

자소 접촉특성 분석에 의한 한글패턴의 부분분리 및 인식

正會員 高 燦* 正會員 陳 庸 玉**

Separation of Subpattern and Recognition of Hangul Patterns by Analysis of Feature of Contacting Phonemes

Chan KOH*, Yong Ohk CHIN** *Regular Members*

要 約 본 논문은 한글문자 인식을 위하여 새로운 접촉자소의 분리 및 자획연결 특징추출 알고리즘을 제안하였다. 자소 접촉특성을 분석하여 자소접촉 형태별로 분류하고 자획연결특징 추출, 접촉자소의 분리, 문자형식 분류를 시행한다. 분리된 자소로부터 설정된 표준패턴으로 정규화하고 자소별 굴곡특징의 상대위치값으로부터 특징을 입력패턴으로 신경망을 이용하여 인식 실험을 하였다. 여기에서의 학습은 BEP 알고리즘을 이용하였다. 접촉자소의 분리, 형식분리, 자획연결특징 추출 및 인식 실험에서 제안된 알고리즘이 좋은 결과를 나타내었다.

ABSTRACT In this paper a new algorithm for separation of contacting subpattern and connective feature extraction of strokes is proposed. This algorithm is able to classification of the types of contacting parts, connective feature extraction of strokes, separate the phoneme of contacting parts between strokes, classify the character types by feature classification of connecting parts and analysis of connecting attribute. Also, shape normalize into formal patterns and decide on the input pattern from position value of bending feature of this normalized shape and make an recognition experiment by neural network using BEP learning algorithm. This algorithm represents the good achievement ratio by separation of phoneme, classification of character type, connective feature extraction of stroke and recognition experiment.

I. 서 론

한글 정보처리 자동화를 위해 한글 패턴의

자동인식에 관한 연구가 많이 진행되고 있다. 한글은 자음과 모음을 포함하여 2~7개의 자소가 조합하여 최대 14,364자를 이루고 보통 사용하는 문자만도 1,600여자에 이른다. 한글 문자인식에서 패턴을 분석하여 문자 기본요소를 찾아 이의 구조분석에 의해 분리 조합하는 방식으로 인식하는 구조적 방법은 크게 나누어 세션화하여 인식하는 방법과^{1) 2)} 형상패턴의 윤곽선에 의하여

* 서울産業大學 電子計算學科
Seoul National Polytechnic University
Dept. of Computer Science

** 慶熙大學校 電子工學科
Dept. of Electronics Eng. Kyung Hee University
論文番號 : 90-62 (接受 1990. 6. 25)

인식하는³⁾ 방법이 연구되었다. 세션화 처리시 변형과 오인식의 원인을 초래하게 되는 문제점으로 인하여 세션화처리를 하지 않고 자소를 추출하려는 시도가 있다. 한글의 기하학적 특징과 구조적 기본 요소간의 상호관계를 분석하여 한글을 표현하여 인식하는 연구와⁴⁾ 필기체에서 필자의 습관에 따른 변형에 유연성을 갖도록 형상패턴에 의한 방법의 연구가 수행되었다⁵⁾. 한글의 형상을 그 윤곽선을 기준으로 하여 내부로의 거리변환값에 따라 특징들을 찾아 패턴을 분류하고 인식하는 방법의 연구등이⁶⁾ 수행되었다. 한글 인식에서 그 형상전체를 대상으로 하여 특징을 찾아 인식하는 것이 변형에 따른 영향을 적게 받게 하고 유사문자 인식에 강력하다. 문자인식에서 가장 중요한 문제점은 자소간의 접촉부분의 불확실성이다. 자소가 접촉되어 다른 자소로 판명되거나 존재하지 않는 자소가 되기도 한다. 이 접촉부분을 찾아 자동분리토록 하여야 한다. 자소분리가 원활히 이루어지는 것이 문자인식에서 중요한 과정이 된다.

본 연구에서는 한글패턴중 자소가 접촉되는 부분을 분석하여 그 접촉모양을 형태별로 분류하고 자소접촉 특징을 추출한다. 자획을 연결특성에 의해 자소로 구성하고 자소간 접촉부분은 추출된 접촉특징에 의해 자소분리한다. 또한 여러가지의 자획연결 특징을 추출하고 문자를 형식별로 분류한다. 분리된 자소를 자소별 번호를 부여하고, 각 자소에서 굴곡 특징의 상대위치 정보를 이용하여 인식 실험을 위한 신경망의 입력패턴으로 만들어준다. 문자인식에서, 인식실험에서의 데이터양의 감소를 위해 보통 전처리단계를 두어 특징을 추출하여 이를 이용하게 된다. 이 전처리 단계에서 주로 해석적 방법이 많이 적용되고 있으므로 이를 위해 문자를 일정한 표준형태로 정규화하면 많은 장점을 갖는다. 입력패턴을 한글패턴형식의 표준형상을 정하고 이 표준형에 맞도록 형상변형을 실시한다. 기본 표준형에 맞도록 변형된 패턴에서 각 문자들이 특징점들의 상대위치 정보를 추출해 낼 수 있다. 이 특징값 상대위치 정보들로 하나의 입력패턴을 구성한다. 이 입력패턴을 신경망을 이용

