

다해상도 영상과 개선된 RBF 네트워크를 이용한 계층적 영문 명함 인식

김 광 백[†] · 김 영 주[†]

요 약

본 논문은 영문 명함의 다해상도 영상을 이용한 계층적 영상 처리를 통해 문자를 추출하고 개선된 신경망 기법을 이용하여 문자를 인식하는 새로운 계층적 명함 인식 알고리즘을 제안하였다. 계층적 인식 알고리즘은 명함 인식 과정을 구성하는 각 처리 단계별로 처리 시간을 단축함과 동시에 성능 향상을 위해 입력된 명함 영상을 해상도가 서로 다른 영상들로 분리하여 적용한다. 우선 1/3배 축소 영상에 가로 스미어링 기법을 적용하여 명함 영상 내에서 문자들을 포함하는 문자열 영역을 추출하고, 문자열 영역으로부터 개별 문자를 추출하기 위하여 1/2배 축소 영상에 세로 스미어링 및 윤곽선 추적 마스크를 적용한다. 마지막으로 추출된 문자를 인식하기 위해서 문자의 형태학적 특성을 그대로 가지고 있는 원 영상을 사용하며, 다양한 형태를 가진 명함상의 문자를 인식하기 위해 ART1 기반의 개선된 RBF 네트워크를 제안하고 인식 과정에 적용하였다. 제안된 인식 알고리즘을 실제 영문 명함 영상에 적용하여 실험한 결과, 기존의 방법과 비교하여 문자 추출 및 인식 성능이 크게 향상됨을 확인하였다.

Hierarchical Recognition of English Calling Card by Using Multiresolution Images and Enhanced RBF Network

Kwang-Baek Kim[†] · Young-Ju Kim[†]

ABSTRACT

In this paper, we proposed the novel hierarchical algorithm for the recognition of English calling cards that processes multiresolution images of calling cards hierarchically to extract individual characters and recognizes the extracted characters by using the enhanced neural network method. The hierarchical recognition algorithm generates multiresolution images of calling cards, and each processing step in the algorithm selects and processes the image with suitable resolution for lower processing overhead and improved output. That is, first, the image of 1/3 times resolution, to which the horizontal smearing method is applied, is used to extract the areas including only characters from the calling card image, and next, by applying the vertical smearing and the contour tracking masking, the image of a half time resolution is used to extract individual characters from the character string areas. Lastly, the original image is used in the recognition step, because the image includes the morphological information of characters accurately. And for the recognition of characters with diverse font types and various sizes, the enhanced RBF network that improves the middle layer based on the ART1 was proposed and applied. The results of experiments on a large number of calling card images showed that the proposed algorithm is greatly improved in the performance of character extraction and recognition compared with the traditional recognition algorithms.

키워드 : 계층적 인식 알고리즘(Hierarchical Recognition Algorithm), 스미어링 기법(Smearing Method), 윤곽선 추적 마스크(Contour Tracking Masking), 개선된 RBF 네트워크(Enhanced RBF Network)

1. 서 론

최근 핸드폰 또는 PDA 등과 같은 모바일 장치에 카메라 장치가 일반화되고, 투명성을 가진 포토 센서를 통해 문서를 스캔하는 입력장치 등이 개발됨에 따라 휴대용 모바일 장치가 단순한 영상 입력 기능뿐만 아니라, 영상의 처리 기능까지도 제공하도록 하는 연구가 활발히 이루어지고 있다.

특히, 개인의 정보를 가지는 명함이나 쪽지 등을 카메라를 통해 영상으로 입력받아 개인 정보를 항목별로 추출하고 인식하여 문자로의 변환 및 저장, 관리할 수 있는 연구도 활발히 이루어지고 있다.

명함 인식[1]은 크게 명함 영상에서 의미 있는 정보의 추출과 이런 정보들을 인식하는 부분으로 구성된다. 명함 영상에서의 정보 추출 과정은 잡음 제거 등의 전처리, 문자열 추출, 개별 문자 추출 등의 작업을 수행하여 명함 영상에 포함되어 있는 개별 문자를 추출하고, 정보 인식 과정에서

[†] 정 회 원 : 신라대학교 컴퓨터공학과 교수
논문접수 : 2003년 5월 22일, 심사완료 : 2003년 8월 1일

추출된 문자들에 대해 다양한 인식 기법을 적용하여 코드화된 정보로 변환한다.

본 논문에서는 명함 영상을 다해상도로 분리하고 윤곽선 추적(contour tracking) 알고리즘과 스미어링(smearing) 알고리즘을 적용하여 잡영과 기울어짐, 문자열의 정렬 방향 등에 상관없이 정확하고 빠르게 문자를 추출하는 방법을 제안하고, 문자 인식 단계에서는 잡음에 강하고 새로운 패턴에 잘 적용할 수 있는 RBF(Radial Basis Function) 네트워크를 개선하여 개별 문자의 인식에 적용하였다.

본 논문은 2장에서 관련 연구를 기술하고, 3장에서 본 논문에서 제안하는 명함 인식 알고리즘을 구체적으로 설명하며, 4장에서 제안된 알고리즘에 대해 성능을 평가한 다음, 5장에서 결론으로 마무리한다.

