

클러스터 신경망을 이용한 우편번호 인식 시스템의 설계

Design of Zip Code Recognition System Using Cluster Neural Network

김종석* · 홍연찬**

Jong-Seok Kim and Yeon Chan Hong

* LG전자(주)

** 인천대학교 전기전자공학부

요 약

최근에는 대부분의 우편물 봉투가 창이나 색깔을 포함하고 있다. 본 논문에서는 창이 있는 봉투와 색깔이 있는 우편 봉투 영상에서 구조적 방법을 분석하여 수취인 주소 영역을 자동적으로 추출하는 시스템을 제안하였다. 제안된 방법은 이치화 전 에지 검출을 이용하고 문자열 추출 후 검출된 블록에 대해 적응 이치화를 적용함으로써 이치화 후 우편 번호를 검출할 때보다 우편 봉투의 숫자 패턴이 밝기 및 주변 환경에 의한 영향을 적게 받는다는 점에서 더 효율적이다.

ABSTRACT

In these days, most of the mail envelopes have windows or colors. In this paper, a system that can analyze the structural method of image from windowed or colored envelopes and can extract the address range of addressee automatically is suggested. By extracting edge before binarization and applying adaptive binarization on the extracted block after binarization, the proposed method is more effective than using extracting zip code after binarization, because the brightness and circumstance have a less effect on number patterns of envelopes.

Key Words : 우편번호인식, 클러스터 신경망, 소벨 마스크, 적응적 이진화, 히스토그램

1. 서 론

현대 사회가 고도 정보화 사회로 급진적으로 발전해감에 따라 정보의 전송 필요성이 날마다 늘어가고 있다. 이에 따라 우편물의 양은 날마다 폭발적으로 증가하고 있으며 이러한 우편물을 발송하기 위하여 각 지역별로 우편물을 분류하는데 많은 인력과 시간이 필요하기 때문에 보다 신속하고 정확한 우편 소통을 위하여 우편물 자동 분류에 대한 연구가 진행되고 있다.

현재까지 발표된 이진화 알고리즘은 크게 전역적 이진화와 지역적 이진화로 나눌 수 있다. 전역적 이진화 방법에는 클래스간 분산을 이용한 방법과 엔트로피를 이용한 방법, 히스토그램을 변형한 방법, 모멘트를 유지하는 방법 등이 있는데, 이 방법들은 알고리즘 자체가 간단하여, 단순한 영상에는 효율적으로 사용할 수 있으나, 잡영이 심한 우편영상이거나, 우편 영상에서 문자와 우편봉투의 명암 변화가 불규칙하면 결과적으로 연계 되는 이진영상의 질은 매우 떨어지게 된다 [1-3].

지역적 이진화 방법에는 로컬 윈도우를 이동하면서 임계값을 결정하는 방법과 경계(edge)정보를 이용한 방법으로 분류할 수 있다. 이 방법들은 실행속도가 전역적인 방법보다

느리지만 흐린 문자와 잡음이 섞인 영상에서 우편 문자를 검출할 수 있다는 장점이 있다[4,5].

우편 문자의 패턴 인식을 위한 설계방법으로 템플릿 비교방법[6], 문자를 구성하는 화소의 통계적 특징에 근거한 통계적 방법[7], 문자의 구문과 구조에 근거한 구조적 방법[8] 및 신경 회로망을 이용한 방법[9,10] 등으로 나눌 수 있다. 템플릿 비교에 의한 방법은 입력패턴과 가장 유사한 원형 패턴을 찾아내어 그 원형패턴이 속한 부류로 인식한다. 이 방법은 인쇄체 숫자와 같이 고정된 형태의 패턴을 인식하는데 사용될 수 있으며, 필기체 숫자의 경우에는 적합하지 않다. 통계적 문자인식 방법에서는 학습단계를 통하여 특징벡터의 통계적인 분포 특성을 구하고 그것을 이용하여 특징벡터 공간을 각 부류별로 분리시킨다. 이 방법은 패턴의 표현 방법에 의하여 시스템의 인식능력이 크게 좌우될 뿐만 아니라 인쇄체 문자패턴에 대해서는 적용이 가능하지만 글자의 모양이 다양한 필기체 문자의 인식에는 부적합한 단점이 있다. 구조적 문자인식 방법에서는 문자패턴을 구성하고 있는 자획 등과 같은 기본원소들의 집합으로 패턴을 표현한다. 그러나 이 방법에서는 모든 경우의 생성규칙을 기술해야만 하고 생성규칙이 복잡하여 수정이 어렵다. 또한 생성규칙에 정의되지 않은 패턴은 인식할 수 없으며, 문법의 기술이 방대하고 소프트웨어의 작성에 많은 시간이 소요되는 단점이 있다. 위와 같은 기존의 문자인식에 대한 방법론 외에, 최근에는 신경망 모델을 이용한 연구가 활발하게 진행되고 있다.

본 논문에서는 흰색의 일반 봉투는 물론 최근에 증가된

접수일자 : 2000년 9월 15일

완료일자 : 2001년 3월 2일

창이 있고 다양한 그레이 값을 갖는 칼라 및 창 봉투 영상에서의 이진화시 발생하는 에러를 감소시키기 위하여 우편 번호 숫자의 에지를 검출하는 방법으로 먼저 소벨 마스크를 이용하여 수평과 수직으로 미분을 하고 미분된 영상을 반전함으로써 우편영상을 추출한다. 이치화 전 에지 검출을 하고 문자열 추출 후 검출된 블록에 대해 적응 이치화를 적용함으로써 이치화 후 우편번호를 검출할 때보다 우편 봉투의 숫자 패턴이 밝기 및 주변 환경에 의한 영향에 대해 더 효율적임을 알 수 있다. 신경망 인식기의 구현의 어려움과 계산의 복잡성 및 정확도 문제 등을 고려하여 단일 인식기를 다수개의 부인식기들로 나누고, 각 부인식기의 입력으로 고차의 입력 벡터를 다수개의 저차 벡터로 분할하여 사용함으로써 인식기의 한계성과 서로 다른 입력 특징들간의 보완성을 충족시킬 수 있는 클러스터 신경망을 이용한 인식기를 구성한다[11].

2. 우편 번호 추출

우편번호 인식을 위해 먼저 우편봉투 영상에서 수신영역(ROI)을 추출한다. 우편봉투 영상이 가지는 에지를 수평 및 수직 방향으로 각각 Sobel 연산자를 이용하여 추출함으로써 우편봉투의 색상이나 구조, 주변환경의 영향을 받지 않는 방법을 택하였다.

일반적인 방법에서는 영상을 먼저 이치화하고 이치화된 결과에 대하여 히스토그램을 이용하여 개별 문자를 추출하였다. 이러한 방법의 경우 특히 문자의 영상이 흐린 경우에 이치화 과정에서 정보 손실과 잡음의 가능성이 항상 존재함으로써 문자 블록의 검출시 에러를 초래 할 수 있다. 따라서 이러한 단점을 해결하고, 우편 봉투 영상의 에지 특성에 의한 문자블록을 검출함으로써 일관성 있는 알고리즘을 적용시킬 수 있는 알고리즘 방법을 보인다.