하여 인식실험 하였다. 여기에서의 학습은 BEP 알고리즘을 이용하였다. 자소의 분리가 원활히 이루어지면 인식을 크게 향상시킬 수 있다. 실험에서의 사용 데이터는 참고문헌(8)의 일상사용문자 조사를 기준으로 977자중에서 접촉부위가 발생하는 형태별로 분류하여 실험하였다. 본 연구에서의 실험결과들은 한글문자인식 시스템에서 활용될 수 있는 형태로 출력하였다. 입력장치로는 가격이 저렴한 간이스캐너를 사용하였다. 2장에 굴곡 특징점 및 접촉특징점 추출, 3장에 자소인식처리, 4장에 실험 및 결과고찰, 5장에 결론을 논하였다.

II. 굴곡특징점 및 접촉특징점 추출

굴곡특징점 및 접촉 특징점 추출은 스캐너로 입력된 문자 패턴을 문자가 존재하는 최소장방형을 취하여 이 사각형을 처리하도록 한다. 이 영역의 문자정보만을 대상으로 하여 처리시간을 단축하도록 한다. 이 장방형의 네모퉁이를 최상단점(t), 최하단점(b), 최우측점(r), 최좌측점(l)로 표시한다.

(1) 접촉 형식 분류

많이 발생하는 자소접촉 형태를 분석하여 다음의 5가지 형식으로 분류한다.

- i) 접촉 1형식 : 수직 종모음과 초성자음이 접촉된 형식

건	견	면
.....77777...777777....	9999222221....
.....77777....777779....	9999233321....
.....7779.....12333399999777779....
.....123399999991232229999912343999999
.....12229999999123211199.9912332999999
.....922119999997777.....12321999999
...7977.....	...92211.....7777.....
	7777.....

상단(↑) 하단(↓) 좌측(←) 우측(→) 특징

777	777	다른굴곡 존재시 ***	blk	← ***
-----	-----	-----------------	-----	-------

***좌측에 굴곡 성분이 존재시 1형식 접촉

ii) 접촉 2형식: 수직 종모음과 종성자음이 접촉된 형식

갑	격	갈
.....7777..777777.	...777....
....977777.777777.	...979....
999999977777.	..999977777.	.999799...
999999934321.	999999934321.	99992991..
999999977777.	999999977777.	99991991..
.....777777.777777.

상단(↑) 하단(↓) 좌측(←) 우측(→) 특징

777	***	하단연결에 굴곡	하단연결이 굴곡	↓*** ←→
-----	-----	-------------	-------------	------------

***하측으로 2/3하단에, 좌우에 굴곡 있으면 굴곡부분이 2형식접촉.

iii) 접촉 3형식: 수평모음과 초성자음이 접촉된 형식

고	구	그
....777...7777...	...11....
....777....7779....	...11....
...979....	...92399999	..779....
.999399999	11992299999	..9299999
9999299999	21991199999	999299999
9999199999	7.....	999199999
.....

상단(↑) 하단(↓) 좌측(←) 우측(→) 특징

777	blk	999	999	*** ↑
-----	-----	-----	-----	----------

***상단에 굴곡 존재시 3형식 접촉.

iv) 접촉 4형식: 수평모음과 종성자음이 접촉된 형식

눅	꿀	물
..7777.....	..77777..	..77777...
..9779.....9779..	...7777...
99922999999	999933999979...
99911999999	999922999	9999939999
99.....	999911...	9999929999
.....	999991....

상단(↑) 하단(↓) 좌측(←) 우측(→) 특징

777	blk	999	999	*** ←→
-----	-----	-----	-----	-----------

***좌우에 굴곡 성분 존재시 4형식 접촉.

v) 접촉 5형식: 수평모음과 수직모음이 접촉된 형식

과	괘	귀
...77777..		
...77777..7777797777.
...123329977777.7777.
...1234399	..1777777.	..111777777.
...1233299	111777777.	111...7777.
11777777..	1...77777.7777.
11777777..		