2. 관련 연구

명함 영상에서 문자열 및 개별 문자를 추출하기 위한 방법으로는 히스토그램 정보를 이용한 방법[2], 윤곽선 추적 알고리즘을 이용한 방법[3,4], 스미어링 방법[5-7] 등이 있다.

히스토그램 정보를 이용한 방법[2]은 텍스트 영역을 화소 행(pixel row) 방향으로 스캔하면서 각 행의 흑화소의 개수가 정해진 임계 값(threshold)보다 크면 문자열에 속하는 것으로 간주하고 그렇지 않은 경우에는 문자열 사이의 공백으로 간주하여 문자열 영역을 분리하여 추출한다. 그리고 추출된 각 문자열 영역에 대해 화소 열(pixel column) 방향으로 흑화소의 개수에 대한 히스토그램(수직 히스토그램)을 구한 후, 히스토그램의 결과로 백화소의 연속된 열의 개수가 임계 값 이상이면 문자 사이의 공백으로 간주하여 문자를 추출하는 방법이다. 이 방법은 문자의 기울어짐 정도가 심한 경우와 문자들의 정렬 방향을 모르는 경우, 그리고 잡영이 많은 경우 등에는 추출율이 저하된다.

윤곽선 추적 알고리즘을 이용한 방법[3,4]은 윤곽선을 추출하기 위해 디자인된 마스크를 이용하여 문자들을 추출한다. 그리고 추출된 문자가 임계 값보다 작으면 잡영으로 간주한다. 이 방법은 기울어짐, 잡영, 문자열의 정렬방향과 상관없이 문자를 정확히 추출한다. 그러나, 시간이 많이 소요되고, 'l'와 같은 문자의 경우에는 위에 위치하는 점이 잡영으로 간주되어 오인식할 수 있다.

스미어링 알고리즘을 이용한 방법[5-7]은 크게 RLSA(Run Length Smoothing Algorithm) 방법[5]과 RXYC(Recursive X-Y Cuts) 방법[6]이 있다. RLSA 방법은 두 흑화소의 사이에 임계 값보다 작거나 같은수의 백화소가 있을 때, 이 백화소들을 흑화소로 대체시키는 방법이다. 예를 들어, 흑화소가 1, 백화소가 0, 그리고 임계 값이 3이라고 했을 때, 영상의 데이터 열이 "1000110000011 00100001"이면 RLSA 방법을 적용하였을 때의 영상 데이터 열은 "111111000001111

100001"이 된다. 그리고 문서 영상을 종방향과 횡방향으로 RLSA 방법을 적용한 다음, 2개의 결과 영상을 AND 연산으로 결합하면 문서 영상이 블록화된다. RXYC 방법[6]은 문서 영상을 수평 방향과 수직 방향으로 스캔한 다음, 스캔한 방향에 대한 흑화소의 수가 임계 값보다 큰 영역을 절단하는 방법이다. 따라서, RLSA 방법은 대블록 보다는 소블록에 효과적이며, RXYC 방법은 소블록 보다는 대블록에 효과적이다. 이 방법은 잡영이 많은 경우에 잘못된 블록을 추출할 수 있다.

추출된 문자들을 인식하기 위하여 원형 비교 방법(template matching method), 통계적 방법(statistical method), 구조적 방법(structural method), 그리고 인공 신경망(artificial neural networks)을 이용한 방법 등이 제안되었다[8-11].

원형 비교 방법은 인식 대상이 되는 모든 문자의 영상을 2차원 배열로 저장하고, 인식하고자 하는 입력 문자의 영상을 저장된 각각의 문자 영상과 대응되는 화소 단위로 비교하여 불일치된 화소의 개수가 가장 작은 문자를 입력된 문자의 결과로 판단하는 방법이다[8]. 이 방법은 단순하고 하드웨어 구현이 적합한 반면에, 명함과 같은 다양한 폰트를 가지는 문자들의 개수가 많은 경우에는 인식 시간이 길어지고, 특정 활자체의 문서에만 적용 가능하다.

통계적 방법은 인식 대상인 각 문자에 대한 충분히 많은 문자 영상을 추출하고, 각각의 문자 영상을 모두 분류할 수 있는 N 개의 특징 값을 추출하여 N 차원의 공간의 벡터로 표현한다. 이 벡터들의 평균 벡터를 구해 문자에 대한 특성 벡터를 저장하여 가장 가까운 거리에 있는 특성 벡터에 해당 문자를 인식하는 방법이다[9]. 이 방법의 가장 중요한 개념은 서로 가까운 벡터끼리는 밀집되고 서로 다른 벡터끼리는 멀리 떨어져 나타날 수 있는 특징을 선택하여야 하고, 빠른 속도로 특징을 계산할 수 있어야 한다. 통계적 방법은 인식속도가 빠르고 이론적 정립이 잘되어 있는 반면에 유사한 글자가 많고 복잡하고 계층적인 구성 방식을 갖는 문자에 대해서는 본질적인 특징을 계산하기가 어렵다.

구조적 방법은 문자의 구성 원리에 입각하여 자획 등과 같은 문자를 구성하는 기본 요소와 그들 간의 연관성을 추론하여 문자를 인식하는 방법이다[10]. 이론적으로는 잘 정립되어 있으나 다양한 폰트를 가지는 경우에는 인식 시간이 오래 걸리는 단점이 있다.

