

# 숫자의 형태 이해와 분할된 FSOM을 이용한 필기 숫자 인식에 관한 연구

정회원 서석배\*, 김대진\*\*, 강대성\*

## A Study On Handwritten Numeral Recognition Using Numeral Shape Grasp and Divided FSOM

Seok-Bae Seo\*, Daijin Kim\*\*, Dae-Seong Kang\* *Regular Members*

### 요 약

본 논문에서는 필기숫자의 형태와 FSOM을 이용한 새로운 필기숫자 인식 알고리즘을 제안한다. 제안한 알고리즘은 비슷한 형태를 가지는 숫자는 같은 그룹으로 분류된다는 사실에 기초한 알고리즘으로, 필기숫자의 외접선에 의한 형태결정과 템플레이트 매칭을 이용하여 필기숫자 데이터를 여러 개의 그룹으로 분할하고 분할된 각 그룹별로 인식 알고리즘을 적용한다. 본 실험에서는 필기 숫자의 데이터를 16개의 그룹으로 분류하였으며, 분류된 그룹별로 각각의 특징추출과 SOM의 단점을 보완한 FSOM을 적용하였다. 분할된 16개의 FSOM은 각 그룹별로 독립적인 학습이 가능한 특징뿐만 아니라 한 그룹씩 단계적으로 학습을 할 수 있는 장점이 있다. 제안한 알고리즘의 성능을 입증하기 위해서 Concordia 대학의 필기 숫자 데이터를 사용하여 실험하였다.

### ABSTRACT

This paper proposes a new handwritten numeral recognition method using numeral shape grasps and FSOM (Fuzzy Self-Organizing Map). The proposed algorithm is based on the idea that numeral input data with similar shapes are classified into the same class. Shapes of numeral data are created using lines of external-contact and the class of numeral data is determined by template matching of the shapes. Each class of numeral data has FSOM and feature extraction method, respectively. In this paper, we divide the numeral database into the 16 classes. The divided FSOM model allows not only an independent learning phase of SOM but also step-by-step learning. Experiments using Concordia University handwritten numeral database proved that the proposed algorithm is effective to improve recognition accuracy.

### I. 서론

필기숫자 인식은 30년 이상 꾸준히 연구되고 있으며, 높은 인식률을 갖는 여러 가지 알고리즘이 많이 제안되었다. 제안된 방법들을 살펴보면 크게 두 가지 부류로 나눌 수 있다.

첫 번째 부류는 필기 숫자의 효과적인 특징 추출을 위한 새로운 알고리즘을 제안하는 것으로, 분류

및 인식의 부분은 특징 추출의 특성에 의해 결정하거나 기존의 알고리즘을 응용하는 방법이다. 데이터의 정보를 수치적으로 표현하는 특징추출은 전체 인식 성능의 관건이 되므로 매우 중요하다고 할 수 있다. 필기 숫자 인식에서는 입력 영상을 이진 영상으로 변환한 다음 일정한 영역으로 분할하여 각 영역내의 흑화소의 수를 특징으로 추출하는 방법이나 이진 영상에 대한 좌우 히스토그램 분포를 특징으로

\* 동아대학교 전기전자컴퓨터공학부(sbseo@vlsi.donga.ac.kr, dskang@daunet.donga.ac.kr)

\*\* 포항공과대학교 컴퓨터공학과 (dkim@postech.ac.kr)

논문번호 : 98399-0910, 접수일자 : 1998년 9월 10일

로 추출하는 방법이 널리 사용된다. 이 방법의 경우 알고리즘 구현이 간단하고 처리시간이 빠른 반면 모든 영상에 대하여 일률적인 특징을 추출하므로 많은 저장 공간이 필요한 단점이 있다. 이를 개선한 효율적인 특징 추출에 관한 연구로 Gabor feature에 의한 방법<sup>[1]</sup>, 필기 숫자의 Contour 정보를 이용한 방법<sup>[2]</sup>, 필기 숫자의 구조적 정보(structural information)를 이용한 방법<sup>[3]</sup> 등 많은 방법이 제안되었다. Gabor feature를 이용한 방법은 spatial 영역과 spatial-frequency 영역에 대하여 optimal joint localization을 갖는 Gabor filter를 이용하여 인간의 시각 시스템(human visual system)과 유사한 구조로 특징을 추출하는 것으로, 필기 숫자뿐만 아니라 인쇄체 한자 인식에도 적용되어 높은 인식률을 보였다. 필기 숫자의 구조적 정보를 이용한 방법에서는 입력 필기 숫자의 벡터화(vectorization), line-drawing 생성, line-drawing 구조 해석, structure graph 생성, 프로토타입 매칭(prototype matching)의 순서로 특징 벡터를 추출하며, 생성된 특징 벡터를 그래프 매칭법과 신경회로망을 이용하여 인식하였다. Contour 정보를 이용한 방법에서는 필기 숫자에 대하여 contour smoothing, CFAB(Contour Following Algorithm with Backtracking), Fourier descriptor 등을 사용하여 특징을 추출한 다음 기하학적 성질(topological property), 출발점의 위치, 통계학적 해석 등에 기반하여 인식하였다.