상단(↑) 하단(↓) 좌측(←) 우측(→) 특징

777	777	999	blk	***←
-----	-----	-----	-----	------

***좌측에 굴곡 있으면 5형식 접촉.

(2) 모음의 특징 분류

굴곡특징부분을 중심으로 다음의 특징 만족여부 검사

a) 상단(↑) 하단(↓) 상단좌측(←)

777	777	blk
-----	-----	-----

b) y방향 1/3상단 y 방향 1/3 하단에 좌, 수평 성분이 없어야 수직성분으로 판정

c) 수직방향으로 시작점부터 수직성분의 연결이 상대길이 80% 이상 일 때 수직 장모음으로 판정(문자형식 1, 문자형식 5존재)

d) 굴곡특징 부분 중심으로 좌우측에 수평성분 존재하고 상대길이 x축방향 80% 이상이며

1) y축 상대길이 70% 이하에 존재시: ㄴ, ㄹ, ㄴ로 좌측

2) y축 1/3~2/3 사이에 존재시: ㄷ, ㅌ로 판정

e) 모음의 형태와 받침 유무 검사

f) 가지의 갯수, 수평, 수직 장모음에 가지부분의 연결내용

g) 수직장모음은 tbrl의 1/3부근에서 시작하며 "l"을 기준으로 한다.

중심부근에 우측, 좌측에 수평가지가 존재하면 "ㄷ", "ㄹ"이다.

h) 중심상하에 대칭으로 가지가 좌우측에 존재하면 "ㅈ", "ㅋ"로 판정한다.

- i) 다시 우측에 수직장모음이 존재하면 “ㅈ”, “ㄱ”, “ㅊ”, “ㅋ”로 판정한다.
- j) 수평장 모음은 tbrl의 x방향의 1/3부근에서 시작하여 “ㅡ”를 기준으로 한다. 중심부근에서 상하로 수직거리가 있으면 “ㄴ”, “ㄷ”이고 중심 좌우대칭으로 가지가 존재하면(2개) “ㄹ”, “ㅍ”로 판정한다.

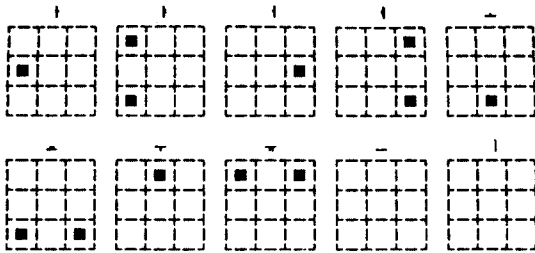
1) 시작점 특징 :

수직좌 ㅏ, ㅑ 수평상 ㅍ, ㅑ
 수직우 ㅓ, ㅕ 수평하 ㄴ, ㄷ

2) 굴곡갯수 특징

굴곡갯수	수직(9)	수평(5)	복합(7)
0	ㅣ	ㅡ	ㅏ
1	ㅏ ㅑ ㅓ	ㄴ ㄷ	ㅓ ㅕ
2	ㅑ ㅕ ㅗ ㅛ	ㄹ ㅍ	ㅗ ㅛ ㅜ
3			ㅜ
4	ㅗ		

1) 굴곡의 상대위치 분포



(3) 특징점 검출

굴곡특징 부분(굴곡부분길이 가로 a×세로 b)을 중심으로 상, 하, 좌, 우, 좌상, 우상, 좌하, 우하로 a×b 만큼씩 특징을 검사하여 아래의 성분이 만족시 특징점으로 판정한다. 8방향의 가로, 세로의 크기를 굴곡부분의 크기와 같게 하였으며 8방향의 성분중 표준설정 내용과 80% 이상이면 해당되는 것으로 판정한다. 특징점을 중심으로 3×3 마스크를 4분면으로 회전시키며 검사한다. 특징점 중심으로 우상단의 3×3 화소, 하상단으로 3×3 화소, 좌하단으로 3×3 화소, 좌상단으로 3×3 화소 순으로 검사한다.

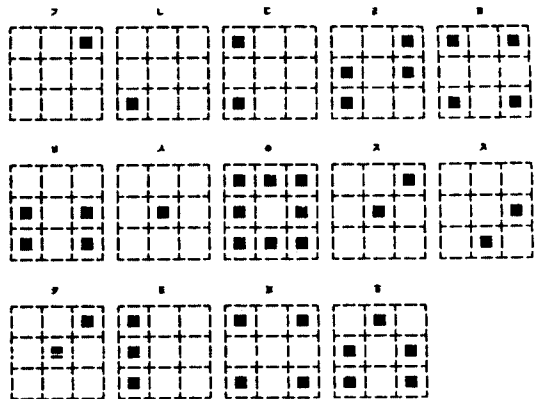
표에서 “7”은 수직성분 화소이고 “9”는 수평성분 화소이며 각각 가로 X 세로 방향으로 굴곡 특징의 가로×세로 길이만큼이 검사 대상이 된다. 여기에서는 예를 나타내기 위하여 가로로 3개 화소값만 표시하였다.