인공 신경망을 이용한 방법은 인간의 신경망 조직을 모델로 하여 많은 수의 단순한 프로세서들을 망으로 연결한 시스템을 사용하는 방법[11]으로, 학습에 의해 새로운 패턴에 적응할 수 있으며, 패턴의 국부적인 변형 및 잡음에 민감하지 않는 장점을 가지는 반면, 패턴의 크기가 큰 경우에는 학습 시간이 많이 소요되고 인식 후보 대상의 개수가 많은 경우에는 성능이 저하된다.

오늘날 신경망 모델의 커넥션니즘(connectionism)에 관해

관심이 집중되고 있다. 그 이유는 신경망 모델의 학습 능력 때문이다. 신경망 학습 알고리즘의 학습 시간과 일반화 능력을 향상시키기 위한 연구가 진행되어 왔다. 그 결과, 통계학의 다변량 분석(multivariate analysis) 및 보간(interpolation) 문제 해결에 이용되었던 RBF를 신경망 모델을 구성하는데 이용함으로써 RBF 네트워크가 제안되었다[12]. RBF 네트워크는 비선형 모델링등에 적용되고 있다. 오류 역전파(error-backpropagation)와 같은 다층 퍼셉트론(multilayer perceptron)의 연결 가중치들이 지역 최소화에 위치하여 학습 패턴들이 분류되지 않는 경우를 방지할 수 있으며 학습 시간을 단축시킬 수 있다. RBF 네트워크의 구조는 입력층, 중간층, 출력층의 3개의 계층으로 구성되며 RBF 네트워크의 중간층으로 적용되는 클러스터링 방법으로는 SOM(Self Organizing Map)이나 FCM(Fuzzy C-Means) 알고리즘 등이 있다[13, 14].

본 연구에서는 입력층과 중간층 사이의 학습 구조로 ART1을 개선하여 RBF 네트워크의 중간층으로 적용하는 개선된 RBF 네트워크를 제안하고 영문 명함 인식에 적용한다.

3. 계층적 접근과 개선된 RBF 네트워크를 이용한 명함 인식

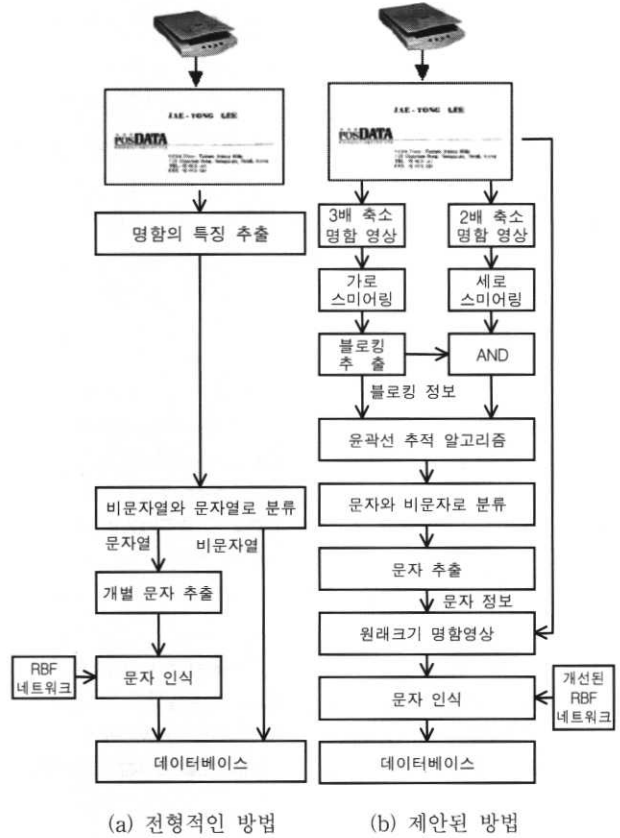
3.1 명함 인식 시스템

명함 인식[1]은 크게 명함 영상으로부터 의미가 있는 정보를 추출하는 부분과 추출된 정보들을 인식하는 부분으로 구성된다. 일반적인 명함 인식 시스템은 (그림 1)(a)와 같은 구성도로 이루어진다.

일반적인 명함 인식 방법은 문자 식별이 가능할 정도의 큰 명함 영상을 이용하여 잡음을 제거하기 위한 전처리 단계, 문자열 추출 단계, 개별 문자 추출 단계로 구성되는데 수행 시간이 많이 소요되는 단점이 있다. 이에 본 논문에서는 전처리 및 추출 시간을 단축하고, 기존의 문자열 추출 방법들과 개별 문자의 추출 방법들의 문제점을 개선할 수 있는 방법을 제안한다.

명함 영상에서 정보를 추출하기 위해서는 명함 영상을 다해상도 영상으로 분리하는데, 문자열 추출을 위한 1/3배 축소 영상, 개별 문자 추출을 위한 1/2배 축소 영상, 인식을 위한 원본 명함 영상으로 분리한다. 축소된 영상에 대해서 블로킹 영역 추출과 개별 문자 추출을 수행함으로써 시간을 단축할 수 있을 뿐만 아니라 하드웨어로 병렬 처리 구현이 가능해진다. 그리고, 기존의 문자열 추출 알고리즘의 문제점을 보완하기 위해 잡음, 기울어짐, 문자열의 정렬 방향 등에 상관없는 윤곽선 추적 알고리즘과 스미어링 알고리즘을 이용한다. 그리고 시간 소요가 많은 윤곽선 추적 알고리즘을 축소된 영상과 영상 내에서 추출된 블록 부분에서만 적용하여 수행 시간을 단축한다. (그림 1)(b)는 본 논

문에서 제안한 다해상도 분리를 이용한 명함 인식 시스템의 개괄적인 구성을 보여준다.



