이 이루어져 왔다.

명암도 영상에서 특징을 추출하는 방법에는 영상의 기하학적인 특징을 이용하는 방법과 블록의 질감 특징을 이용하는 방법[10,11] 등이 있다.

질감영상은 한 화소와 그 이웃 화소들 사이의 관계에 의해 그 성질이 결정된다. 그러므로 단순히 영상 내에 존재하는 화소 값의 빈도수만을 표시하는 히스토그램에서 추출되는 1차 통계 특성만으로는 질감 영상을 분류하기 어렵다. 따라서 이웃 화소들과의 관계를 나타내는 공간 정보(Spatial Information)를 표현할 수 있는 2차 통계 특성을 고려하여야 한다. Julesz[12]의 실험에서도 인간의 시각능력은 2차 통계 특성이 있어야 두 종류의 서로 다른 질감을 잘 구별할 수 있음을 보였다.

이러한 명암도 영상에서 블록의 질감특징을 이용하는 방법들은 사진, 그래프, 도표, 글자블록 등의 분할에는 효과적이지만, 여러 가지 질감특징을 추출하기 위하여 많은 시간이 소요된다는 단점이 있다.

본 논문에서는 블록의 속성과 질감특징을 이용하여 효과적인 블록 분류 방법을 제안하였다. 제안한 방법에서는 먼저 명암도 문서영상을 이진화를 하여 평활화 기법을 적용함으로써 질감특징을 이용하여 블록을 분할하는 것보다 신속하게 블록을 분할하는 동시에 그 블록의 위치정보와 본 논문에서 사용할 특징 중에 하나인 각 블록의 내부에 있는 작은 블록의 최대 높이 값을 구할 수 있도록 하였다. 그리고 분할된 각 블록들의 공간 명암도 의존 행렬로부터 문서블록들의 7가지 질감특징인 평균, 분산, 상관계수, 에너지, 엔트로피, 대비, 국부균일성을 구하고, 이를 블록의 속성과 더불어 최소거리 규칙(Nearest Neighbor Rule)[13,14]에 학습시킴으로써 문서 블록들을 상세하게 분류해 낼 수 있도록 하였다.

문서블록들을 우선적으로 블록의 내부블록 최대 높이 값을 이용하여 문자와 비문자 블록으로 구분하고 다시 큰 문자, 작은 문자, 표, 그래픽 및 사진 등

다섯 가지 유형으로 상세 분류함으로써 문서인식을 위한 구조 해석뿐만 아니라 다양한 응용 분야에 효과적으로 이용될 수 있도록 하였다.

제안한 블록분류 방법의 성능을 확인하기 위해 대한 전자 공학회 논문지와 한국 정보과학회 논문지의 다양한 실제 문서들에 대하여 실험하였으며, 제안한 블록의 속성과 질감특징을 이용한 문서영상의 블록분류 방법을 컴퓨터 시뮬레이션하고, 동일 데이터를 문서자동인식 시스템에 적용하여 그 결과를 비교, 검토하였다.

본 논문은 2. 문서 영상의 블록분류, 3. 블록의 속성과 질감특징을 이용한 블록 분류, 4. 실험 및 고찰 그리고 5. 결론 등으로 구성되어 있다.

2. 문서영상의 블록분류

문서의 블록분류 과정은 크게 블록분할, 특징추출 및 유형분류 등으로 구성된다. 스캐너나 카메라로 입력된 문서영상에서 이진화나 기울기보정 등의 전처리 과정을 거쳐 동질의 영역으로 블록을 분할하고, 분할된 각 블록으로부터 분류에 필요한 특징들을 추출하여 분류기를 통하여 블록의 유형을 분류한다. 입력된 문서영상의 블록분류 과정에서 출력되는 블록의 종합적인 정보는 문서자동인식 시스템의 문자인식에서 블록의 구조 해석뿐만 아니라 문서 데이터베이스 시스템, 컴퓨터 통신 등에 활용된다.

일반적으로 문서자동인식 시스템의 문자인식에서 사용되는 유형 분류는 대부분 문자와 비문자 블록으로 단순분류해도 되지만, 보다 나은 성능을 보이기 위하여 최근에는 문자영역을 좀 더 세분화하고 도표 영역과 그림 영역으로 구분하기도 한다. 문자 블록은 흔히 문자추출을 거쳐 인식되고 후처리 과정을 통하여 인식결과가 보완되며, 비문자 블록은 효율적인 기억장치 활용을 위해 압축 저장된다. 또한 문서 내의 각 블록정보로부터 순서와 의미적 관계해석을 거쳐 문서를 인식하며, 이렇게 인식된 문서정보들은 수정 및 탐색을 위해 문서의 재구성 및 편집과정을 거쳐 저장된다.

2.1 블록분할

블록분할 방법은 크게 이진영상을 이용하는 방법과 명암도 영상을 이용하는 방법으로 나눌 수 있다.

이진영상을 이용하는 방법에는 투영에 의한 방법[9], 평활화에 의한 방법[3,7,9], 그리고 흑화소 순번화에 의한 방법[15,16] 등이 있으며, 명암도 영상을 이용하는 방법에는 질감특징을 이용하여 분할하는 방법[17] 등이 있다.

일반적으로 블록분할 과정에는 투영에 의한 방법과 흑화소 순번화 방법 등이 널리 이용되고 있으나 투영에 의한 방법은 작은 블록을 구하기 위해서는 여러 번 반복 투영을 해야 한다는 단점이 있다. 그리고, 흑화소 순번화 방법은 이진영상에서 여러 방향의 흑화소 연결을 구하여 영역으로 만든다. 이 영역을 만들기 위해 각 흑화소에 대해 여러 번의 스캔이 필요하고, 각 영역에 하나의 고유값을 할당하기 위해 큰 용량의 메모리가 요구된다.

평활화에 의한 방법은 사이에 임계값보다 작은 수의 백화소가 끼어있는 두 흑화소는 흑화소의 연속으로 만드는 방법으로 수직방향 평활화한 영상과 수평방향 평활화한 영상을 이용하여 블록을 분할한다.

명암도 문서영상을 블록분할 하는데 있어서 Jain[17] 등은 각 화소를 중심으로 일정한 크기의 마스크에 여러 개의 내칭 필터를 적용하여 질감특징이 나타나는 영상을 구한 다음, 이 영상에 제곱에러 클러스터링(Squared-error Clustering) 알고리즘을 적용하여 영상을 블록으로 분할하였다. 이러한 방법에서는 명암도 문서영상을 일반 사진영상의 질감분할과 같은 방법으로 전체영상을 주사하여 서로 다른 질감특징을 가진 영역을 찾아서 블록으로 분할하기 때문에 시간이 많이 소요된다.