상단↑ 하단↓ 좌측← 우측→ 사선↗

a(ㄱ)	blk	777	blk	999	blk
b(ㄴ)	blk	777	999	blk	blk
c(ㅏ)	777	blk	999	blk	blk
d(ㅑ)	777	blk	blk	999	blk
e(ㅓ)	777	777	blk	999	blk
f(ㅕ)	777	777	999	blk	blk
g(ㅗ)	777	blk	999	999	blk
h(ㅛ)	blk	777	999	999	blk
i(ㅜ)	777	blk	099	990	blk
	우상	좌하	좌상	우하	하단
	blk	999	blk	999	blk

(4) 자음의 특징 분류

a) 굴곡 상대위치 분포



ㅈ, ㅊ과 ㅍ, ㅑ은 굴곡 특징이 유사하고 ㅇ, ㅎ은 원형성분으로 구별에 문제점을 갖고 있으므로 입력데이터 구성시 변형을 추가하였다.

b) 굴곡특징 갯수

굴곡갯수	초성(19)		중성(27)	
	단자음(14)	복자음(5)	단자음(14)	복자음(13)
0	ㅇ		ㅇ	
1	ㄱ ㄴ ㄷ		ㄱ ㄴ ㄷ	ㅇ
2	ㄷ ㅌ ㅍ	ㄱ ㅋ	ㄷ ㅌ ㅍ	ㄱ ㅋ ㄴ
3	ㅌ		ㅌ	ㄴ
4	ㄹ ㅁ ㅂ ㅍ	ㄷ ㅌ	ㄹ ㅁ ㅂ ㅍ	ㄷ ㅌ ㄹ
5				ㄹ ㄹ ㅁ
6				
7				ㄹ
8		ㅁ		ㅁ ㄹ ㄹ ㄹ

c) 시작점

좌상측 (ㄱ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅌ, ㅍ, ㅌ, ㅍ)

중앙상측 : ㅌ

독립점 : ㅌ, ㅎ

원요소 : ㅇ

III. 자소인식처리

문자인식처리에서 문법적방법의 단점을 해결하고자 인공신경망을 이용한 실험을 하였다. 문법적인 방법에서는 모든 경우의 생성규칙을 기술해야만 하고 생성규칙이 복잡하고 수정이 어렵다. 또한 생성규칙에 정의되지 않은 패턴은 인식될 수 없고, 문법의 기술이 방대하고 소프트웨어 작성에 많은 시간이 소요된다. 한글은 문자수가 방대하고, 문자간 유사성으로 인하여 이전 패턴을 바로 신경회로망에 입력으로 사용하기에는 데이터양이 과중하여 입력패턴의 특징추출값을 상대 위치벡터로 구성한 입력값을 신경회로망의 입력으로 사용하였다. 최대 64개가 입력노드의 최대수가 된다. 출력은 초성, 중성, 종성에서 최대 5비트를 필요하므로 5개 출력노드로 하였다. 문자인식에서 유사성의 제거와 정보분산을 유도하여 한글인식에 도움이 되도록 출력층의 노드를 초성, 중성, 종성의 갯수가 49개가 되도록

한 연구가⁽¹⁶⁾ 있으나 본 논문에서는 자소가 같은 경우 초성모음이나 중성모음의 모양이 달라도 굴곡 특성과 상대위치를 이용하였으므로 5개의 출력노드로 하여 이의 조합으로 32개 이내에서 정할 수 있도록 하였다. 여기에 적용한 신경회로망은 다층인식자이고 적용된 훈련방법은 BEP 알고리즘을 사용하였다^(14, 15). 다음은 BEP 알고리즘의 간략화한 내용이다. BEP 알고리즘은 지도 학습알고리즘으로 패턴에 대하여 기대출력치와 신경망의 실제출력치와의 차이를 최소화하려고 가중치를 구해내는 알고리즘으로 일반화 델타를 사용하며 내부층을 포함하는 전진다중계층 인공신경망 구조를 갖는다.

학습단계는

- 1) 각 유니트에 대해서 출력 O_i 계산
- 2) 기대 출력치와 실제 출력치와의 차이를 계산해서 이 차이를 역전파시키며 가중치를 조절한다.
- 3) 전체 오류의 합이 오류범위를 만족시키기까지 반복 시행

$$E = \sum (t_i - O_i)^2$$

입력층위의 중간층과 출력층의 각 노드 i 의 실제 출력값은 다음 형태의 Sigmoid 함수로 주어지며 실제 출력값의 범위는 0.0에서 1.0사이이다.

$$O_i = \frac{1}{1 + \exp(-nt_i)}$$

$$nt_i = \sum w_{ij} o_j + \theta_i$$

여기서

nt_i : 입력치

O_i : 앞의 실제출력치

W_i : 가중치

θ_i : 유니트 i 의 bias

출력층 유니트의 $\delta_i = (t_i - O_i) O_i(1 - O_i)$

내부층 유니트의 $\delta_i = (O_i(1 - O_i) \sum_k \delta_k W_{ki})$

가중치 변화량 계산식은 다음과 같다.

$$\Delta w_{ij}(n+1) = \eta \delta_i o_j + \alpha \Delta w_{ij}(n)$$

여기서 n 은 시간증분, η 는 학습율, α 는 관성항이다.