(그림 1) 명함 인식 시스템의 구성

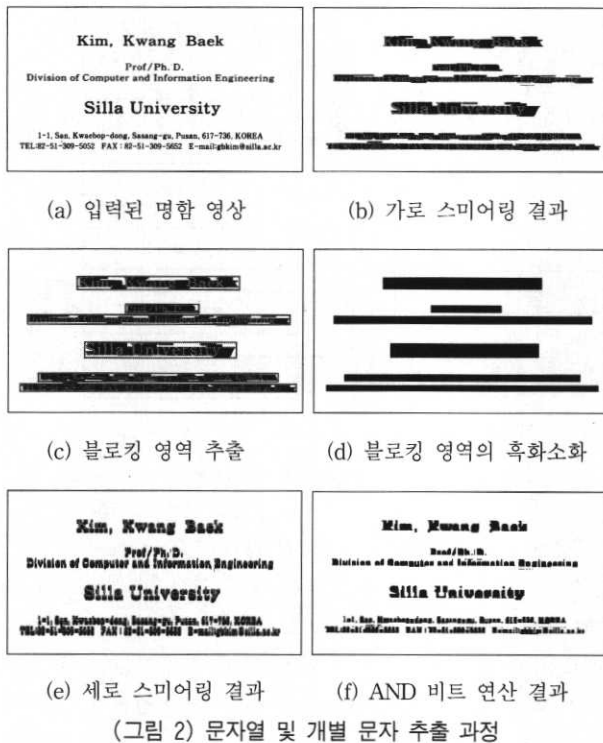
3.2 명함 영상 전처리와 블로킹 영역 추출

명함 인식에서는 우선 입력된 명함 영상으로부터 잡음을 제거하고, 명함의 사진, 로고, 실질적인 문자 등을 포함하는 블로킹 영역을 추출한다. 여기서 문자 이외의 요소를 포함하는 비문자열 영역은 압축된 이미지로 저장하고, 문자를 포함하는 문자열 블로킹 영역은 인식을 위해 개별 문자를 추출하여 정규화한다. 입력받은 명함 영상은 문자와 그림으로 이루어져 있으며, 명함 인식의 관심 영역인 개별 문자들을 정확히 추출하기 위해 명함 영상을 이진화하는 전처리 작업이 필요하다. 이진화된 명함 영상을 다해상도로 분리하는 작업은 잡음을 제거해주는 효과를 가져온다.

명함 영상의 블로킹 영역 추출 방법은 빠른 속도로 연산할 수 있고, 문자들간의 간격에 상관없이 때문에 다해상도의 영상 중 가장 작게 축소된 1/3배 축소 영상에 대해 가로 스미어링을 수행한 후에 블로킹 영역을 추출한다. 이는 명함 영상의 문자들이 가로 방향으로 일정한 크기와 간격으로 나열되어 있으므로 문자들간의 간격을 제거하는 가로 스미어링 방법을 이용하여 문자들을 연결시켜 블로킹 영역을 추출한다. 스미어링 방법은 문지르기 기법이라고도 하며, 문자간

의 여백이 설정 값 이하이면 글자간의 여백을 제거하는 전처리 방법이다. (그림 2)(b)은 가로 스미어링의 결과이다.

스미어링 수행 후에 추출된 각 블로킹 영역의 외곽 사각형(boundary rectangle)의 좌표와 그 안의 내용을 저장한다. 그리고 추출된 블로킹 영역에 대해 다시 세로 스미어링을 수행한 결과와 원래의 블로킹 영역사이에 AND 비트 연산을 수행하여 블로킹 영역의 외곽 사각형 안의 모든 내용을 흑색으로 설정한다. 이는 개별 문자 추출 단계에서 일부 문자를 잡음으로 간주하는 문제점을 해결하기 위한 것으로 다음 절의 문자 추출 알고리즘에서 자세히 논의한다. (그림 2)(b)~(그림 2)(f)는 문자열 영역 및 개별 문자의 추출 과정을 나타낸다.



3.3 문자 추출

문자 추출은 속도를 개선하기 위해 문자의 간격이 어느 정도 존재하는 1/2배 축소된 영상의 블로킹 영역에 대해 세로 스미어링 기법을 적용하여 문자들을 추출한다. 그러나 단순한 세로 스미어링 작업만으로는 문자를 정확히 추출할 수 없는 경우가 발생한다. 이에 본 논문에서는 문자를 형태대로 정확하게 추출하기 위해서 윤곽선 추적 알고리즘을 이용하여 문자를 추출한다. 그러나 윤곽선 추적 알고리즘을 적용하면 하나의 문자가 두 개로 분리되어 문자 크기 비율을 만족하지 못하고 잡음으로 간주되어 문자 영역이 삭제되는 경우가 발생하는데, 이러한 문제를 해결하기 위해서 본 논문에서는 스미어링 상태에서 윤곽선 추적 알고리즘을 적용한다. (그림 3)은 블로킹 영역에서 세로 스미어링을 수

행하여 문자를 추출한 결과와 본 논문에서 제안한 방법으로 문자를 추출한 결과를 나타내었다.