2.1 전처리

전처리부는 흰색의 일반 봉투는 물론 최근에 증가된 창이 있고 다양한 그레이 값을 갖는 칼라 봉투 영상에서의 이진화시 발생하는 에러를 감소시키기 위하여 먼저 소벨 마스크를 이용하여 수평과 수직으로 미분을 하고 미분된 영상을 반전함으로써 우편영상을 추출하였다.

① 우편 봉투 영상

우편 봉투 영상은 칼라가 있는 우편봉투와 창이 있는 우편 봉투 영상 50장을 400 dpi의 스캐너로 입력 받아 640 × 360 크기의 그레이 값을 갖는 Raw 파일로 저장하여 사용하였다.

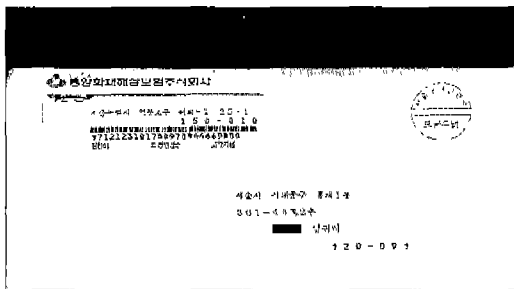


그림 1. 우편봉투 영상

② 소벨 미분 연산자에 의한 영상 미분

영상에서의 에지는 상당한 밝기의 차이가 있는 곳이고 이것은 대개 물체의 윤곽선에 해당하는 곳이므로 대개 픽셀값의 불연속이나 픽셀 미분값의 불연속점에 존재한다. 스텝 불연속점에 해당하는 에지를 스텝 에지(step edge)라고 하고, 라인 불연속점에 해당하는 연산을 라인 에지(line edge)라고 한다. 보통 에지 검출 연산을 하기 전에 잡음을 제거하기 위하여 먼저 가우시안 필터를 사용하고, 에지 검출을 한다. 이럴 경우 평활화 연산의 영향으로 픽셀값의 변화가 둔화되면서 스텝 에지가 램프 에지로 변화되지만 우편 봉투 영상에서는 배경이 어느 정도 일정하고 문자와 배경의 급격한 에지를 검출해야 하므로 이 과정을 거치지 않아도 된다.

윤곽선을 검출하기 위해서 수학적으로 표현하면 대부분의 윤곽선 검출 기법은 편미분 연산자 계산에 근거한다. 영상에서의 기울기를 그래디언트(gradient)라고 하는데 영상에서 그래디언트의 크기를 구하면 윤곽선을 구할 수 있다. 그래디언트가 윤곽선 검출에 있어서 그래디언트가 큰 곳이 윤곽선이라고 말할 수 있다. 그래디언트를 구하는 공식은 먼저 기울기 벡터를 구하고,

$$\Delta F = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (1)$$

이 벡터의 크기를 구하면,

$$\Delta f = \text{mag}(\Delta F) = [G_x^2 + G_y^2]^{\frac{1}{2}} \quad (2)$$

여기서 자승 관계를 다음과 같이 간단히 하여 사용하였다.

$$\Delta f \approx |G_x + G_y| \quad (3)$$

이 식은 계산량도 적고 직관적인 판독이 가능하므로 민감한 경우의 에지 검출이 아니라면 프로그래밍 하는데 매우 편리하다. 그래디언트 방향 역시 중요한 값인데 방향 값은 영상 안의 여러 개의 윤곽선들 중에서 윤곽선을 구분하는 기준이 된다. 즉 방향 값이 서로 같은 윤곽선은 하나의 윤곽선으로 구분되어 진다.

그래디언트의 방향 값은,

$$\alpha(x, y) = \tan^{-1} \left(-\frac{G_y}{G_x} \right) \quad (4)$$

위 수식에서 우편봉투 영상은 2차원이므로 미분도 x 축과 y 축으로 각각 미분을 해주었다.

여러 개의 윤곽선 알고리즘에서 본 논문은 편미분 연산에 의한 직접적인 연산보다 마스크를 이용하여 빠른 계산을 하였다.

마스크의 형태는 3 × 3이며 원 영상의 3 × 3 픽셀 단위로 연산을 하였으며 2차 미분으로 x축, y축으로 각각 한번씩 미분을 하였고, 기울기의 크기는

$$G = \sqrt{(G_x)^2 + (G_y)^2} \quad (5)$$

(단 G_x 는 x축 편미분, G_y 는 y축 편미분)

다음은 프로그램에 사용한 소벨 마스크 이다.

-1	0	1
-2	0	2
-1	0	1

1	2	1
0	0	0
-1	-2	-1

(왼쪽 G_x 미분 연산자, G_y 미분 연산자)
그림 2. 소벨 마스크

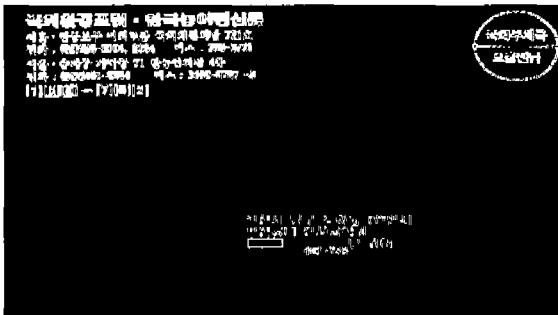


그림 3. 소벨 마스크를 적용한 영상

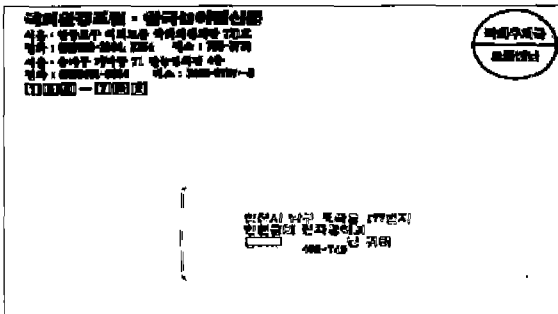


그림 4. 반전된 우편봉투 영상

원색의 일반봉투는 물론 최근에 증가하고 있는 비닐창과 다양한 그레이 값을 갖는 칼라 봉투 영상에서의 이진화시 발생하는 에러를 감소시키기 위하여 먼저 소벨 마스크를 이용하여 수평과 수직으로 미분을 하고 미분된 영상을 반전함으로써 우편 번호 영역을 추출하였다.

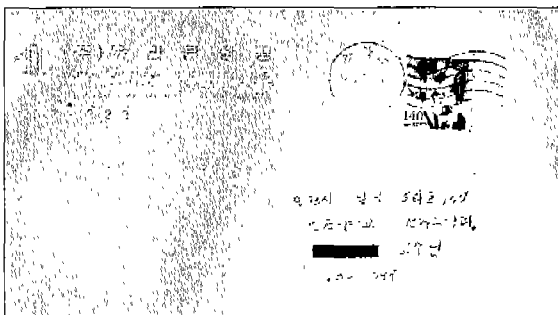


그림 5. 명도가 높은 우편봉투

그림 5는 노란색의 우편 봉투로 명도 값이 높고 문자가 흐린 영상의 경우에 대한 예이다.