문서의 블록분할에 있어서는 명암도 문서영상의 질감특징을 이용하기보다는 일단 이진화를 하여 각 블록의 속성을 이용하는 것이 더 단순하고 빠름을 알 수 있다. 즉, 줄 간격이나 다른 유형블록과의 간격 등 문서의 형식과 구조상의 특징을 이용하여 블록분할 하는 것이 더 빠르고 효과적인 뿐만 아니라 부수적으로 각 블록의 속성 및 위치정보도 얻을 수 있다.

2.2 이진영상에서의 블록의 특징추출 및 유형분류

분할된 블록을 유형 분류하기 위해서는 각 블록의 특징을 추출하여야 한다. 문서영상의 각 블록에서 특징을 추출하는 방법에는 이진영상에서 블록의 속성을 특징으로 추출하는 방법과 질감을 특징으로 추출하는 방법이 있다.

이진영상에서 블록의 속성을 특징으로 추출하는 방법에는 연결 화소간의 특징을 찾기 위하여 여러 방향으로 주사하는 방법과 이웃 화소를 찾는 방향을 단순화한 방법 등이 있다. 이러한 블록의 추출된 속성에는 정규화 문자의 크기, 블록의 크기, 블록 속에 있는 화소수 등과 블록의 좌표, 가로 세로의 길이 및 비율 등이 주로 사용되었다.

이진영상에서 질감을 특징으로 추출하는 방법[9]에서는 입력 문서 한 페이지 전체를 이진화하여 이 영상을 특정한 형태의 반복적인 규칙을 갖는 질감 영상으로 간주하고 흑-백 및 흑-백-흑화소의 연결 길이가 거의 같고, 그 분포가 비슷한 것을 하나의 단위로 이들의 빈도수를 구하여 흑-백 쌍의 연결길이(black-white pair run length) 및 흑-백-흑 조합의 연결 길이(black-white-black combination run length)를 요소값으로 사용하는 각각의 행렬을 만든 다음, 이 행렬들을 이용하여 세 가지 질감특징을 추출하였다. 이 방법은 특징 추출시 각 연결 길이에 대한 임계값을 잘 선정하여야 정확한 특징이 추출되며, 또 이진영상을 사용하므로 이진화시 많은 정보가 손실되는 단점도 있다.

유형의 분류과정에서는 큰 문자, 작은 문자, 표, 그림 등 문서에 포함된 여러 가지 블록을 그 특징에 따라 분류한다. 기존의 방법으로는 이진영상에서 블록의 형태 속성을 이용하여 분류하는 방법과 질감특징을 이용하여 분류하는 방법, 명암도 영상에서 질감특징을 이용하여 영상의 블록을 분할하고 분류하는 방법 등이 있다.

이진영상에서 블록의 형태 속성을 이용하여 블록의 유형을 분류하는 방법은 문자의 크기를 이용하여 문자와 비문자 블록으로 분류하는 방법과 문자의 크기, 흑화소의 강조, 면적, 위치를 이용하여 문자열을 추출하는 방법, 블록분할과 문자인식을 위해 블록의 좌표, 가로 세로의 길이 및 비율, 총 화소 수와 흑화소 수 및 흑화소의 비율을 구하여 문자블록, 수직 및 수평선 부분, 음영을 가진 블록으로 분류하는 방법[16] 등이 있다. 이 방법들은 이진영상에서 문자크기를 찾기 위해 흑화소 단위에서 문자 단위의 연결 길이를 만들어야 한다. 이 화소의 연결 길이를 찾기 위해서는 여러 번의 문서영상의 스캔이 필요하므로 시간이 많이 소요된다. 그리고, 블록의 유형분류를 위해 미리 문자의 크기나 블록의 흑화소 비율 등을 임계값으

로 미리 선정하여야 된다.

이진 영상에서 질감특징을 이용한 분류에서 D.Wang등[9]은 입력된 문서영상 한 페이지를 이진화를 하여 이 영상을 특정한 형태의 반복적 규칙을 갖는 질감영상으로 간주하고, 흑-백 쌍의 연결 길이와 흑-백-흑 조합의 연결 길이 행렬을 만들어 질감특징을 추출하여 유형을 분류하였다. 그러나 이 방법에서는 효과적인 분류를 위해 몇 차례의 실험을 거쳐 가장 분류가 잘되는 결정면상수를 구하여야 하며, 다섯 가지의 유형으로 분류하기 위해서는 네 개의 결정면 상수를 구하여야 한다. 그리고 적절한 임계값을 선정하더라도 이진화에 의한 정보손실은 피할 수 없다.

2.3 명암도 영상에서의 블록의 특징추출 및 유형분류

명암도 영상을 이용하여 질감특징을 추출하는 방법에는 Jain[17] 등이 게이버 필터(Gabor filter)를 적용하여 국부에너지(Local Energy)를 질감특징으로 추출하면서 유형분류도 동시에 되도록 하였다. 이 방법은 명암도 문서영상의 일정한 크기 마스크를 적용하여 문자와 비문자 블록을 구분하는 방법이지만, 마스크를 전체 영상에 주사하여야 하기 때문에 영상을 분할 및 분류하는데 매우 많은 시간이 소요된다. 또 마스크의 크기는 영상분할을 하기 위하여 질감특징을 찾는 성능에 매우 중요한 매개변수로 작용하여 너무 큰 마스크는 분할된 영역을 문지르기 때문에 분류능력이 떨어진다. 또한 다른 방법으로는 명암도 차이 행렬(Gray Level Difference Matrix, GLDM), 명암도 연결 길이 행렬(Gray Level Run Length Matrix, GLRLM), 그리고 공간 명암도 의존 행렬(Spatial Gray Level Dependence Matrix, SGLDM)을 이용하는 방법[10,11] 등이 있다. 이러한 방법들은 일정한 크기의 마스크로 전체영상을 주사하면서 한 화소에 대해 공간정보를 나타내는 행렬들을 구하여 질감특징을 추출하는 방법이다. 이들도 일반 사진영상의 질감추출 방법을 사용하면서 문서영상을 질감특징 추출 및 분류를 한다면 시간이 너무 많이 걸리고 문서영상의 줄이나 간격을 이용할 수 없는 단점이 있다.

지금까지의 연구를 살펴보면, 이진영상을 사용할 경우에는 이진화시에 정보손실의 문제점이 있고 명암도 문서영상을 사용할 때는 일반 사진영상에서와

같이 질감특징을 추출하는 데에 많은 시간이 걸린다. 따라서 명암도 문서영상에서 문서 특징을 이용한 빠른 분할방법이 필요하며, 좀 더 효과적으로 블록의 유형을 분류할 수 있도록 할 필요가 있다. 또한 문서영상을 가급적 상세히 분류하여 여러 가지 응용분야에 이용될 수 있도록 할 필요성도 있다.