자소분리된 패턴에서 자소별 tbrl 윈도우를 정의하여 각 윈도우별 크기를 $8 \times 8 = 64$ 입력패턴이 되도록 규격화한다. 규격화 방법은 다음과 같다.

1) 윈도우내에 굴곡부분의 특성값 부위를 가로길이 \times 세로길이의 평균값을 구하고 자소별 tbrl에 비례로 가로 $8 \times$ 세로 8화소로 사상한다.

사상의 방법은 다음식으로 하고 정수값을 취하고 동일점으로 정의되는 부분을 결정하여 시행한다.

자소윈도우의 x, y 방향 최대, 최소점의 좌표값 ($X_{w,max}, y_{w,max}$ 과 $X_{w,min}, y_{w,min}$)을 규격화 입력패턴 가로 \times 세로 (8×8 화소)로 사상한다.

8×8 규격화 윈도우내의 점(X_v, Y_v)로 변환식은

$$(X_v, Y_v) = (X_w, Y_w) \cdot \begin{pmatrix} s1 & 0 \\ 0 & s2 \end{pmatrix}$$

$$s1 = \frac{8}{X_{w,max} - X_{w,min}}$$

$$s2 = \frac{8}{Y_{w,max} - Y_{w,min}}$$

$$X_v = X_w \times s1$$

$$Y_v = Y_w \times s2$$

소수 이하를 버리고 정수를 취하고 동일점은 단일값으로 처리한다. 8화소를 3, 2, 3화소씩 9등분 하여 각각의 4각형 안의 화소수의 비율을 0에서 1사이의 값으로 하여 입력패턴으로 하였다.

IV. 실험 및 결과 고찰

특징점검출 실험은 입력패턴에서 모음의 인식을 먼저하고 자음을 인식하도록 실험하였다. 거리변환 패턴에서 수평, 수직 방향값을 부여하여 굴곡부분을 검출하고 이를 다시 a(ㄱ), b(ㅋ), c(ㆁ), d(ㄴ), e(ㄷ), f(ㅌ), g(ㄹ), h(ㅍ), i(ㅅ) 굴곡점으로 특징 추출하였다. 자소접촉부분의 검출은 자소가 접촉된 특징별로 접촉 5형식으로 구분하여 j(접촉1), k(접촉2), l(접촉3), m(접촉4), n(접촉5)으로 특징 추출하였다. 형식분류 및 특징벡터 추출 실험은 자소접촉형식 분류와 수평, 수직, 장모음의 검출, 받침유무, 굴곡갯수, 시작점위치, 장모음과 가지의 연결내용, 굴곡점의 상대위치값으로 문자의 형식을 분류하고 이들 값을 이용하여 입력패턴을 구성하였다. 입력패턴은 자소별 $8 \times 8 = 64$ 화소가 되도록 규격화하였다. 가로, 세로로 3, 2, 3으로 9등분하여 각 4각형안의 화소수의 비율을 0.0에서 1.0사이의 값으로 변환시켜 입력 패턴 값으로 하였다. 인식 실험은 BEP 학습 알고리즘을 이용하여 인식 실험하였다. 특징점 추출, 접촉분리, 형식분류 실험에서 대부분의 입력패턴에서는 좋은 결과를 얻었으나 일부 심하게 편향 입력된 패턴이나 접촉부분이 중복되어 나타난 경우에는 실패하였다. 그림 1은 입력패턴, 그림 2는 거리변환 패턴, 그림 3은 굴곡특성 패턴, 그림 4는 특징 추출 패턴, 그림 5는 자소별 입력값, 출력값 및 학습값 예, 그림 6은 특징의 상대위치 및 입력값, 그림 7은 인식처리결과 예를 나타낸다.