2.2 수신 영역(ROI) 추출

소벨 연산된 우편 봉투 영상에서 수신영역(ROI)을 찾기 위해 화소 투영(Pixel Projection)을 수평과 수직으로 투영함으로써 수신영역을 추출하였다.

① Horizontal Pixel Projection

수신영역 투영시 수신 블록에 대한 길이와 폭의 일정한 제한을 두어 추출하였으며, Pixel Projection시 소벨 연산된 영상에서 수평 화소의 누적값을 비교하는데, 영상의 계산과 ROI영역의 추출에 빠른 연산을 위하여 수평 화소값이 임계값 T 이상 크면 좌표의 값을 평균값을 취하는데 대입하고, 그렇지 않으면 수평 좌표의 값을 무시하는 방법으로 수신영역의 수평 블록을 추출하였다. 여기에서 $f(x_i, y)$ 은 소벨 연산된 입력 영상이고 $b_{hi}(y)$ 는 일정 임계값 이상이 되는 영상의 수평 좌표이다.

if ($f(x_i, y) = BlackPixel$,

$$\sum_{i=1}^n f(x_i, y) > T \quad \text{then } y = b_{hi}(y)$$

$$\text{else } y \neq b_{hi}(y)$$
(6)

$$\sum_{i=1}^n \frac{b_{hi}(y)}{N}$$
(7)

② Vertical Pixel Projection

수평으로 화소 투영된 블록에 대해 수직 영역도 길이와 폭의 일정한 제한을 두어 추출하였다. 수신 영역의 수직 투영 추출도 빠른 연산을 위하여 수평 화소값이 임계값 T 이상 크면 좌표의 값을 평균값을 취하는데 대입하고, 그렇지 않으면 수직 좌표의 값을 무시하는 방법으로 수신영역의 수평 블록을 추출하였다. 여기에서 $f(x, y_i)$ 은 소벨 연산된 입력 영상이고 $b_{vi}(x)$ 는 일정 임계값 이상이 되는 영상의 수평 좌표이다.

if ($f(x, y_i) = BlackPixel$,

$$\sum_{i=1}^n f(x, y_i) > T \quad \text{then } y = b_{vi}(x)$$

$$\text{else } y! = b_{vi}(x)$$
(8)

2.3 우편 번호 영역 추출

ROI영역 검출 후 문자 블록을 검출하기 위하여 추출된 영역에 대해 다음과 같이 문자 블록을 추출한다. 수신영역(ROI)에서 수평과 수직으로 투영함으로써 문자 블록을 추출하였다.

① 문자열 검출

먼저 수평으로 위에서 Pixel Projection을 함으로써 문자 영역의 위의 좌표를 검출하고 수평 수신영역 투영시 문자 블록에 대한 여러 임계값을 적용해 추출하였다.

$$\sum f(x_i, y) > T_1 \quad \text{then } B_1 = y$$
(9)

$$\text{if} (\sum f(x_i, y) > T_2 \quad \text{and} \quad (B_1 - y) > T_3$$

$$\text{and} \quad (B_1 - y) < T_4)$$

$$\text{then } B_2 = y$$
(10)

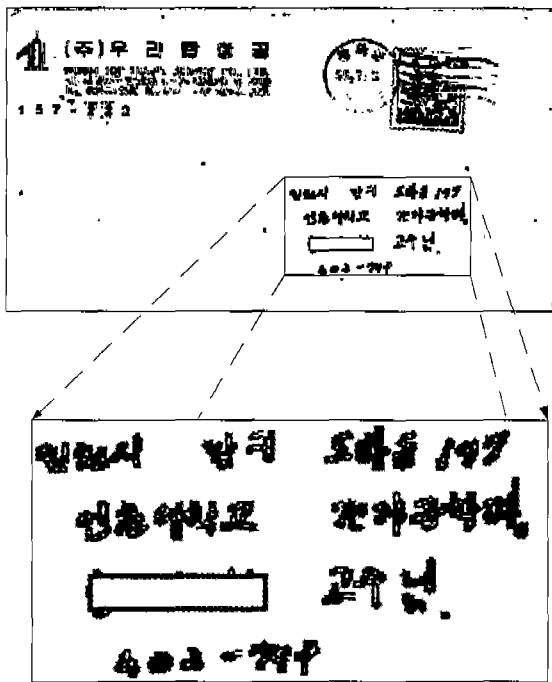


그림 6. ROI 영역의 추출

$$\text{if}((B_1 - y) > T_5 \text{ or } (\sum f(x_i, y) > T_6))$$

$$\text{then } B_1 = y \tag{11}$$

여기서 B1은 문자 블록의 시작 수평 좌표이고, B2는 각각의 문자 블록의 좌표이다.

(9) 식에서와 같이 수평 Pixel Projection의 일정 합이 임계값 T1 이상이면 그때의 수평좌표를 문자블록의 위의 수평좌표로 설정하고, 계속적으로 투영을 해 나가면서 (10) 식과 같이 문자 블록의 일정 폭의 길이와 Pixel Projection의 값이 임계값 T2 보다 클 때까지 투영함으로 문자 블록을 추출하였다. 또한 우편 봉투의 소인이 아래로 위치해 있는 경우와 문자가 아닌 다른 영상이 있는 경우와 노이즈에 의한 영상에 대해서는 (11) 식과 같은 문자 블록의 제한을 두어 추출하였다.

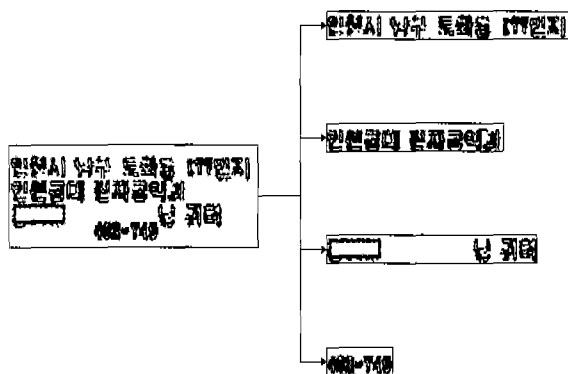


그림 7. 문자 블록의 추출

2.4 우편 번호 문자 분리

Pixel Projection에 의해 분리된 문자열 영역은 수취인의 주소뿐만 아니라, 이름과 우편번호를 모두 포함한다. 따라서

우편번호 영역의 분리 및 인식을 위해서는 우편번호 영역을 추출하고 추출된 우편 번호 영역에서 우편 번호 숫자를 분리하는 방법을 기술한다.

① 적응적 이진화

전체적으로 이진화는 하나의 임계값을 결정하기가 매우 어려울 뿐만 아니라 우편 봉투 영상의 주소 영역이나 배경이 특정한 색을 갖는 경우가 많기 때문에 우편 봉투 영상의 이진화는 효과적이지 못하다. 추출된 우편 번호 영역의 블록 평균보다 조금 높은 임계값을 이용해 우편 영상을 이진화 하였고 우편 봉투 영상에서 추출된 문자열 블록의 명도 영상을 전체적으로 이진화하였다.