3. 블록의 속성과 질감특징을 이용한 블록 분류

일반적인 문서영상은 문자, 표 및 그림 등 서로 다른 다양한 유형의 블록들로 복합 구성되어 있다. 최근 문서자동인식 시스템에서는 여러 가지 블록으로 분류하지만, 상세히 분류하지는 못한다. 또한 각 블록의 유형에 따라 최적의 알고리즘을 적용하여 인식 및 압축하는 것이 효과적이고, 또 문서구조 분석의 경우에도 상세한 구조 분석을 위해서는 문서를 구성하는 여러 유형의 블록들을 가급적 상세하게 분류하는 것이 요구된다. 따라서 전문적인 응용분야의 요구 조건을 만족시키기 위해서는 문서를 구성하는 여러 유형의 블록들을 잘 구별해낼 수 있는 효과적인 특징추출과 상세 유형분류에 관한 연구가 필요하다 [4,5].

본 장에서는 블록의 속성과 공간 명암도 의존 행렬을 이용하여 문서 영상을 효과적으로 블록을 분류하는 방법을 제시하였다. 문서 영상이 가지고 있는 배열이나 간격 등의 정보를 이용하기 위하여 먼저 문서영상을 이진화를 하여 평활화에 의한 방법으로 블록을 분할함으로써 각 블록의 위치정보와 속성을 빠르게 구할 수 있도록 하였다. 이와 같이 추출된 각 블록에 대한 위치정보와 대응되는 명암도 문서영상의 위치정보를 이용하여, 각 블록에서 단일방향의 공간 명암도 의존 행렬을 구한 후, 질감을 표현하는 평균, 분산, 상관계수, 에너지, 엔트로피, 대비 및 국부 균일성 등 일곱 가지의 질감특징을 추출하였다.

추출된 일곱 가지의 특징과 각 블록의 속성값 중에서 먼저 내부블록의 최대 높이를 최소거리 규칙을 이용하여 문자와 비문자 블록으로 구분하기 위하여 사용하고, 다시 공간 명암도 의존 행렬의 일곱 가지 특징을 최소거리 규칙에 적용하여 큰 문자, 작은 문자, 표, 그래픽 및 사진 블록으로 상세 분류할 수 있도록 하였다.

3.1 명암도 문서영상의 효과적인 블록분할

블록분할 과정은 특징추출을 위한 전처리 과정으로 문서영상 내에 존재하는 서로 다른 특징을 갖는 영역들을 문단단위로 분할한다.

본 논문에서는 적은 메모리로도 고속의 분할이 가능하도록 평활화에 의한 방법으로 블록을 분할하였다. 이 방법에서는 먼저 명암도 문서영상을 이진화를 한 후, 이진화된 영상에 대해 수평방향 평활화를 먼저 수행하고 이 영상에 대하여 다시 수직방향으로 평활화를 수행하여 최종적으로 사각형 블록으로 분할된 영상을 구한다. 그림 1에는 평활화를 이용한 블록분할의 각 단계별 수행 결과를 나타내었다. 또한 이 방법은 수평 방향으로 평활화 한 후, 각 블록의 속성인 내부 블록의 최대 높이를 추출하는 다음 과정에서도 효과적으로 이용되며, 문단 단위로 문서를 분할하므로 마스크를 이용하여 각 블록의 질감특징을 추출하는 과정에도 적합하다는 장점이 있다.

평활화에 의한 블록 분할 방법에서는 먼저 명암도 문서영상(그림 1(a))을 이진화한 후, 이 이진영상(그림 1(b))에서 수평방향 평활화(그림 1(c))를 하고, 이 영상을 다시 수직방향 평활화(그림 1(d))를 한 후, 이 영상에 외곽 탐색 방법을 이용하여 최종적인 사각형 블록(그림 1(e))과 같은 사각형 블록으로 분할된 영상을 구한다.

문서영상은 수평 평활화에 의해 문장과 같은 수평성분에 따라 분할되며, 수직 평활화를 통해 문단과 같은 수직성분에 따라 분할됨을 그림을 통해 확인할 수 있다. 이와 같이 평활화된 영상에 대하여 외곽 탐색 방법을 이용하여 각 블록의 사각 좌표를 구하고, 이렇게 구하여진 좌표들 중에 중첩이 되는 좌표들은 서로 병합함으로써 사각형 블록으로 분할됨을 그림을 통해 확인할 수 있다.

또한 수평 평활화를 통하여 각 세분되는 내부 블록의 높이 속성을 특징을 구할 수 있고, 질감특징을 이용한 일반영상 분할과 달리 본 논문에서는 이진영상에서 사각형 블록으로 분할된 영상의 4점의 좌표를 추출한 후 이를 명암도 문서영상의 블록분할에 이용함으로써 명암도 문서영상의 신속한 분할도 가능하도록 하였다.