(그림 3)(a)와 같이 일반적인 문자 추출 방법은 가로 스미어링을 통해서 블로킹 영역을 추출한 후, 세로 스미어링을 이용하여 분리된 문자들을 투영에 의한 방법이나 윤곽선 추적 알고리즘으로 추출하게 된다. 이때, 문자의 특성에 따라 둘로 분리된 하나의 문자의 경우에는 잡음으로 제거되어 추출되지 않는다. (그림 3)(a)의 결과에서 문자 'i'를 정확히 추출하는데 실패하였고, 이는 인식 결과에까지 영향을 미치게 된다.



(그림 3) 블로킹 영역과 개별 문자 추출 과정 및 결과

본 논문에서 제안된 문자 추출 방법은 가로 스미어링으로 추출된 블로킹 영역의 외곽 사각형을 모두 흑색으로 설정한 영역과 블로킹 영역에 대해 세로 스미어링한 결과를 AND 비트 연산을 수행한 다음에 윤곽선 추적 알고리즘을 이용하여 각 개별 문자들을 추출한다. AND 비트 연산에 의해 추출된 결과는 하나의 문자의 특성이 분리되어 나타

나더라도 세로 스미어링한 결과에 의해서 묶여서 나타나므로 정확히 개별 문자가 추출된다. 1/2배 축소 영상에서 추출된 문자 영역들의 위치 정보를 이용하여 원래의 명함 영상에서 개별 문자들을 추출한다

3.4 개선된 RBF 네트워크를 이용한 명함 인식

RBF 네트워크는 빠른 학습 시간, 일반화(generality), 단순화(simplicity) 등의 특징을 가진 알고리즘으로 학습 데이터를 분류하는 작업과 비선형 시스템 모델링 등에 적용된다[12-14]. RBF 네트워크는 기존의 오류 역전파 학습 알고리즘과 같은 다층 퍼셉트론(multilayer perceptron)의 연결가중치들이 지역 최소화에 위치하여 학습 패턴들이 잘 분류되지 않은 경우를 방지할 수 있으며 학습 시간을 단축시킬 수 있다. RBF 네트워크에서 입력층과 중간층 사이에는 클러스터링 하는 역할을 수행하게 된다. 중간층에 있는 노드는 그 자신의 중심을 가지고, 입력 $X = \{x_1, x_2, \dots, x_n\}$ 에 대해 자신과 중심간의 거리 벡터를 구한다. 이 노드의 출력은 거리 벡터에 대한 일정 형태의 비선형 함수로 얻어진다. 입력이 중간층의 노드에 가까울수록 그 출력 값은 증가한다. 중간층의 노드를 구하는 것은 식 (1)과 같다. 여기서, c_k 는 각 클러스터의 중심 값이고, σ 는 중심에 대한 넓이 값이다.

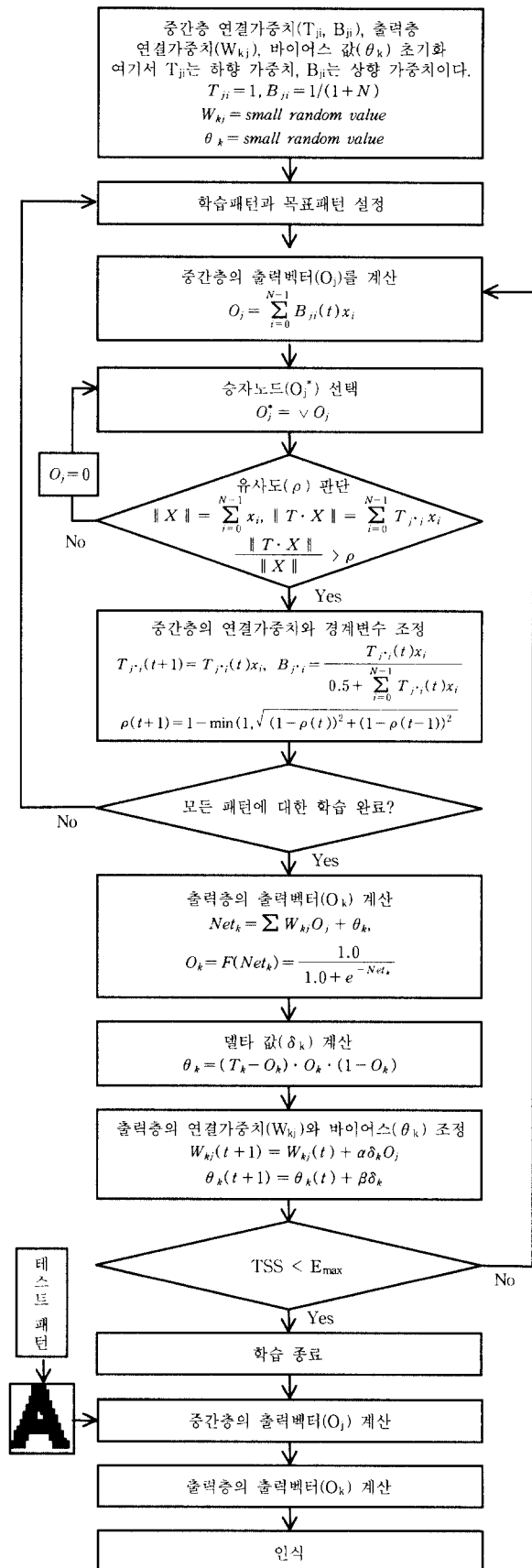
$$\rho_k = \exp\left(\frac{-\|X - c_k\|^2}{\sigma_k^2}\right) \quad (1)$$