② 연결 요소의 추출

연결요소는 임의의 흑화소에 대하여 8방향 인접화소가 흑화소일 경우 흑화소들을 모두 연결하여 추출하고, 연결 요소를 포함하는 최소 사각형의 위치, 최소 사각형의 중형비, 사각형 내 흑화소의 밀도 등의 특징을 추출하여 기술어진 영상 교정이나 수취인 주소 영역을 추출할 때 이용한다. 본 논문에서는 추출된 문자의 크기가 작은 경우에는 연결요소를 추출하지 않고, 그대로 문자 패턴을 사용하였고, 우편 문자 분리에 앞서 추출된 우편 번호 블록의 이진화시 발생하는 노이즈를 제거하고 우편 문자를 정확하게 분리할 목적으로 사용하였다.

③ 우편 번호 추출

우편 번호 문자를 신경망의 인식기에서 인식할 수 있는 6개의 숫자 영상을 추출하기 위하여 이치화되고 연결요소가 추출된 우편 번호 블록에서 수직 방향으로 투영하여 검은 화소의 히스토그램을 구하고, 구한 히스토그램의 투영값에서 문자의 일정 폭과 연결요소의 농도값에 의해 문자를 분리하였다.

④ RL 알고리즘에 의한 정규화

우편 영상에서 숫자 패턴은 동일한 패턴이라 할지라도 데이터의 형태가 다양하고 그 변화를 거의 예측할 수 없다. 또한 지나친 이미지 정보는 많은 특징벡터로 인해 인식시스템의 효율을 저하시키게 된다. 이러한 점에서 입력된 이진 영상패턴에 대해 정규화 하는 과정이 필요하다. 이는 특징 추출 및 인식기에 효과적인 처리효과를 가져온다.

본 논문에서는 다음과 같은 두 가지 단계를 거쳐 입력 패턴을 정규화 하였다. 먼저, 입력 영상으로부터 숫자 영상을 나타내는 정보점으로 처리영역을 처리하여 설정하였으며, 처리된 영역을 영상 패턴에 대해서 RL(run length) 알고리즘을 이용하여 28 × 25의 규격에 따라 스케일링(scaling)하였다[6].

RL 방법은 처리 영역으로 설정된 이진 영상에서 1과 0을 run으로 일련된 1 또는 0의 개수를 length로 설정하여 이것을 비율 C로 나누어 스케일링한 후 스케일링된 run length를 다시 일련의 1과 0으로 나타냄으로써 정규화된 영상을 얻는다. 압축 과정은 압축 비율 C_H에 의해 먼저 수평방향으로 수행되며, 그 다음에 C_V에 의해 수직 방향으로 수행된다. 정규화 비율 C는 다음과 같은 식으로 구해진다.

$$C = C_H = \frac{\text{입력 영상의 쉼련수}}{\text{원하는 압축영상의 쉼련수}} \tag{12}$$

$$C = C_v = \frac{\text{입력 영상의 열 수}}{\text{원하는 압축영상의 열 수}} \quad (13)$$

그림 8은 입력 영상으로부터 처리영역을 설정한 후 스케일링을 수행한 전처리 과정을 나타낸다.

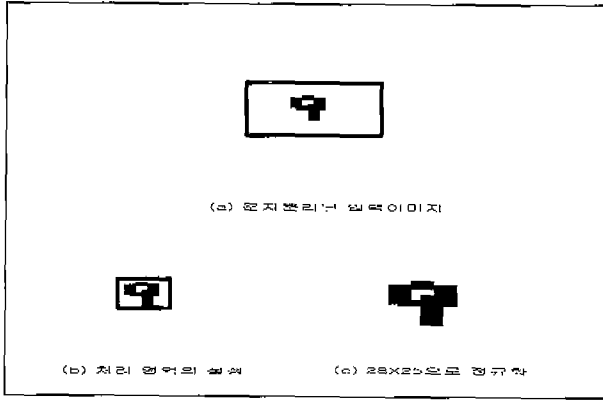


그림 8. RL 알고리즘

3. 우편 번호 인식 및 시스템 구현

본 장에서는 우편 영상에서 추출된 우편번호 인식 시스템에 적용한 클러스터 신경망과 우편번호 숫자의 특징 추출 및 본 논문에서 구현한 시스템에 관하여 기술한다.

3.1 신경망 인식기

추출된 우편번호는 RLC 알고리즘을 이용하여 규격의 이미지로 정규화 하였으며 이렇게 정규화된 우편번호는 망 특징 벡터와 수평, 수직, 사선, 역사선의 4방향 특징을 입력으로 하는 클러스터 신경망 분류기를 구성하였다.

① 클러스터 신경망

클러스터 MLP 신경회로망이란 각 층에 있는 유니트들이 여러 개의 군집으로 군집화되고, 상위층과 하위층에 있는 군집이 서로 쌍을 이루어 국부적인 네트워크를 구성한 신경회로망을 의미하며 각각의 국부적인 네트워크는 서로 독립적으로 완전 연결된다. 그림 9는 클러스터 신경망의 구조를 나타낸다. 클러스터 신경망의 학습도 일반적인 MLP 신경망과 같이 BP 알고리즘으로 학습이 되며, 본 논문에서는 델타 법칙을 이용한 수정된 BP 학습 알고리즘을 이용하여 수행하였다.

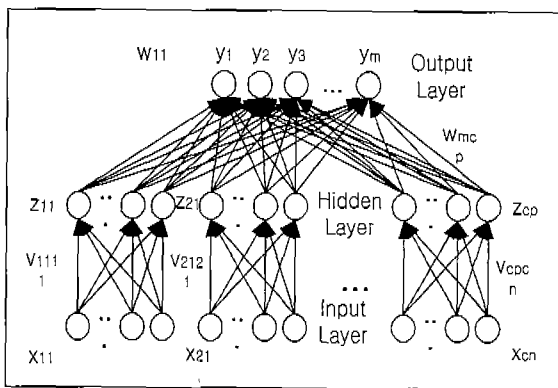


그림 9. 클러스터 신경망

우편 번호 인식기에 대한 블록도는 그림 10에 나타났다. 제안된 숫자 분류기는 전처리, 망 특징 추출, 인쇄체 인식기, 4방향 특징 추출로 구성된다.