3.2 문서영상에서 블록의 속성과 질감특징 추출

먼저 문서 블록의 유형 분류에 사용될 특징으로는

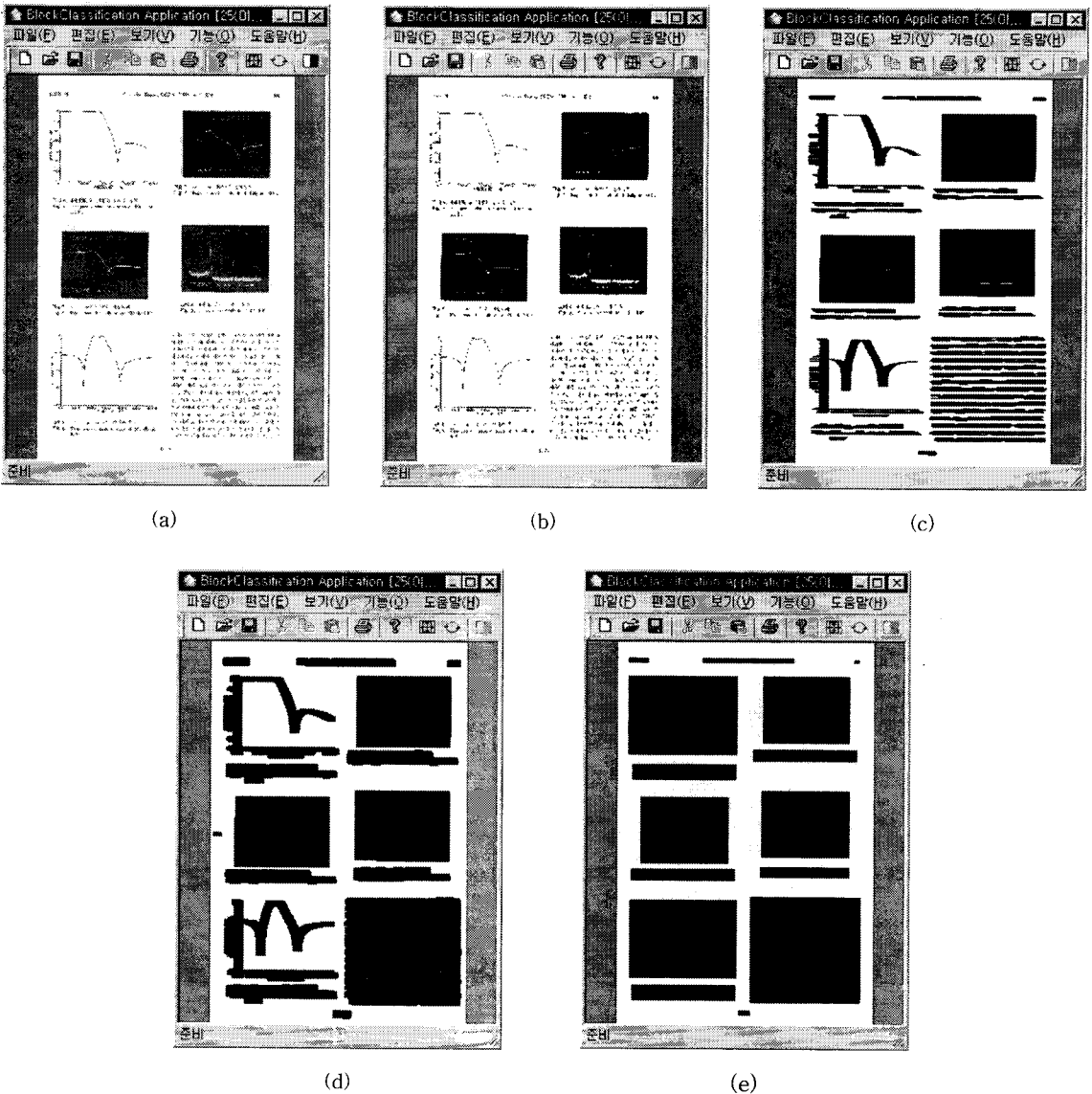


그림 1. 문서영상의 블록 분할 (a) 명암도 문서 원 영상 (b) 이진화된 영상 (c) 수평 평활화 된 영상 (d) 수직 평활화 된 영상 (e) 최종 사각형 블록으로 분할된 영상

각각의 블록에서 문자와 비문자를, 그리고 또 큰 문자와 작은 문자의 구별을 위해 블록별 내부 블록의 최대 높이를 특징으로 사용하였다. 내부 블록의 최대 높이는 수평방향 평활화로 처리된 영상에서 각 문장 단위의 높이를 계산한 후, 최대 높이 값을 구할 수 있었다. 그림 2에서는 블록에서 내부 블록의 최대 높이를 구하는 과정을 보여준다. 그림 2(a)는 입력 원 영상이며, 그림 2(b)는 수평으로 평활화를 수행한 문서의 하나의 블록이며, 그림 2(c)는 수평으로 평활화

된 영상에서 내부 블록으로 분할된 예이다. 이렇게 나누어진 블록은 각 내부 블록별 높이를 계산한 후, 최대 값을 블록 분할의 특징 값으로 사용하게 된다.

영상내의 화소들의 상호관계에 의해 영상의 성질을 결정하는 질감영상 분류나 분할에서는 질감영상을 효과적으로 표현할 수 있는 특징추출이 중요하다. 이와 같은 특징 추출에 영상 내에 존재하는 화소 값의 빈도수만을 표시하는 히스토그램을 이용한 단순 통계 특성은 질감영상 분류를 위한 특징으로는 적합

재구성가능 메쉬에서 이진영상과 경계코드 사이의 변환

(Transformation between a Binary Image and a Boundary Code on a Reconfigurable Mesh)

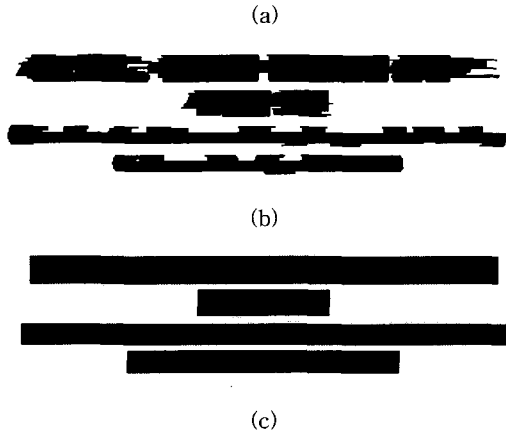


그림 2. 문서영상에서 블록의 속성 중 내부블록의 최대 높이 계산 과정 (a) 명암도 문서 원 영상 (b) 수평 평활화된 영상 (c) 문장단위 블록 분할된 영상

하지 않다. 영상내의 서로 다른 두 화소 사이의 공간 정보를 행렬로 표현하며, 이것을 나타내는 방법에는 명암도 차이 행렬, 명암도 연결 길이 행렬, 그리고 공간명암도 의존 행렬 등이 있다. 이들은 2차 통계특성을 가지고 있다. 본 연구에서는 공간 명암도 의존 행렬을 이용하여 질감특징을 추출하는 방법을 사용되었다.

공간 명암도 의존 행렬은 Haralick[18]에 의해 제안되었으며 이는 θ 방향에 대하여 거리 d 만큼 떨어진 두 화소 사이의 의존성을 이용하여 공간정보를 표현하는 방법으로 다음과 같이 표현할 수 있다.

사각형 모양의 영상이 있을 때, N_c 를 수평방향의 영상크기, N_r 를 수직방향의 영상 크기, N_g 를 최대 양자화 레벨이라 하면, 수평 방향의 공간 영역 L_c , 수직 방향의 공간 영역 L_r 및 화소값의 레벨 G 는

$$\begin{aligned} L_c &= \{1, 2, \dots, N_c\} \\ L_r &= \{1, 2, \dots, N_r\} \\ G &= \{1, 2, \dots, N_g\} \end{aligned} \tag{1}$$

와 같이 표현할 수 있다. 위와 같이 정의하면 $L_c \times L_r$ 위치에 G 를 할당하는 함수 I 는

$$I : L_c \times L_r \rightarrow G \tag{2}$$

와 같다. 그리고 공간명암도 의존 행렬은 상대적 빈도수로

$$\Phi(d, \theta) = [f(i, j | d, \theta)] \tag{3}$$

와 같이 나타낼 수 있다. 여기서 $f(i, j | d, \theta)$ 는 공간 명암도 의존 행렬의 한 원소를 의미하고 θ 방향으로 거리 d 만큼 떨어진 두 화소사이의 명암도 레벨 값이 i 에서 j 로 변하는 빈도수를 의미한다. 공간 명암도 의존 행렬을 구한 후, 사용한 전체 화소수로 공간 명암도 의존 행렬의 각 원소를 나누면 확률밀도 함수를 구할 수 있다. 그림 3은 공간 명암도 의존 행렬을 구

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

(a)