중간층과 출력층 사이에는 일반적인 학습 과정을 수행한다. RBF의 출력 노드는 각 중간층 노드의 출력 값을 연결가중치를 곱한 후, 이들을 더한 값이 출력 값으로 식 (2)와 같이 표현된다.

$$O_i = \sum_{k=1}^n \rho_k w_{ik} \quad (2)$$

그리고, 일반적인 학습 방법인 최소 자승(least mean square) 방법으로 목표 값과 출력 값간의 오차를 최소화하는 방향으로 학습한다.

RBF 네트워크에서 중간층은 학습 패턴들을 클러스터링 하는 층이다. 이 중간층의 목적은 주어진 학습 벡터들에 대한 특징 벡터 공간 사이에서 한 클러스터 내의 벡터들간에 거리를 측정하여 정해진 반경 내에 존재하면 같은 클러스터에 속하고 정해진 반경 내에 존재하지 않으면 다른 클러스터에 속하게 된다. 그러나 정해진 반경 내에서 클러스터링 하는 것은 잘못된 클러스터를 선택할 수 있는 가능성을 내재하고 있어 잘못 선택된 클러스터의 중간층 노드는 출력층의 입력으로 구성되어 인식을 저하시키는 원인이 된다[15]. 따라서, 본 논문에서는 입력층과 중간층 사이의 학습 구조로 맞게 ART1을 개선하여 RBF 네트워크의 중간층으로 적용한다.



(그림 4) 개선된 RBF 네트워크의 학습 및 인식 구성도

ART1에서 경계 변수(vigilance parameter)는 임의의 패턴과 저장된 패턴과의 불일치 허용도를 역으로 결정한다 [16]. 이 경계 변수가 크면 입력 패턴과 기대 패턴 사이에 약간의 차이만 있어도 새로운 카테고리(category)로 분류하게 된다. 반대로 경계 변수가 작으면 입력 패턴과 기대 패턴 사이에 많은 차이가 있다라도 이를 허용하므로 입력 패턴들을 대략적으로 분류한다. 그런데 ART1를 명함 인식에 적용하기 위해서는 경계 변수 값을 경험적으로 설정하여야 하므로 인식률에 부정적인 영향을 준다. 본 논문에서는 경계 변수를 패턴간의 유사성에 따라 동적으로 조정하기 위하여 퍼지 논리 연산자 중에서 Yager의 교 접속 연산자를 적용하여 경계 변수를 동적으로 조정하도록 ART1를 개선하였다. 퍼지의 교 접속 연산자는 총체화된 값이 모든 입력 중 가장 작은 값보다 크지 않은 성질을 갖는 연산자로서, Yager의 교 연산자는 식 (3)과 같다[17].

$$\mu_{x_1 \cap x_2} = 1 - \min \left\{ 1, \left((1 - \mu_{x_1})^p + (1 - \mu_{x_2})^p \right)^{\frac{1}{p}} \right\} \quad (3)$$

여기서, 식 (3)은 $p \in (0, \infty)$ 값에 따라 단조 감소하는 함수이다. Yager의 교 연산자를 사용하여 경계 변수 (ρ)를 동적으로 조정하는 식은 다음과 같다.

$$\rho(t+1) = 1 - \min \left(1, \sqrt{((1 - \rho(t))^2 + (1 - \rho(t-1))^2)} \right) \quad (4)$$

본 논문에서는 식 (4)를 적용한 개선된 ART1을 RBF 신경망의 입력층과 중간층의 학습 구조로 적용하여 개선된 RBF 네트워크를 제안하고 명함 인식에 적용한다. 개선된 RBF 네트워크의 학습 및 인식 구성도는 (그림 4)와 같다.

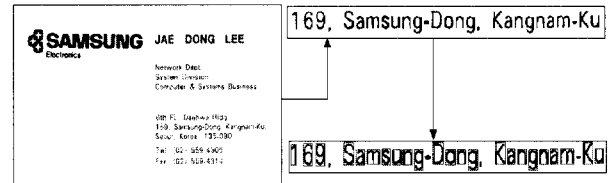
4. 실험 및 결과 분석

본 논문은 제안된 명함 인식 시스템을 IBM 호환 기종의 펜티엄III에서 C++ Builder로 구현하고, 1500×800픽셀 크기의 영문 명함 영상 50개를 대상으로 인식 실험을 수행하였다. <표 1>은 참고 문헌 [1]의 블로킹 영역 및 개별 문자 추출 방법과 본 논문의 제안된 방법간에 추출된 블로킹 영역의 총 개수와 개별 문자의 추출 개수를 비교하여 나타내었다. 그리고 제안된 RBF 네트워크를 이용한 두 방법간의 인식 결과도 나타내었다.