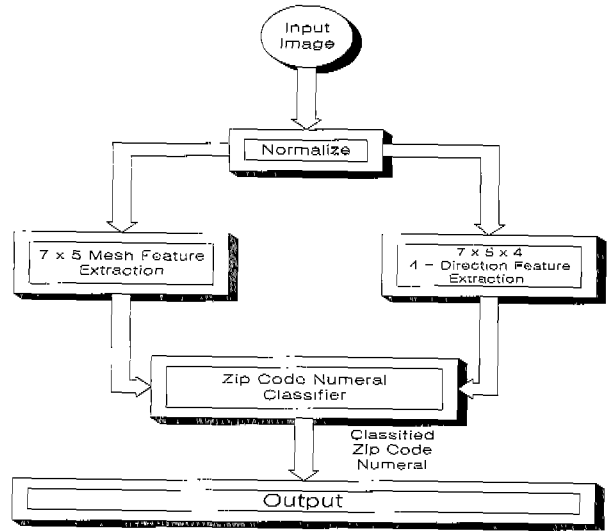


그림 10. 인식 시스템의 블록도

입력된 숫자 영상은 전처리 과정을 통하여 망 특징 추출 및 4방향 특징 추출 단계에 보내진다. 여기서 추출된 복합특징 벡터는 모듈화 된 클러스터 신경망에 의해 인식됨으로써 신뢰도가 높은 분류를 할 수 있다.

인식기의 구성 방법은 인식 시스템의 규모 및 인식 성능을 결정하는 것으로서 실제 구현시 여러 가지 사항을 고려해야 한다. 숫자를 인식하기 위하여 사용할 수 있는 가장 원시적인 형태의 신경망 구조는 매우 많은 수의 유니트를 완전 연결시키는 것이다. 그러나 사용되는 은닉 유니트의 개수를 몇 개로 해야 하는 지에 대해서는 아직 명확한 해결책이 제시되지 않고 있다. 만약 가중치의 개수를 적게 한다면 신경망의 정확한 학습은 이루어지지 않으며, 반면, 가중치의 개수가 학습 집합에 비해 너무 크면 불필요한 독립 변수만 늘어나게 되어 계산의 복잡성 및 정확도 문제 등이 발생하여 되어 인식 시스템의 효율성을 저하시키게 된다. 이러한 문제점은 단일 인식기를 다수개의 부인식기들로 나누고, 각 부인식기의 입력으로 고차의 입력 벡터를 다수개의 저차 벡터로 분할하여 사용함으로써 해결될 수 있다. 이와 같은 부인식기를 결합한 인식기는 구현 및 계산의 용이성뿐만 아니라 부인식기 각각의 인식 결과에 비해서 우수한 성능을 나타낸다. 다수개의 부인식기를 사용한 결합형 인식기는 부인식기들 사이에 상호 보완성, 각 부인식기의 전문성 여부 및 각 부인식기의 입력으로 사용하는 특징들간의 상호 보완성을 가져야한다. 실제 구현시 이상의 조건을 충족시키는 결합형 인식기의 구성은 용이하지 않다.

따라서 본 연구에서는 위와 같은 조건을 만족하기 위해 최근 연구에서 제시된 입력층과 은닉층 사이의 유니트를 국부적으로 연결시킨 클러스터 MLP 신경회로망을 인식기로 사용하였다.

입력층은 5개의 7×5 유니트로 구성되는데, 이는 4방향에 대한 특징 벡터와 망 특징에 대한 특징 벡터를 입력으로 하기 때문이다. 각 방향에 대한 특징 벡터 및 망 특징 벡터는 그림 11에 있는 국부적인 은닉 유니트와 군집을 이루어 완전 연결되어 있다. 따라서 각 그룹의 은닉 유니트는 4개의 7×

5 국부적인 특징 벡터와 1개의 7×5 전역적인 특징 벡터 중 어느 하나에 대한 특징 벡터를 입력으로 한다. 즉 첫 번째 은닉 유닛 그룹은 수평 방향으로의 특징 벡터에 의해 출력 값을 결정하게 되며, 두 번째 은닉 유닛 그룹은 사선 방향으로의 특징 벡터, 그리고 세 번째와 네 번째 그룹, 다섯 번째 그룹은 각각 수직 방향과 역사선 방향으로의 특징 벡터 그리고 전역적인 특징 벡터에 의해 출력 값을 결정하게 된다.

각 특징 벡터에 따른 클러스터 구조는 각 부류의 입력 영상으로부터 서로 구별되는 특징을 다룰 수 있다. 예를 들어, '0'과 '6'을 분류한다고 할 때 수직 방향으로의 특징 벡터만으로는 두 부류를 정확히 구별해내기가 어렵게 된다. 즉, 세 번째 은닉 유닛 그룹의 출력 값은 '0'과 '6'에 대해서 거의 유사한 값을 출력하게 된다. 그러나, 수평 방향으로의 특징 벡터에 의해 반응하도록 학습된 첫 번째 은닉 유닛 그룹과 역사선 방향으로의 특징 벡터에 의해 반응하도록 학습된 네 번째 은닉 유닛 그룹은 두 부류에 대하여 서로 구별되는 값을 출력하므로 결과적으로 '0'과 '6'을 구별해 낼 수 있게 된다.

출력층은 모두 10개의 유닛으로 구성되며 은닉 유닛 그룹의 출력 값에 의해 각 부류의 숫자를 분류하도록 학습된다. 그리고 그림 9에서와 같이 은닉층과 출력층 사이는 완전 연결된다. 출력층에 있는 10개의 유닛은 입력 부류에 해당하는 유닛만이 1 값을 갖고 그 외의 유닛은 0 값을 갖도록 하였다. 예를 들어 '3'에 대한 정규화 영상이 입력되었을 때, '3'을 나타내는 유닛만이 1 값을 갖고 그 외의 유닛은 0 이된다.

② 망 특징 추출

최근의 연구에서도 알 수 있듯이 신경망의 입력으로써 숫자 영상을 그대로 사용할 경우보다 숫자 영상에 대한 사전 정보를 이용할 경우 즉, 입력 영상으로부터 특징을 추출하여 신경망의 입력으로 사용할 경우 신경망의 인식 성능이 매우 향상된다. 그러므로 인식에 필요한 중요한 정보의 손실을 최소화하는 특징 추출기를 설계하여 특징 벡터를 추출할 필요가 있다.

따라서 본 논문에서는 최근의 특징추출 성능 비교 연구에서 이와 같은 조건을 만족하는 특징 추출 방법으로써 최적의 성능을 나타내는 것으로 알려진 4방향 특징 및 망 특징을 특징 벡터로 사용하였다. 따라서 필기체 우편 번호 문자 패턴의 다양함과 변형을 흡수 할 수 있는 4방향 특징 및 숫자 영상의 전역적인 탐색을 위한 망 특징을 복합특징으로 사용하였다[11].

망 특징은 숫자 영상을 수직과 수평으로 균등한 크기나 비균등적 크기의 부 영역으로 나누어 각 부 영역에 대하여 획의 수를 누적하여 특징으로 이용한다. 이 특징은 획의 이동에 약한 단점을 가지고 있으나 인쇄체의 경우에는 획에 따른 입력 패턴이 거의 일정하므로 망 특징을 통한 전역적인 탐색으로 인식이 가능하다. 따라서 28×25 로 정규화된 영상에서 7×5 의 망 특징을 추출하여 신경망의 입력 벡터로 사용하였다.