Gray Level	0	1	2	3
0	$N_o(0,0)$	$N_o(0,1)$	$N_o(0,2)$	$N_o(0,3)$
1	$N_o(1,0)$	$N_o(1,1)$	$N_o(1,2)$	$N_o(1,3)$
2	$N_o(2,0)$	$N_o(2,1)$	$N_o(2,2)$	$N_o(2,3)$
3	$N_o(3,0)$	$N_o(3,1)$	$N_o(3,2)$	$N_o(3,3)$

(b)

$$\begin{aligned} \Phi_h &= \begin{bmatrix} 4210 \\ 2400 \\ 1061 \\ 0012 \end{bmatrix} & \Phi_v &= \begin{bmatrix} 6020 \\ 0420 \\ 2222 \\ 0020 \end{bmatrix} & \Phi_{135} &= \begin{bmatrix} 2130 \\ 1210 \\ 3102 \\ 0020 \end{bmatrix} & \Phi_{45} &= \begin{bmatrix} 4100 \\ 1220 \\ 0241 \\ 0010 \end{bmatrix} \\ & 0^\circ & 90^\circ & 135^\circ & 45^\circ \end{aligned}$$

(c)

그림 3. 공간 명암도 의존 행렬의 예 (a) 4x4 영상 (b) 공간 명암도 의존 행렬 구조(최대 명암도 레벨은 3) (c) 4방향에 대한 공간 명암도 의존 행렬

하는 하나의 예이다. 그림 3(a)는 크기가 4×4이고 최대 명암도 레벨이 3인 원 영상의 명암도 레벨이고, 그림 3(b)는 공간 명암도 의존 행렬의 구조를 나타낸 것으로서 여기서 N_0 는 빈도수를 나타낸다. 그림 3(c)는 두 화소의 방향이 0°, 45°, 90°, 135°인 경우에 구한 각 공간명암도 의존 행렬을 나타낸다. 여기에서 두 화소사이의 거리는 모두 1로 하였다.

각각의 방향에 대하여 정규화 된 공간 명암도 의존 행렬 $p(i, j | d, \theta)$ 는

$$p(i, j | d, \theta) = \frac{f(i, j | d, \theta)}{\sum_{x=0}^{N_0} \sum_{y=0}^{N_0} f(x, y | d, \theta)} \quad (4)$$

이 되며 확률을 의미한다. 이는 공간 명암도 의존 행렬의 각 원소의 확률로서 이를 모두 합하면 1이 된다.

$f(i, j | d, \theta)$ 를 0° 방향에 대하여 표시하면

$$f(i, j | d, 0^\circ) = N_0 \{ (k, l), (m, n) \in (Lr \times Lc) \times (Lr \times Lc) \mid \begin{matrix} k-m=0, \\ |l-n|=d, \\ I(k, l)=i, \\ I(m, n)=j \end{matrix} \} \quad (5)$$

과 같다. 0°, 45°, 90°, 135 방향에 대하여 공간 명암도 의존 행렬로부터 얻을 수 있는 질감특징은

(1) 평균

$$\mu(\Phi(d, \theta)) = \mu = \sum_{i=0}^{N_0} i \sum_{j=0}^{N_0} p(i, j | d, \theta) \quad (6)$$

(2) 분산

$$\sigma^2(\Phi(d, \theta)) = \sigma^2 = \sum_{i=0}^{N_0} (i - \mu)^2 \sum_{j=0}^{N_0} p(i, j | d, \theta) \quad (7)$$

(3) 상관계수

$$C(\Phi(d, \theta)) = \frac{\sum_{i=0}^{N_0} \sum_{j=0}^{N_0} (i - \mu)(j - \mu)p(i, j | d, \theta)}{\sigma^2} \quad (8)$$

(4) 에너지

$$E(\Phi(d, \theta)) = \sum_{i=0}^{N_0} \sum_{j=0}^{N_0} p^2(i, j | d, \theta) \quad (9)$$

(5) 엔트로피

$$H(\Phi(d, \theta)) = \sum_{i=0}^{N_0} \sum_{j=0}^{N_0} -p(i, j | d, \theta) \log p(i, j | d, \theta) \quad (10)$$

(6) 대비

$$T(\Phi(d, \theta)) = \sum_{i=0}^{N_0} \sum_{j=0}^{N_0} (i - j)^2 p(i, j | d, \theta) \quad (11)$$

(7) 국부균일성

$$I(\Phi(d, \theta)) = \sum_{i=0}^{N_0} \sum_{j=0}^{N_0} (1 + (i - j)^2)^{-1} p(i, j | d, \theta) \quad (12)$$

등이다. 본 연구에서는 문서의 특징을 잘 반영하는

0° 방향의 거리벡터 (0,1)인 공간 명암도 의존 행렬만을 질감특징 추출에 이용하여 식(6)에서부터 식(12)에 나타낸 일곱 가지 질감특징을 추출함으로써 속도가 빠르고 효과적인 문서영상의 블록분류가 가능하도록 하였다.

3.3 블록의 속성과 질감특징을 이용한 블록분류

본 논문에서는 블록의 유형을 크게 문자 블록, 문자 포함 블록 및 비문자 블록으로 구분하고, 문서의 구조 해석을 효율적으로 할 수 있게 하기 위하여 이를 다시 큰 문자, 작은 문자, 표, 그래픽 및 사진 블록의 다섯 가지 유형으로 상세히 분류하였다. 문자 블록은 문서에서 표제 및 내용을 가려내기 위해 큰 문자, 작은 문자의 두 가지 블록으로 세분하였고, 문자 포함 블록은 직선과 문자만으로 이루어진 표 블록과 도형과 문자 및 부호 및 선으로 이루어진 그래픽 블록으로 세분하였으며 사진은 비문자 블록으로 구분하였다.

유형분류는 추출된 일곱 가지 질감특징과 블록의 속성을 최소거리 규칙에 적용하여 학습, 인식과정을 통하여 각 블록을 분류함으로써 유형에 따른 임계값과 결정면 상수를 구해야 하는 어려움을 해소하였다.

본 논문에서 사용한 최소거리 규칙은 패턴분포에 관한 정보로 식별함수를 계산하는 대신에 미리 저장 및 학습해 놓은 기본패턴과의 거리를 계산하여 가장 가까운 기준패턴의 클래스를 미지패턴의 클래스로 결정하는 방법으로, 이와 같은 기본원리에 따른 최소거리 규칙을 의미한다. 학습데이터의 특징 추출에서 추출된 특징들을 먼저 학습시키고, 분류하고자 하는 블록의 특징이 입력되면, 각 특징별로 최소거리 규칙에 의거하여 블록을 분류한다.