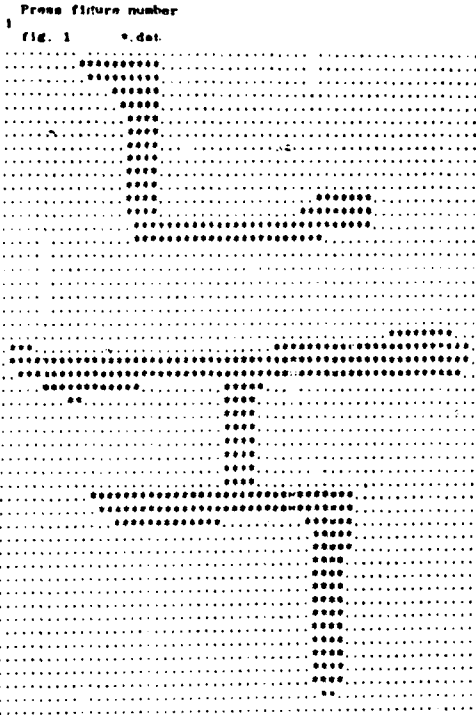


그림 1. 입력패턴
Input pattern

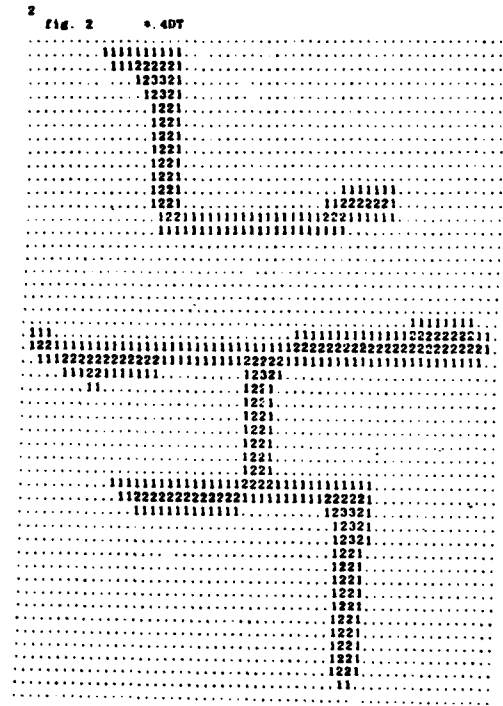


그림 2. 거리변환패턴
Distance transform pattern

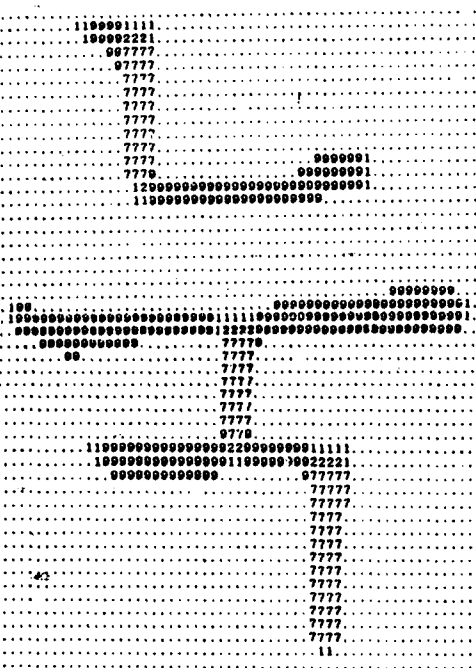


그림 3. 굽곡특징 패턴
Bending feature pattern

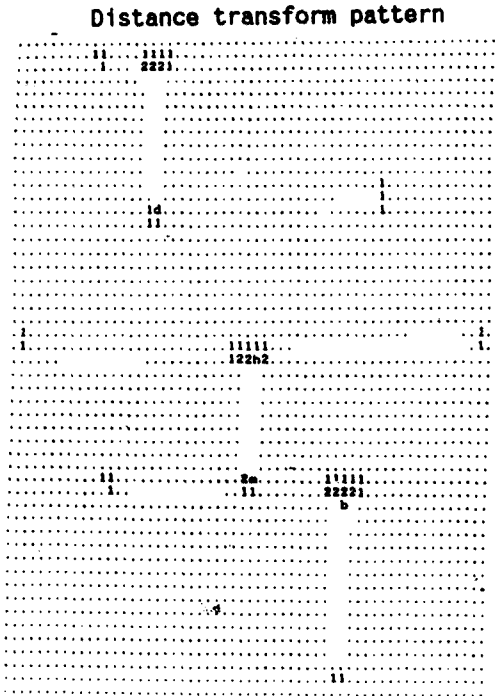


그림 4. 특징추출 패턴
Feature extraction pattern

"ㄱ"	입력값 :	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
	출력값 →	0.1	0.1	0.1	0.9					
	인식값 →	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
"ㄴ"	입력값 :	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
	출력값 →	0.1	0.1	0.9	0.1					
	인식값 →	0.100001	0.100256	0.900551	0.105390					
"ㄷ"	입력값 :	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
	출력값 →	0.1	0.1	0.9	0.9					
	인식값 →	0.099999	0.099859	0.899423	0.895469					
"ㄹ"	입력값 :	0.0	0.0	1.0	1.0	0.0	1.0	1.0	0.0	0.0
	출력값 →	0.1	0.9	0.1	0.1					
	인식값 →	0.100002	0.899850	0.099470	0.094513					
"ㅁ"	입력값 :	1.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	1.0
	출력값 →	0.1	0.9	0.1	0.9					
	인식값 →	0.100000	0.900007	0.100469	0.902650					
"ㅂ"	입력값 :	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	출력값 →	0.1	0.9	0.9	0.9					
	인식값 →	0.100001	0.900033	0.900083	0.901242					