③ Kirsch 마스크를 이용한 4방향 특징

우편 번호 숫자의 경우, 숫자 영상의 정해진 위치에서 일정한 방향으로의 선분 존재에 의해 구별될 수 있으므로, 위치 정보와 그 위치에서의 선분의 유무 및 이 선분의 방향을 일반화된 특징 벡터로서 사용할 수 있으며 이러한 특징 벡터는 다양한 에지 검출 방법을 이용하여 얻을 수 있다. 본 연구

에서는 빠른 계산을 고려하여 3×3 마스크를 사용하고, 주위 8개의 이웃을 모두 고려함으로써 다른 마스크에 비하여 4 방향에 대한 보다 정확한 에지를 검출할 수 있는 Kirsch 마스크를 사용하였다[7]. Kirsch 에지 검출식은 인접화소 (i, j) 의 8 이웃인 a_k 에 따라 다음과 같다.

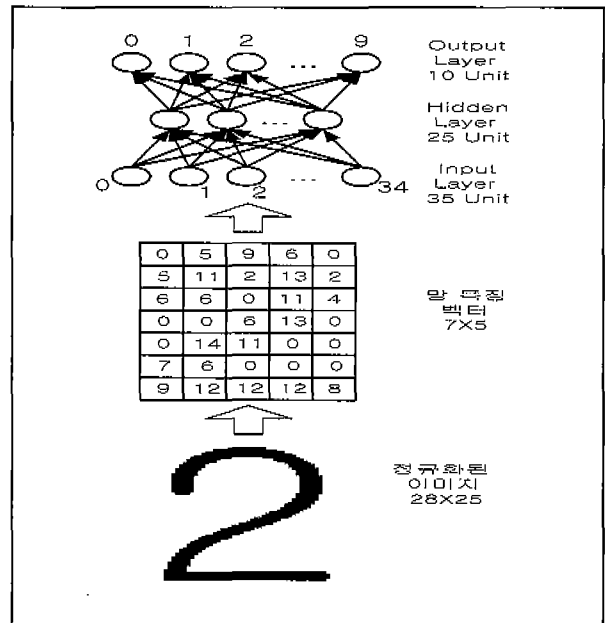


그림 11. 망 특징 추출

$$m(i, j) = \max\{1, \{|5s_k - 3t_k| : k=0, \dots, 7\}\} \quad (14)$$

단, $s_k = a_k + a_{k+1} + a_{k+2}$ 그리고 $t_k = a_{k+3} + a_{k+4} + a_{k+5} + a_{k+6} + a_{k+7}$.

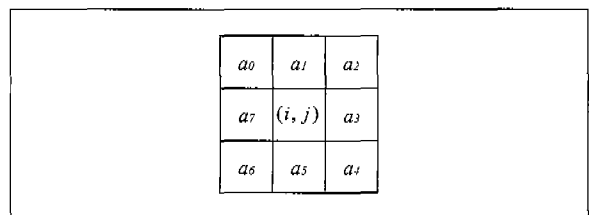


그림 12. 8 이웃 화소 (i, j)

(14) 식에서 $m(i, j)$ 는 화소 (i, j) 에서의 기울기를 나타내며, 첨자 k는 modulo 8연산을 통해서 구한 값이다. 여기서 수평(H), 수직(V), 사선(R), 역사선(L)의 방향 성분에 따른 국부적인 특징 벡터를 추출하기 위해 다음과 같은 수식으로 방향 성분을 계산하였다.

$$\begin{aligned} m(i, j)_H &= \max(|5s_0 - 3t_0|, |5s_2 - 3t_4|) \\ m(i, j)_V &= \max(|5s_2 - 3t_2|, |5s_6 - 3t_6|) \\ m(i, j)_R &= \max(|5s_1 - 3t_1|, |5s_5 - 3t_5|) \\ m(i, j)_L &= \max(|5s_3 - 3t_3|, |5s_7 - 3t_7|) \end{aligned} \quad (15)$$

(15) 식에 따른 4방향의 Kirsch 마스크는 그림 13과 같다.

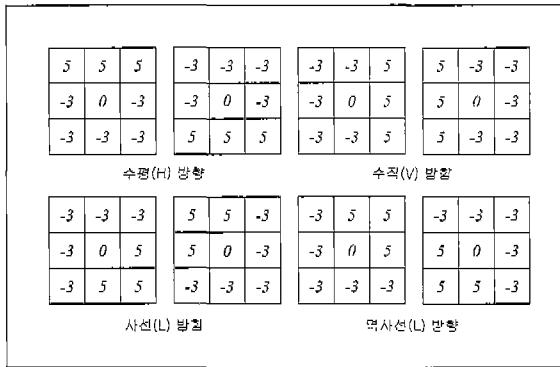


그림 13. 4 방향 요소에 대한 Kirsh 마스크

28 × 25로 정규화된 입력 영상으로부터 4쌍으로 이루어진 Kirsch 마스크를 통해 추출된 방향정보는 평균화 연산을 통해 다시 7 × 5 × 4의 국부적인 특징 벡터로 압축되어 우편번호 인식을 위해 클러스터 신경망의 입력으로 사용된다.

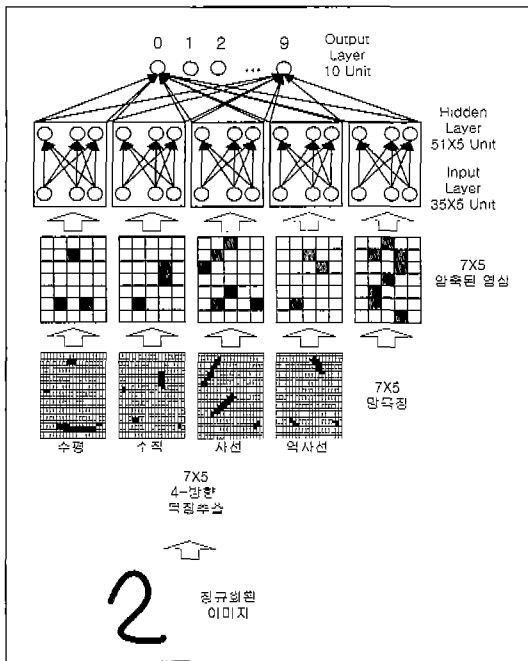


그림 14. 우편번호 인식 시스템의 구조

3.2 우편 번호 인식기 설계

우편 번호는 필체에 따라 획의 변화 및 변형이 심하다. 따라서 인식에 필요한 중요한 정보의 손실을 최소화하는 특징 추출기를 설계하여 특징 벡터를 추출할 필요가 있다. 또한 단일 특징만으로는 높은 인식률을 기대하기 어려우며, 인식률의 개선을 위하여 다양한 특징을 조합한 복합특징을 사용할 경우 인식률을 개선시킬 수 있다. 따라서 필기체 우편 번호의 다양한 형태 변형을 흡수할 수 있는 최적의 특징을 찾는 문제가 제기된다.

① 시스템 구성

시스템의 전체 처리과정은 4개의 모듈, 즉 우편 봉투 영상

전처리부, 수신영역 추출부, 우편번호 추출부, 그리고 제안된 신경망을 이용한 우편번호 인식부로 구성되어 있다.