제안한 유형분류 순서도는 그림 4와 같다. 스캐너를 이용해 입력된 명암도 문서영상을 이진화한 후, 먼저 수평방향으로 평활화를 수행하고 수직방향으로 평활화를 수행한다. 이와 같이 평활화된 영상에서 외곽 탐색을 통하여 블록의 좌표를 구하고, 서로 중첩되는 블록을 병합함으로써 사각형 블록으로 분할된 영상을 구하게 된다. 사각형 블록으로 분할된 영상의 각 사각형 좌표를 추출한 후, 이를 명암도 문서영상의 블록분할에 이용함으로써 신속한 명암도 문서영상 분할이 가능하도록 하였다. 분할된 각 블록의 유형분류를 위한 특징에는 수평방향 평활화된 영상

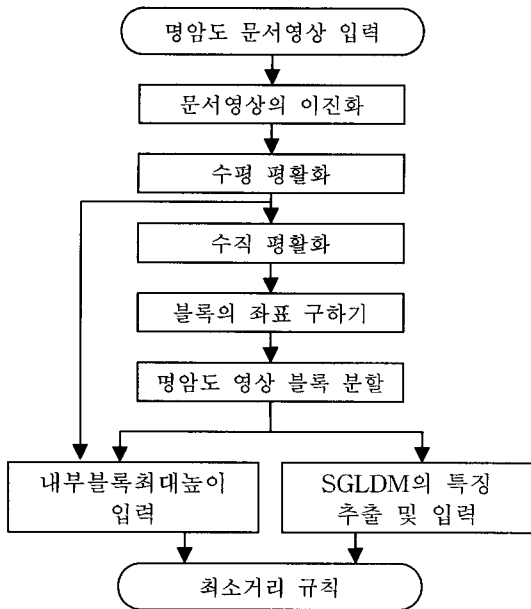


그림 4. 제안한 유형분류 절차

에서 추출해 낸 블록별 내부블록의 최대 높이와 거리 벡터가 (0,1)인 공간 명암도 의존 행렬로부터 추출된, 식(6)에서부터 식(12)에 나타난 일곱 가지 질감특징을 이용하였으며, 이 특징을 최소거리 규칙을 통하여 효과적인 블록 유형 분류가 가능하도록 하였다.

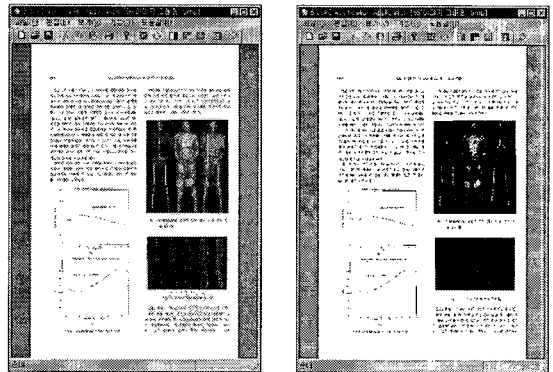
4. 실험 및 고찰

본 연구는 Windows 환경의 PC의 Microsoft Visual C/C++ 6.0 하에서 문서영상의 블록의 속성과 질감특징을 이용하여 블록분류 실험하였다. 실험에 사용한 영상은 Hewlet Packard Scanjet 5300C 스캐너를 이용하여 대한 전자 공학회 논문지와 한국 정보과학회 논문지에서 300dpi의 256명암도 영상으로 스캔하여 사용하였다. 실험에 사용한 표준 데이터는 58개의 문서, 총 733개의 블록으로 하였다. 평활화 방법을 이용하여 블록을 분할하였으며 평활화의 간격은 가로 65화소, 세로 34화소로 하였다.

실험에 사용한 대표적인 명암도 문서영상은 그림 5(a)에 나타내었으며, 이를 이진화한 결과 영상은 그림 5(b)와 같다. 수평 평활화의 결과는 그림 5(c)에 도시하였으며, 수평 평활화된 영상에 수직 평활화한 영상은 그림 5(d)에, 각 영역을 외곽 탐색의 방법으로 순번화하고 각 영역의 좌측 상단 좌표와 우측

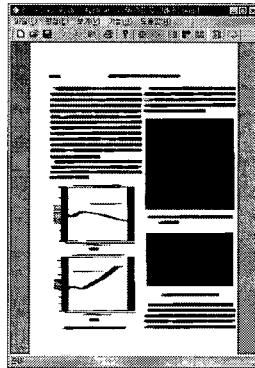
하단 좌표를 구하여 블록으로 만든 결과는 그림 5(e)에 나타내었다. 그림 5(f)에는 작은 문자, 그래픽, 그리고 사진 블록으로 유형 분류되었다.

평활화 간격은 문단사이 간격이나 블록간 간격 등

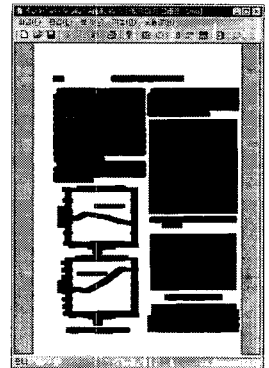


(a)

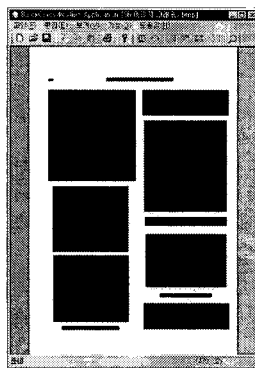
(b)



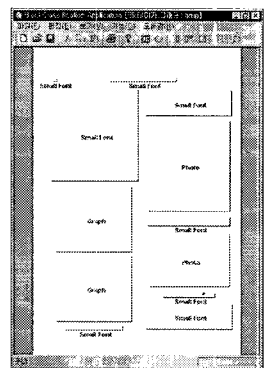
(c)



(d)



(e)



(f)

그림 5. 제안한 방법에 의한 블록분할과 유형분류 (a) 원 영상 (b) 이진화된 영상 (c) 수평 평활화된 영상 (d) 수직 평활화된 영상 (e) 평활화 방법에 의해 분할된 영상 (f) 제안한 방법에 의해 분류된 결과

을 고려하여 가로 세로 각각 65, 34화소로 설정함으로써 문장단위가 아닌 문단 단위로 효과적인 블록분할이 됨을 그림 5(e)를 통해 확인할 수 있었다.

또한 그림 5에서 보는 바와 같이 질감특징을 이용한 일반영상 분할과 달리 본 논문에서는 이진화된 영상에서 분할된 사각형 블록의 좌표를 추출한 후, 이를 명암도 문서영상의 블록분할에 이용함으로써 신속한 명암도 문서영상 분할이 가능하도록 하였다.

또 다른 특징 값으로 사용된 블록의 속성은 이진

화된 영상에서 수평 평활화를 통하여 문장 단위의 블록으로 분할한 후, 각 블록의 내부블록의 높이 중에서 가장 큰 값을 해당 블록의 속성으로 사용함으로써 문자 블록과 비문자 블록으로 크게 분류하였다. 그림 6부터 그림 8까지는 문자 블록과 비문자 블록들의 영상으로부터 블록의 속성을 추출해내는 과정을 나타내었다.