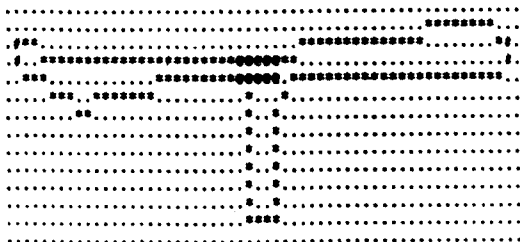
그림 5. 자소별 입력값, 출력값 및 학습값 예
Example of I/O and learning value of each phoneme



(a) 굴곡 특징의 상대위치

scf=	4.950000
f1=	0.000000
f2=	0.000000
f3=	0.000000
f4=	0.000000
f5=	0.000000
f6=	0.000000
f7=	0.000000
f8=	0.000000
f9=	0.000000

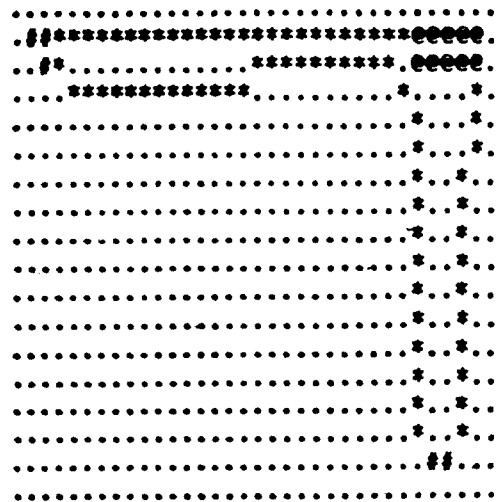
(b) 자소의 계산된 입력값



(a) 굴곡 특징의 상대위치

scf=	7.560000
f1=	0.000000
f2=	1.000000
f3=	0.000000
f4=	0.000000
f5=	0.000000
f6=	0.000000
f7=	0.000000
f8=	0.000000
f9=	0.000000

(b) 자소의 계산된 입력값



(a) 굴곡 특징의 상대위치

scf=	6.912001
f1=	0.000000
f2=	0.000000
f3=	0.960000
f4=	0.000000
f5=	0.000000
f6=	0.000000
f7=	0.000000
f8=	0.000000
f9=	0.000000

(b) 자소의 계산된 입력값

그림 6. 특징의 상대위치 및 입력값
Relative position of feature and Input value

"ㄴ"	ㄴ :	입력값	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0
		출력값	0.100001	0.100256	0.900551	0.105390					
		인식	0.1	0.1	0.9	0.1					
"ㄷ"	ㄷ :	입력값	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		출력값	0.100001	0.900033	0.900083	0.901242					
		인식	0.1	0.9	0.9	0.9					
"ㄱ"	ㄱ :	입력값	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
		출력값	0.099998	0.099958	0.100008	0.899549					
		인식	0.1	0.1	0.1	0.9					
"ㄹ"	ㄹ :	입력값	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0
		출력값	0.099998	0.098684	0.140777	0.885383					
		인식	0.1	0.1	0.1	0.9					
"ㅁ"	ㅁ :	입력값	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		출력값	0.100001	0.900033	0.900083	0.901242					
		인식	0.1	0.9	0.9	0.9					
"ㅂ"	ㅂ :	입력값	1.0	0.0	0.5	0.0	0.0	0.0	0.3	0.0	0.1
		출력값	0.099998	0.114321	0.391878	0.960908					
		인식	0.1	0.9	0.1	0.9					
"ㅅ"	ㅅ :	입력값	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		출력값	0.100001	0.878044	0.894717	0.895933					
		인식	0.1	0.9	0.9	0.9					
"ㅈ"	ㅈ :	입력값	0.0	0.0	1.0	1.0	0.0	0.5	0.4	0.0	0.0
		출력값	0.100001	0.819871	0.093095	0.227665					
		인식	0.1	0.9	0.1	0.1					

그림 7. 인식 처리 결과 예
Example of result of recognition experiment

V. 결 론

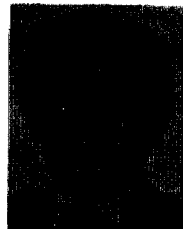
본 논문에서는 신경망 구조중의 하나인 전진다중계층 인공 신경망에 BEP 학습 알고리즘을 이용하여 인식실험을 하였다. 간이스캐너를 사용하여 입력된 2진문자 영상을 굴곡 특징 분석에 의해 구한 상대위치값을 여기에서의 입력패턴으로 하였다. 실험은 IBM PC/AT 호환기종(16MHz)에서 Turbo C2.0을 사용하여 시뮬레이션 하였고 1문자당 특징 추출은 평균 2초가 소요되었고 학습시간은 자소당 평균 12분이 소요되었다. 자소가 접촉되는 부분을 접촉행태별로 분류하고 자소를 분류하여 자소별 특징추출과 자회연결 특징을 추출하는 알고리즘과 이 특징들로 신경망의 입력패턴으로 구성하도록 하는 알고리즘을 제안하였다. 실험결과 굴곡특징을 이용한 특징추출이 입력 문자 변형에도 유연성을 보임을 알 수 있었다. 접촉부분의 분리가 원활히 이루어질 수 있었으나 접촉부분이 심한 중복 경우나 자획의 굵기가 수직, 수평 굵기보다 연결부분이 상대적으로 가늘게 되어 거리변환값 정보를 상실한 경우에 실패의 경우가 많았다. 입력문자의 변형에도 강력하고 자체 변형이 많은 필기체 인식에서도 사용할 수 있는 좀더 유연성을 갖는 알고리즘의 개발이 요구된다.