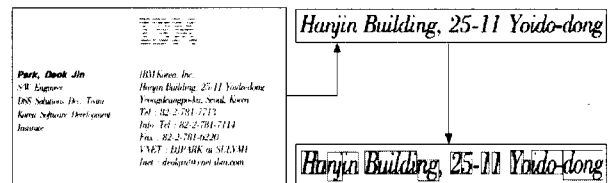
<표 1> 추출 방법에 따른 블로킹 영역과 개별 문자 영역의 추출 결과

	블로킹 영역	문자 영역	문자 인식
총 영역 개수	629	7641	7641
참고문헌[1]의 추출 방법	627	7073	6812
참고문헌[1]의 오류	2	568	829
제안된 추출 방법	627	7495	7386
제안된 방법의 오류	2	146	255

<표 1>에서 첫 번째 행의 총 영역 개수는 실험에 사용된 50장의 명함 영상에 포함된 블로킹 영역과 개별 문자 영역의 총 개수이다. 그리고 두 번째 행은 참고 문헌[1]의 명함 추출 방법을 이용하여 추출된 블로킹 영역과 개별 문자의 개수 그리고 인식이 성공한 개수를 나타내었다. 세 번째 행은 참고 문헌[1]의 추출 방법에서 추출에 실패한 개수와 오인식된 개수를 나타내었고 첫 번째 행에서 두 번째 행을 뺀 결과이다. 네 번째 행은 본 논문에서 제안한 명함 추출 방법의 블로킹 영역 개수와 개별 문자 영역 개수 그리고 인식 성공 개수를 나타내었다. 다섯 번째 행은 제안된 방법으로 추출이나 인식에서 실패한 개수를 나타내었고 첫 번째 행에서 네 번째 행을 뺀 결과이다. <표 1>에서 알 수 있듯이, 제안한 방법이 참고 문헌[1]의 추출 방법과 비교하여 블로킹 영역과 개별 문자들의 추출률이 향상되었고 인식 과정에서도 인식률이 높은 것을 알 수 있다. (그림 5)(a)는 제안된 방법에 의해서 개별 문자 추출에 성공한 것이고 (그림 5)(b)는 개별 문자 추출에 실패한 경우이다.



(a) 개별 문자 추출에 성공한 예제



(b) 개별 문자 추출에 실패한 예제

(그림 5) 개별 문자 추출 실행 예

본 논문에서 제안한 개선된 ART1과 기존의 ART1을 RBF 네트워크의 중간층으로 적용하여 학습 및 인식한 결과를 <표 2>에서 나타내었다.

<표 2> 개선된 RBF 네트워크를 이용한 학습 및 인식 결과

개선된 ART1 기반 RBF 네트워크		
	유니코드 코드수	인식 성능
영문자	380/520	3698/3723
숫자	121/200	2782/2821
특수문자	16/30	906/951
ART1 기반 RBF 네트워크		
영문자	498/520	3682/3723
숫자	175/200	2773/2821
특수문자	21/30	889/951

<표 2>에서 제시된 결과는 개선된 ART1과 기존의 ART1을 각각 RBF 네트워크의 중간층으로 적용한 다음, 개별 문자를 영문자, 숫자 그리고 특수문자로 각각 분리하여 학습한 결과이다. 기존 ART1 기반 RBF 네트워크에서 초기 경계 변수를 0.9로 설정하였고, 개선된 ART1 기반 RBF 네트워크에서는 초기 경계 변수를 0.9로 설정하여 동적으로 조정하였다. 개선된 ART1을 적용한 RBF 네트워크에서 영문자의 경우에는 520개의 입력 패턴에 대해서 380개의 중간층 노드가 생성되었다. 그리고 기존의 ART1을 적용한 RBF 네트워크에서는 영문자의 경우 520개의 입력 패턴에 대해 498개의 중간층 노드가 생성되었다. 따라서, 개선된 ART1을 적용한 RBF 네트워크로 명함 영상의 개별 문자를 학습하는 것이 효율적인 것을 알 수 있다. 50개의 명함 영상에서 개별 문자의 총 수는 7641개이다. 그러나 개별 문자 추출의 실패로 실질적으로 문자 인식이 사용된 개별 문자는 7495개이고, 그 중에서 영문자는 3723개, 숫자는 2821개, 특수문자는 951개이다. 인식 성능을 분석하면, 개선된 ART1을 적용한 RBF 네트워크에서 영문자의 경우에는 3723개의 영문자에 대해서 3698개가 정확히 인식되었고 특수문자의 인식 성능은 영문자와 숫자와 비교하여 그들이 가진 문자의 개수에 비교하여 오인식률이 높게 나타났지만 기존의 ART1을 적용한 RBF 네트워크 보다 학습 및 인식률이 향상된 것을 확인하였다.