그림 5(d), 그림 6(d) 그리고 그림 8(c)를 보면, 내부 블록의 최대높이 값이 서로 차이가 많이 나타난

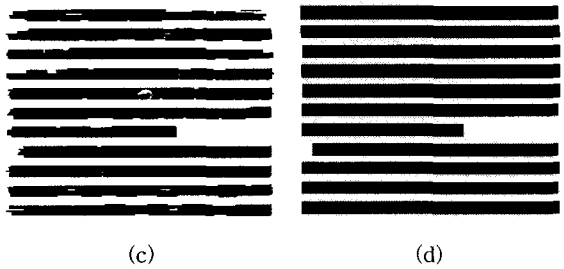
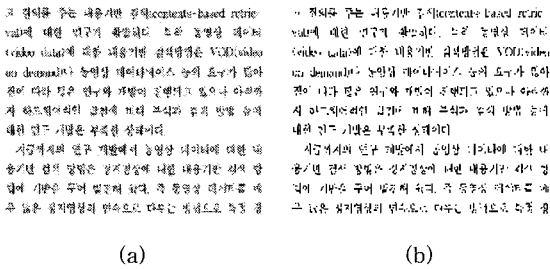


그림 6. 작은 문자 블록의 속성 중 내부블록의 최대 높이 계산 과정 (a) 명암도 문서 원 영상 (b) 이진화된 영상 (c) 수평 평활화된 영상 (d) 문장단위 블록 분할된 영상

동적 영역 히스토그램을 사용한 동영상 데이터의 컷검출 방법
(Cut Detection Method for Video Data using Dynamic Region Histogram)

동적 영역 히스토그램을 사용한 동영상 데이터의 컷검출 방법
(Cut Detection Method for Video Data using Dynamic Region Histogram)

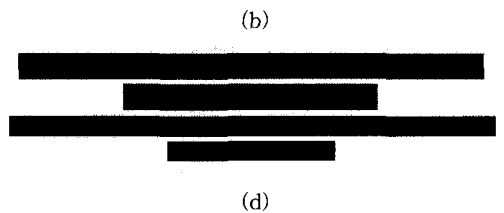
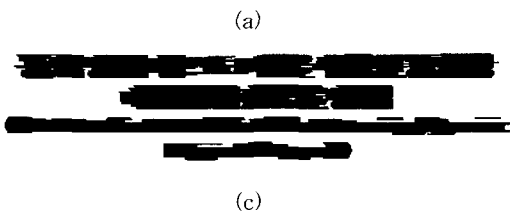


그림 7. 큰 문자 블록의 속성 중 내부블록의 최대높이 계산 과정 (a) 명암도 문서 원 영상 (b) 이진화된 영상 (c) 수평 평활화된 영상 (d) 문장단위 블록 분할된 영상

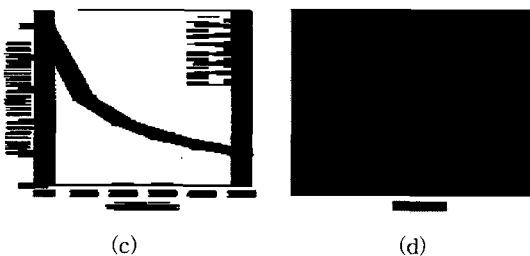
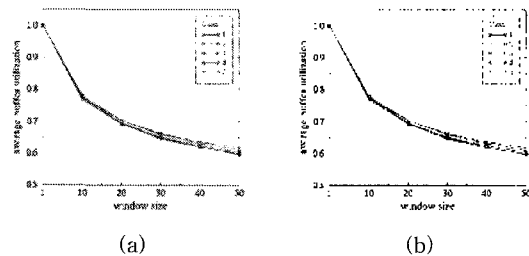


그림 8. 비문자(그래프) 블록의 속성 중 내부블록의 최대높이 계산 과정 (a) 명암도 문서 원 영상 (b) 이진화된 영상 (c) 수평 평활화된 영상 (d) 문장단위 블록 분할된 영상

것을 확인할 수 있다. 그리고 그림 8(d)와 같이 분할된 각 블록으로부터 특징추출에 필요한 행렬들을 만들었다. 이 행렬들을 이용하여 추출된 특징 값들과 그림 6에서 그림 8까지에서 볼 수 있는 블록의 속성을 추출하였다. 먼저 블록의 속성 특성을 이용하여 문자 블록과 비문자 블록으로 분류하였으며, 비문자 블록으로 구분된 블록들은 최소거리 규칙을 사용하여 도표, 그래픽 및 사진블록으로 상세 분류하였다.

분류할 유형의 종류는 작은 문자, 큰 문자, 표, 그래픽 및 사진 블록 등 다섯 가지이다. 학습에는 실험 전체 데이터 733개의 데이터 블록 중에서 각 유형별로 반을 사용하였고, 그 나머지는 분류실험에 사용하였다.

본 연구과정에서는 수평방향의 공간 명암도 의존 행렬로부터 평균, 분산, 상관계수, 에너지, 엔트로피, 대비 및 국부 균일성의 일곱 가지 특징을 사용하였으나, 도표, 그래픽 및 사진 블록은 일곱 가지 질감특징 중에서 평균과 국부 균일성에 의해 잘 분류될 수 있음을 알 수 있었다.

본 저자가 참고문헌[1]에서 사용한 실험결과를 참고적으로 나타내면 표 1과 같다. 실험 데이터의 개수가 이 논문과 서로 다른 것은 실험하는 방법도 문서의 속성을 특징으로 새롭게 추출하여 실험한 때문이다.

표 2에서 볼 수 있듯이 특히 이 방법은 먼저 문자 부분을 미리 문서영상의 속성을 이용하여 분류하기 때문에 작은 문자, 큰 문자 등에서 100%의 정확한 분류를 해낼 수 있다는 것이 장점이다. 또한 상세 분류에서는 질감방법의 특징을 사용함으로써 문자이외의 영상도 잘 분류를 하는 것으로 파악된다. 그 예로 사진 블록도 오분류가 없었으며, 하나의 도표 블록이 그래픽 블록으로, 그래픽 블록의 하나가 도표 블록으로 오분류 되었다. 질감특징만을 사용한 방법 보다는 전체적으로도 분류의 성능도 뛰어난을 알 수 있다.