參 考 文 獻

- 남궁재찬, "Index-window 알고리즘에 의한 한글 pattern의 부분분리와 인식에 관한 연구", 박사학위논문, 인하대학교 대학원, 1982. 2.
- 이주근외 2인, "논리연산에 의한 fast thinning algorithm", 대한 전자공학회 추계학술대회 논문집 : Vol.4, No. 3, pp. 47~49쪽 1981.10.
- 박종욱, 이주근, "Shape pattern에 의한 필기체의 한글 인식", 대한전자공학회지, 제22권, 제5호, 1985. 9.
- 한상기, "구조적 방식에 의한 한글인식", KAIST 석사논문, 1984.
- 박종욱, "도형패턴의 벡터스 검출법과 필기체 한글세이프패턴의 인식" 박사학위논문, 인하대학교 대학원, 1985. 2.
- 고찬, 이대영, "거리변환법에 의한 한글패턴의 특징분류", 한국통신학회 논문지, 제14권, 제6호, 650~662쪽, 1989. 12.
- 고찬, 이대영, "중첩윤곽 형상에 의한 한글패턴의 정점검출", 한국통신학회 논문지, 제15권, 제2호, 112~123쪽, 1990. 2.
- 김상운, "방향투영에 의한 인쇄체 한글의 기본선소추출 및 자동인식에 관한 연구", 박사학위논문, 연세대학교 대학원, 1987. 12.
- 이근하, "속성에 구속받는 문법을 이용한 문자패턴인식", 박사학위논문, 인하대학교 대학원, 1981. 2.
- 오길남, 신승호, 진용옥, "2차원 디지털 필터링에 의한 한글자모의 인식 알고리즘", 전자공학회지, 제21권, 제3호, 55~59쪽, 1984. 5.
- 홍성민, 김은원, 박종국, 조원경, "구조해석에 의한 필기체 한글의 온라인 인식", 전자공학회 논문지, 제23권, 제1호, 1986. 1.
- 최병욱외 2인, "한글인식에 있어서의 자소추출", 대한 전자공학회지, 제18권, 제2호, 36~43쪽, 1981. 4.
- 김태균외 2인, "Stroke 조합에 의한 필기체 한글의 표현과 인식", 전자공학회지, 제25권, 제1호, 18~26쪽, 1988. 1.
- 정주성, 김춘석, 박충규, "윤곽선 추적에 의한 고딕체 한글의 신속인식에 관한 연구", 전기학회논문지, 제37권, 제8호, 579~587쪽, 1988. 8.
- 최정훈, 권희용, 황희용, "통합사용자 인터페이스에 관한 연구 : 인공신경망 모델을 이용한 한글필기체 on-line 인식", 한국어 정보처리 학술발표논문집, 126~131쪽, 1989. 10.
- 김상우, 전운호, 최종호, "신경회로망을 이용한 인쇄체 한글문자의 인식", 전자공학회 논문지, 제27권, 제2호, 228~235쪽, 1990. 2.
- D. E. Rumelhart and J. L. McClelland, "Parallel distributed processing, Vol. 1, Vol.2", MIT press, pp. 318~364, 1986.
- Yoh-Han Pao, "Adaptive pattern recognition and Neural Networks", Addison-Wesley, pp. 111~140, 1989.



高 燦(Chan KOH) 正會員
1951年 9月14日生
1974年 2月：慶熙大學校 工科學 卒業
1987年 8月：慶熙大學校 大學院電子工
學科 博士課程 修了
1987年：情報處理 技術士 資格取得
1974年10月～1978年 2月：海軍大學電算
教官(海軍將校 服務)
1982年 2月～1989年 現在：서울産業大學
電算學科 助教授



陳 庸 玉(Yong Ohk CHIN) 正會員
1943年 3月21日生
1968年：延世大學校 工科學 電氣工學
科 卒業
1975年：延世大學校大學院 電子工學科
工學碩士
1981年：延世大學校大學院 電子工學科
工學博士
1980年：通信技術士
現在：慶熙大學校 工科學 電子工學科 教授