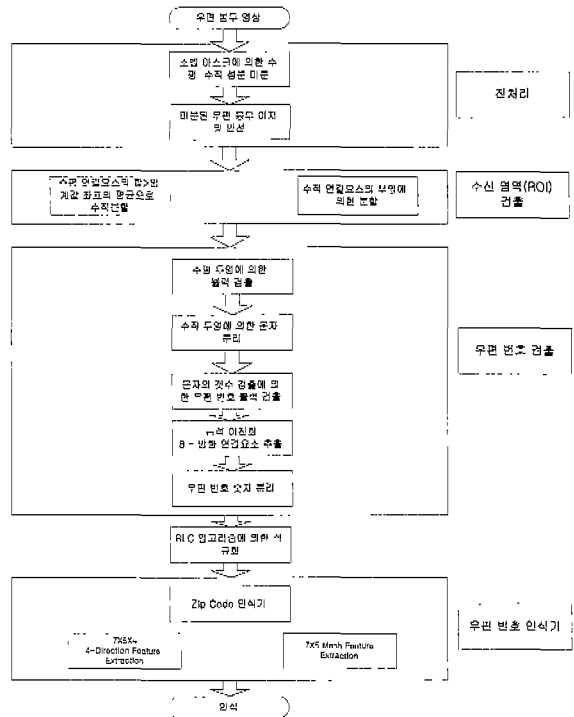


그림 15. 우편번호 인식 시스템

전처리부는 흰색의 일반 봉투는 물론 최근에 증가된 창이 있고 다양한 그레이 값을 갖는 칼라 봉투 영상에서의 이진화 시 발생하는 에러를 감소시키기 위하여 먼저 소벨 마스크를 이용하여 수평과 수직으로 미분을 하고 미분된 영상을 반전함으로써 우편영상을 추출하였다. 또한 우리나라 우편 봉투에서의 주소영역들은 서로 근접하기 때문에 수신영역 추출부에서는 연결요소의 구조적 수평 투영에 의한 합과 평균값을 적용해 수신자 영역을 추출하였다. 우편번호 추출부에서는 수신자 영역에 대해 주소 블록의 추출을 위하여 블록의 밀도, 길이, 거리등의 임의의 임계값을 가지고 수평과 수직투영을 함으로써 수신영역의 주소블록을 추출하였다. 그리고 추출된 주소블록에 대해 블록단위로 적용적 이진화를 수행하였고 다시 임의의 폭과 투영밀도를 가진 값을 수직으로 투영함으로써 글자단위로 분리하였다. 최소 분산 분포와 글자의 개수에 의해 우편번호 블록에서의 우편번호를 추출한다. 추출된 우편번호는 RLC 알고리즘을 이용하여 규격의 이미지로 정규화 하였으며 이렇게 정규화 된 우편번호는 시스템 처리 시간의 향상을 위하여 망 특징 벡터를 입력으로 하는 볼러스터 신경망 인식기를 구성하였다.

4. 실험 및 고찰

4.1 실험 환경 및 실험 방법

본 논문에서는 우편 번호 인식기의 성능을 검증하기 위하여 인쇄체와 필기체가 혼용된 우편 봉투 영상 50 개를 자체적으로 구성하여 데이터 베이스화하여 인식 실험을 수행하였다. 실험 환경은 노트북 PentiumIII (450MHz)에서 Visual

C++을 이용하여 우편번호 추출 및 인식을 할 수 있는 시스템을 구현하였다.

4.2 우편 번호 인식기에 대한 검증

그림 16은 데이터 베이스화된 우편 봉투로 명도 값이 높고 문자가 흐린 입력 영상에 대해 그림 17은 개선된 지역적 이진화 기법으로 9×9의 윈도우를 취하고 임계값을 다르게 적용하여 특정 부분의 적당한 임계값을 정해줄 수 있는 블록 이진화 기법을 사용한 것과 그림 18은 소벨 필터를 적용한 영상이다. 그림에서처럼 블록 이진화 기법을 적용하여 우편번호 영역의 추출에 유리하게 블록 평균값에 일정한 임계값의 차를 적용하였지만 추출된 우편 번호 패턴이 흐린 문자 '2'와 '0'에서 문자의 방향 성분의 손상

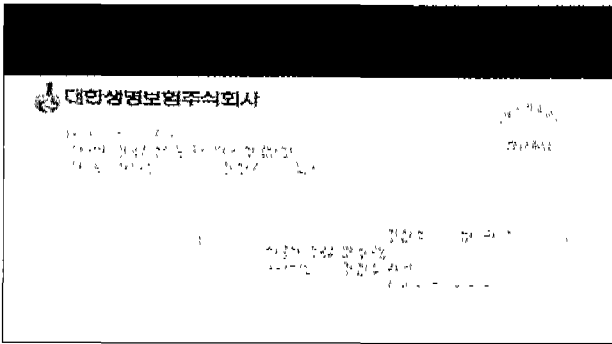


그림 16. 명도가 낮은 우편번호 영상



그림 17. 블록 이진(9×9) 그림 18. 소벨 필터의 적용

이 많음을 알 수 있다. 반면에 소벨 필터를 적용한 것은 우편 번호 문자의 영상이 진하게 나타나고, 흐린 부분에 대한 수평과 사선의 방향 성분이 두드러지게 나타났다.

그러므로 전역적 이진화 보다는 지역적 블록 이진화에 임계치를 적용하면 어느 정도 우편 봉투 영상의 획득이 나옴을 알 수 있다. 그러나 영상마다 문자의 흐림 정도가 동일하지 않기 때문에 우편 영상의 블록 검출시 에러의 유발성이 존재한다. 또한 주소 영역을 살펴보면 우편 봉투 영상이 높은 그레이 값을 갖고 우편 번호 문자가 흐린 경우에는 블록 이진화에 일정한 임계치를 적용해도 이진화시 글자에 대한 영상이 흐려짐을 알 수 있다.

따라서 대부분의 다량 우편물은 수평으로 문자 블록이 일정하므로 우편 번호 블록 검출 시 흐린 영상과 글자와 배경의 그레이 값의 차가 적은 우편물에서도 적용할 수 있게 일반적인 이진화 전 소벨 연산에 의한 우편 봉투를 수평과 수직 방향으로 미분함으로써 이진화 시 문자의 특징이 손상되는 것을 최소화할 수 있으며, 미분된 영상을 반전함으로써 우편 번호 영역 추출시 문자 블록 영역 추출을 손쉽게 구현할 수 있고, 동일한 알고리즘을 적용한 시스템의 프로그램시 보다 손쉬운 알고리즘을 구현할 수 있다.

다음은 50장의 데이터 베이스화된 우편물에 대한 실험 결과이다.

여기서 기각률은 수신자 우편 영역을 검출하지 못하였거나, 문자열 블록에서 우편번호를 잘못 찾은 경우와 우편블록에서 우편 번호 숫자를 제대로 분리하지 못한 경우이다.