표 1. 기존의 방법들의 분류결과

방법 \ 종류	Wang의 방법	GLDM을 이용한 방법	GLRLM을 이용한 방법	SGLDM을 이용한 방법
작은 문자	78/79	64/79	59/79	73/79
큰 문자	32/32	29/32	28/32	30/32
도 표	43/44	40/44	40/44	38/44
그래픽	87/94	89/94	92/94	94/94
사 진	54/55	55/55	55/55	54/55

표 2. 블록의 속성과 질감특징을 이용한 블록분류 결과

블록 유형	실험 대상 블록 수	정분류 블록 수	인식률
작은 문자	250	250	100%
큰 문자	11	11	100%
도 표	37	36	97%
그래픽	33	32	96%
사 진	19	19	100%
결 과	350	348	98.60%

블록의 속성과 질감특징 중 공간 명암도 의존 행렬을 이용한 분류 실험의 결과는 표 2에 정리하였다. 블록을 이진화하여 분할하는데 약 2~3초의 시간이 소요되었으며, 특징을 추출하는데 약 0.7초, 분류하는 과정은 0.1ms미만의 시간이 소요되었다.

5. 결 론

제안한 블록분류 방법에서는 문서의 명암도 영상을 입력하여 이진화한 다음 평활화 기법을 적용하여 각 블록의 위치정보를 찾아내어 최외각 사각형 좌표를 추출 한 후, 이를 명암도 문서영상의 블록 분할에 적용함으로써 신속한 명암도 문서영상 분할이 가능하였으며, 또한 블록의 속성 특징을 추출하는 데도 이용하였다.

이 블록들에서 수평방향의 공간 명암도 의존 행렬을 구하여 평균, 분산, 상관계수, 에너지, 엔트로피, 명암대비 및 국부균일성 등의 7가지 질감특징을 추출하였다.

블록의 내부블록 최대 높이 속성을 사용함으로써 문자와 비문자 블록을 미리 분류하였으며, 공간 명암도 의존 행렬의 질감특징만으로 분류하였을 때의 작은 문자와 도표 블록의 오분류도 줄일 수 있었다.

상세분류에는 7가지 질감특징을 최소거리 규칙에 적용하여 최종적으로 작은 문자, 큰 문자, 도표, 그래픽 및 사진 등 다섯 가지 유형으로 나누었다.

전체 분류 실험대상 350개 블록에 실험한 결과, 하나의 도표 블록이 그래픽 블록으로 분류되었고, 그래픽 블록 하나가 도표 블록으로 분류되는 오분류가 나타났으며, 전체 인식률은 98.6%의 성공률을 나타내었다.

참 고 문 헌

[1] 김중수, "SGLDM을 이용한 문서영상의 블록 분류," 한국멀티미디어 학회 논문지, Vol. 8, No. 10, pp. 1347-1359, 2005.

[2] 장대근, 황찬식 "이미지 필터와 제한조건을 이용한 문서영상 구조분석," 한국정보처리학회 논문지, Vol. 9-B, pp. 311-318, 2003.

[3] F. M. Wahl, K. Y. Wong, and R. G. Casey, "Block Segmentation and Text Extraction in Mixed Text/Image Document," *Computer Vision, Graphics, and Image Processing*, Vol. 22, pp 375-390, 1982.

[4] L. O'Gorman and R. Kasturi, *Document Image Analysis*, IEEE Computer Society Press, Washington, 1995.

[5] 이성환, 문자인식 - 이론과 실제, 홍릉과학출판사, 1993.

[6] G. Nagy, S. Seth, and M. Viswanathan, "A Prototype Document Image Analysis System for Technical Journals," *Computer*, 1992.

[7] T. Pavlidis and J. Zhou, "Page Segmentation and Classification," *Computer Vision, Graphics, and Image Processing: Graphical Model and Image Processing*, Vol. 54, No. 6, pp. 484-496, 1992.

[8] 이인동, 권오석, 김태균, "문서영상에서 문자와 비문자의 분리추출," 한국정보과학회 논문지, Vol. 17, No. 3, pp. 247-257, 1990.

[9] D. Wang and S. N. Srihari, "Classification of Newspaper Image Blocks Using Texture Analysis," *Computer Vision, Graphics and Image Processing*, Vol. 47, pp. 327-352, 1989.

[10] 김중수, "공간명암도의존 행렬의 2차통계 질감 특징을 이용한 문서영상의 블록분류," 경북대학교 박사학위논문, 1996.

[11] Linda Shapiro and George Stockman, "*Computer Vision*," pp. 235-247, 2000.

[12] B. Julesz, "Visual Pattern Discrimination," *IEEE Trans. Inform. Theory*, Vol. IT-8, No. 1, pp. 84-92, 1962.

[13] 김상운, 식별 알고리즘을 중심으로 한 패턴인식

입문, 홍릉과학출판사, 1995.

[14] 이성환, 패턴 인식의 원리, 홍릉과학출판사, 1994.

[15] 장명옥, 천대녕, 양현승, "연결화소를 이용한 문서영상의 분할 및 인식," 한국정보과학회 논문지, Vol. 20, No. 12, pp. 1741-1751, 1993.

[16] 이인동, 권오석, 김태균, "블록영상의 추출 알고리즘," 한국정보과학회 논문지, Vol. 18, No. 2, pp. 218-225, 1991.

[17] A. K. Jain and Y. Zhong, "Page Segmentation Using Gabor Filters for Automatic Document Processing," *Machine Vision and Application*, Vol. 5, pp. 169-184, 1992.

[18] R. M. Haralick, "Statistical and Structural Approaches to Texture Classification," *Proceeding of the IEEE*, Vol. 67, No. 5, pp. 786-804, 1979.



장 영 내

1994년 3월 안동대학교 공과대학 컴퓨터공학과 입학
 2001년 8월 안동대학교 공과대학 컴퓨터공학과 석사
 2001년 8월 ㈜써니벨시스템즈 정보기술연구소 Web Accelerator 개발 연구

원

2002년 11월~현재 ㈜그린벨시스템즈 정보기술연구소 IETM 개발팀 과장
 연구분야 : 영상처리, 웹 컴퓨팅, 네트워크 보안



김 중 수

1982년 경북대학교 전자공학과 학사
 1984년 경북대학교 대학원 전자공학과 석사
 1996년 경북대학교 대학원 전자공학과 전자계산전공 박사

2002년 UTA 교환교수
 1987년~현재 안동대학교 컴퓨터공학과 교수
 관심분야 : 영상처리, 데이터베이스



이 철 희

- 1995년 2월 경북대학교 전자공학과 졸업, 공학사
- 1997년 2월 경북대학교 대학원 전자공학과 졸업, 공학석사
- 2000년 8월 경북대학교 대학원 전자공학과 졸업, 공학

박사

2000년 3월~2003년 8월 경운대학교 컴퓨터공학부 전임강사

2002년 6월~현재 (사단법인)한국화상학회 총무이사

2003년 9월~현재 안동대학교 컴퓨터공학전공 조교수

2004년 7월~현재 산업자원부 기술표준원 ISO/TC-130 전문위원

연구분야 : 칼라영상처리, 프린팅, 디스플레이 색재현