5. 결 론

본 논문에서는 명함 인식의 영상 전처리 과정에서 속도와 정확성을 향상시키기 위하여 명함 영상을 처리 작업에 따라 적절한 해상도의 영상으로 분리하여 처리함으로써 명함 영상에서의 문자 추출을 빠른 시간 내에 정확하게 수행하도록 하였다. 즉, 블로킹 영역을 추출하기 위해서 1/3배 축소된 명함 영상을, 문자 추출을 위해 1/2배 축소된 명함 영상을, 정확한 인식 결과를 유도하기 위해 원래 명함 영상을 이용하여 영상 처리를 수행하는 다해상도 영상 기반의 계층적 영상처리 기법을 적용하였다. 그리고 문자들의 특성을 고려하여 문자 추출에 정확성을 부여하기 위하여 블로킹 영역의 외곽 사각형을 검은색으로 설정한 영역과 세로 스미어링 결과를 AND 연산을 수행함으로써 각 문자들을 분리시켰고, 개별 문자의 특징이 분리된 형태를 갖춘 문자는 묶어서 추출되도록 하였다. 추출된 개별 문자의 인식률을 높이기 위해, 경계 변수를 동적으로 조정하는 개선된 ART1 알고리즘을 적용하여 중간층을 효율적으로 구성할 수 있는 개선된 RBF 네트워크를 제안하고 개별 문자 인식에 적용하였다. 제안된 명함 추출 및 인식 방법의 성능을 확인하기 위해서, 실제 50개의 영문 명함 영상을 대상으로 실험한 결과, 기존의 방법보다 명함 정보 추출 및 인식에서 우수한 성능이 있음을 확인하였다.

참 고 문 헌

- [1] 김광백, 김철기, 김정원, "윤곽선 추적 알고리즘과 개선된 ART1을 이용한 영문 명함 인식에 관한 연구," 한국지능정보시스템학회논문지, 제8권 제2호, pp.105-115, 2002.
- [2] Gonzalez and Woods, Digital Image Processing Second Edition, Prentice Hall, 2002.
- [3] 원남식, 손윤구, "8-이웃 연결 값에 의한 병렬 세선화 알고리즘," 정보처리학회논문지, 제2권 제5호, pp.701-710, 1995.
- [4] 김성영, 권태균, 김민환, "추적에 의한 단순화된 윤곽선 추출," 한국멀티미디어학회, 춘계학술발표논문집, pp.356-361, 1999.
- [5] K. Y. Wang, R. G. Casey and F. M. J. Wahl, "Document analysis system," IBM J. Res. Develop.26, No.6, pp.647-656, 1982.
- [6] G. Nagy, S. C. Seth and S. D. Stoddard, "Document analysis with an expert system," Proceedings of Pattern Recognition in Practice II, Amsterdam, pp.19-21, 1985.
- [7] R. C. Gonzalez and P. Wintz, Digital Image Processing, Addison-Wesley Publishing Company Inc., pp.88-102, 1977.
- [8] K. S. Fu and A. Rosenfeld, "Pattern recognition and computer vision," IEEE Computer, pp.247-282, Oct., 1984.
- [9] Robert Schalkoff, Pattern Recognition : Statistical, Structural and Neural Approaches, John Wiley & Sons, 1992.
- [10] Richard O. Duda, Peter E. Hart and David G. Stork, Pattern Classification, 2nd Ed., John Wiley & Sons, 2001.
- [11] Abhijit S. Pandya and Robert B. Macy, Pattern Recognition with Neural Networks in C++, CRC Press, 1996.
- [12] M. Watanabe, K. Kuwata and R. Katayma, "Adaptive Tree-Structured Self Generating Radial Basis Function and its Application to Nonlinear Identification Problem," Proceedings of IIZUKA, pp.167-170, 1994.
- [13] J. Park and I. W. Sandberg, "Universal Approximation Using Radial Basis Function Network," Neural Comp. Vol.3, pp.246-257, 1991.
- [14] C. Panchapakesan, D. Ralph and M. Palaniswami, "Effects of Moving the Centers in an RBF Network," Proceedings of IJCNN, Vol.2, pp.1256-1260, 1998.
- [15] K. B. Kim and H. W. Yun, "A Study on Recognition of Bronchogenic Cancer Cell Image Using a New Physiological Fuzzy Neural Networks," Japanese Journal of Medical Electronics and Biological Engineering, Vol.13, No.5, pp. 39-43, 1999.
- [16] 김광백, "효과적인 운송 컨테이너 영상의 식별자 인식을 위한 개선된 퍼지 ART 알고리즘," 한국통신학회논문지, 제28권 제5호, pp.486-492, 2003.
- [17] H. J. Zimmermann, Fuzzy set Theory and its Applications, Kluwer Academic Publishers, 1991.



김 광 백

e-mail : gbkim@silla.ac.kr

1993년 부산대학교 전자계산학과(이학석사)

1999년 부산대학교 전자계산학과(이학박사)

1996년~1997년 동의공업대학 사무자동화과
전임강사

1999년~2000년 Biomedical Fuzzy Sys-
tems Association Associate Edi-
tors (Japan)

1997년~현재 신라대학교 컴퓨터공학과 부교수

2003년~현재 한국퍼지 및 지능시스템학회 이사

관심분야 : Neural Networks, Image Processing, Fuzzy Logic,
Biological Signal Processing and Biomedical System



김 영 주

e-mail : yjkim@silla.ac.kr

1988년 부산대학교 계산통계학과(이학사)

1990년 부산대학교 계산통계학과(이학석사)

1999년 부산대학교 전자계산학과(이학박사)

1990년~1995년 큐닉스컴퓨터 응용시스템
연구소

2000년~현재 신라대학교 컴퓨터공학과 조교수

관심분야 : 분산시스템, 멀티미디어, 영상 압축 및 통신