표 1. 우편번호 추출의 결과

데이터 베이스 우편물 실험 결과(50장)	ROI 검출	문자열 검출	우편번호 추출	기각 (%)
옳은 추출	48	48	46	8
비율 (%)	96.00	96.00	92.00	
틀린 추출	2	2	4	
비율 (%)	4	4	8	

이렇게 추출된 우편 번호 영상은 클러스터 신경망으로 입력되어 추출된 우편 번호에 대해 다음과 같은 인식 결과를 얻었다.

학습은 Falhman이 제안한 수정된 BP 학습법을 이용하여 학습률과 관성항을 각각 0.1과 0.5로 설정하여 수행하였다.

검출된 우편 영역에 대한 우편 번호 인식률은 다음과 같다.

표 2. 우편번호 인식의 결과

정인식률(%)	오인식률(%)	기각률(%)
91.3	7.5	1.2

이러한 실험 결과는 최순만과 오일석[12]의 우편번호 인식 시스템과 비교해 볼 때 표 3과 같이 우편 영역 추출률과 인식률이 향상되었다.

그리고 비교 논문은 수평 에지를 검출해 선분을 집단화하는 방법으로 우편 번호 영역을 추출하고 추출된 우편 번호 영역에 히스토그램으로 문자를 분리하는 방법을 택하고 있다 [12]. 그러나, 이 시스템은 규격봉투의 네모칸 안에 쓰여진 문자만을 추출할 수 있어서 최근에 증가된 규격 외 봉투의 다량 우편물에는 적합하지 못하다. 또한 우편번호 추출 시스템만을 구현하였고, 숫자 인식기는 다른 필기 숫자 인식기에서 테스트하는 방법으로 구현하여서 실제 시스템의 적용 시 어려움이 있다. 여기의 필기 숫자 인식기의 구조는 10개의 x-인식기(x=0, 1, 2, ..., 9)로 구성되는데, x-인식기는 입력패턴이 x 부류에 속하는지 아닌지에 따라 '예/아니오' 형태의 답을 하는 이진 분류기이다. 어떤 입력 패턴이 들어오면 이를 0-인식기, 1-인식기, ..., 9-인식기 각각에 입력하여 '예'라고 답하는 인식기의 부류로 분류한다. 그러나 이 인식기는 입력 패턴에 비교적 잡영이 적다는 가정을 하고 개발되었으므로, 우편 봉투의 다양한 필기체 숫자 패턴을 입력하면, 오인식 또는 기각하는 경우가 비교적 많이 발생한다. 상기 논문과의 성능 비교는 다음과 같다. 여기서 A는 위에서 설명한 논문의 시스템이며, B는 본 논문에서 구현한 시스템의 성능이다.

표 3. 추출 및 인식의 성능 비교

구분	우편 번호 추출(%)		우편 번호 인식(%)		기각
	추출	오검출	분류	오분류	
A	87.8	12.2	81.6	2.6	15.8
B	92.0	8.0	91.3	7.5	1.2

5. 결 론

본 논문에서는 시스템의 성능 평가를 위해 창과 명도값이

높은 색깔이 있고 문자가 흐린 다양한 우편물 50장에 대해 실험을 실시하였다.

본 방법은 에지 검출에 의한 이진화를 먼저 이용함으로써 우편 번호 영역 추출 및 인식에서 이진화를 먼저 수행할 때 보다 문자의 방향 특징이 손상되는 것을 최소화할 수 있으며, 미분된 영상을 반전함으로써 우편 번호 영역 추출시 문자 블록 영역 추출을 손쉽게 구현할 수 있고, 동일한 알고리즘을 적용한 시스템의 프로그램시 보다 손쉬운 알고리즘을 구현할 수 있다. 실험 결과 우편 번호의 검출이 향상되었으며 인식부에서는 망 특징 및 복합특징의 신경망으로 인식면에서도 실용적인 시스템을 Visual C++를 이용하여 구현하였다.

최근 이러한 우편 번호 인식 시스템은 OCR, 순로구분기 등을 실제 집중국에서 오래 전부터 사용하여 왔지만, 대부분 NEC, 지멘스, 도시바 등의 수입업체의 제품에 의존하여 왔다. 더욱이 이런 제품은 가격이 높을 뿐만 아니라 기술의 국산화라는 면에서 바람직하지 못하다. 따라서 숫자뿐만 아니라 주소 등의 이론적 연구와 더불어 실제 시스템에 적용될 수 있는 연구가 계속 수행되어야 할 것이다.

참 고 문 헌

[1] Nobuyuki Otsu, "A threshold selection method from gray-level histograms," *IEEE trans. Systems, Man and Cybernet*, vol. SMC-9, pp. 62-66, Jan. 1979.

[2] W.H. Tsai, "Moment-Preserving Thresholding: A New Approach," *Computer Vision Graphics, and Image Processing*, vol. 29, pp. 377-393, 1985.

[3] 오균, "문서영상 이진화 알고리즘에 대한 체계적인 평가," *연세대학교 대학원 석사학위 논문*, Dec., 1996

[4] M.Kamel and A. Zhao, "Extraction of Binary Character / Grapgics Image," *CVGIP: Graphical Models and Image Processing*, vol. 55, no. 3, pp. 203-217.

[5] 전병태, 윤호섭, 배창석, 민병우, "명암 상태가 불균일한 영상의 이진화," *제 1회 문자인식 워크샵 발표 논문집*, pp. 77-80, 1993.

[6] J. R. Parker, *Algorithms for Image Processing and Computer vision*. Wesley Computer Publishing, 1996.

[7] H. Yan, "Design and implementation of optimized nearest neighbor classifiers for handwritten digit recognition," *ICDAR' 93*, pp. 10-13, 1993.

[8] L. Wang and T. Pavlidis, "Direct gray-scale extraction of feature for character recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, Oct. 1993.

[9] Carl G. Looney, *Pattern Recognition Using Neural Networks*. Oxford University Press, pp. 75-107, 1997.

[10] A. Rajavclu, M. T. Musavi and M. v. Shirvaikar, "A Neural Network Approach to Character Recognition," *Neural Networks*, vol. 2, pp. 387-393, 1989.

[11] 엄상수, 김종식, 홍연찬, "인쇄체 숫자와 필기체 숫자의 효율적인 통합인식 시스템에 관한 연구", *한국퍼지 및 지능시스템 학술대회 논문집*

[12] 최순만, 오일석, "한국 우편봉투 영상에서 우편번호 추출 및 인식을 위한 시스템", *정보과학회논문지(C) 제 2권, 제 1호*, pp. 94-104, 1996

저 자 소 개



김 종 석 (Jong-Seok Kim)

1997년 2월 : 인천대학교 전자공학과 (공학사)

2000년 8월 : 인천대학교 전자공학과(공학 석사)

2000년~현재 : LG전자(주) CDMA 단말연 구소(연구원)

관심분야 : 신경망 시스템 및 퍼지이론, 지능제어시스템, 로봇 텍스, 영상인식, 컴퓨터비전

홍 연 찬 (Yeon Chan Hong)

정회원

제 7권 제 3호 참조