필기 숫자 인식의 두 번째 부류는 분류 및 인식의 방법을 개선하는 것이다. 즉, 특징 추출은 단순한 방법을 사용하거나 기존의 방법을 조금 개선시켜 사용하고 분류 및 인식을 위한 새로운 알고리즘을 제안한다. 분류 및 인식을 위한 알고리즘에는 학습 특성이 존재하는 신경회로망이 주로 이용되는데 그 중에서도 자율분류 특성이 있는 SOM(Self-Organizing Map)이 널리 사용된다. N. V. Subba는 필기 숫자 인식을 위하여 MSOM(Modified Self-Organizing Map)과 LVQ(Learning Vector Quantization)을 조합한 3차원 신경회로망(Three-dimensional neural network)을 제안하였다<sup>[4]</sup>. 제안한 알고리즘은 SOM이나 MSOM을 단독으로 적용할 때보다 학습시간, 인식률, reject율이 모두 개선되었고, SOM과 LVQ를 조합한 알고리즘과 비교할 때는 동일한 인식률에서 학습시간과 reject율이 줄어들 수 있었다. 그리고 SOM에 퍼지 알고리즘을 접목한 FRFOM(Fuzzy Region Self-Organizing Map)<sup>[5]</sup>과 SOM의 구조를 개선한 DSNN(Doubly Self-Orga-

nizing Neural Network)<sup>[6]</sup> 등에서도 기존의 방법보다 시간적인 면에서나 성능적인 면에서 모두 개선되어 좋은 특성을 보였다.

CME(Combination of Multiple Experts)에 의한 방법은 여러 가지 다른 알고리즘을 조합해서 인식 성능을 향상시키는 방법으로, 개개의 알고리즘은 앞서 설명한 2개의 범주에 속한다고 할 수 있다. 투표(voting), Bayesian, Dempster-Shafer에 의한 방법이 주로 이용되는데, 투표에 의한 방법은 여러 가지의 알고리즘을 이용해 인식하고 최종 결과는 투표에 의하여 결정하는 비교적 단순한 방법인데 비하여 Bayesian 방법이나 Dempster-Shafer 방법에서는 분류를 위한 식이나 조건 등을 추가하여 구조가 더 복잡해진 반면 인식의 성능이 더욱 개선되었다.<sup>[7-11]</sup>

본 논문에서는 필기 숫자를 형태에 따라서 16개의 그룹으로 분리하고 각 그룹에 대하여 독자적인 특징추출과 인식 알고리즘을 적용하는 방법을 제안한다. 기존의 특징추출 알고리즘에 비하여 매우 단순하면서도 숫자 형태에 따른 효과적인 특징추출을 할 수 있으며, 형태에 따른 그룹화에 의해서 각 그룹별로 독자적인 인식 알고리즘을 적용함으로써 인식을 위한 학습시간이 대폭 단축되는 것이 제안한 알고리즘의 장점이다. 필기 숫자의 형태 파악을 위해서는 숫자에 외접하는 많은 직선을 발생한 다음 외접선이 관통하지 않는 내부영역을 그 숫자의 형태로 파악하였다. 각 숫자는 그의 형태에 따라 미리 설정된 16개의 템플레이트(template) 중 유사한 하나로 분류되며 각 그룹은 각각의 특징추출 방법과 인식 알고리즘을 가지고 있다. 특징추출에서는 각 그룹의 해당 템플레이트 모양을 참고하여 숫자 정보가 존재할 가능성이 적은 부분에는 특징점의 수를 줄여서 컴퓨터 메모리를 감소시킴으로서 전체 인식 시스템의 처리 속도를 향상시킨다. 인식 알고리즘으로는 SOM을 개선한 FSOM을 적용한다. FSOM은 SOM의 학습반경내의 노드(node) 갱신을 위한 학습률 설정에 퍼지 알고리즘<sup>[12]</sup>을 적용하여 승리노드와의 유사성에 의해 학습률을 달리 설정함으로써 SOM의 특성을 개선시킬 수 있었다. 그리고 기존의 알고리즘의 경우 전체 데이터에 대해 하나의 인식 알고리즘을 적용시키는데 비하여 제안한 알고리즘에서는 16개의 그룹에 대하여 작은 규모의 인식 알고리즘을 각각 사용하여 분산 학습함으로써 계산시간을 대폭 단축할 수 있고, 각 그룹별로 독자적인 학습이 가능한 장점이 있다. 제안한 알고리즘의 성능을 평가하기 위해서 Concordia 대학에서 제

공하는 4000개의 필기 숫자 데이터를 사용하였다. 2000개는 학습을 위한 데이터이고 나머지 2000개는 성능 평가를 위한 것이다. 그림 1은 본 연구에서 제안하는 필기 숫자 인식 시스템의 구조를 나타낸다.

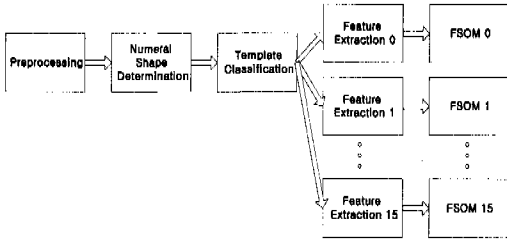


그림 1. 필기 숫자 인식 시스템

II. 전처리

필기체 문자는 왜곡이 심하고 불규칙적 형태를 가지고 있어 이를 분류하고 인식하기 위해서 전처리 과정은 필수적이다. 본 연구의 전처리 필터는 입력 필기체 숫자가 갖는 잡음을 제거하고 여러 가지 특성을 정규화하여 전체 시스템의 성능을 향상시킨다. 그림 2는 전처리 필터의 구성을 나타낸다.

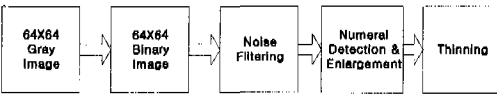


그림 2. 전처리 필터

1. 이진화

입력은 256 그레이 레벨(gray level)의 영상으로 식 (1)에서  $f(n, m)$ 로 표시하였다. 여기서  $N_r$ 과  $N_c$ 는 입력 영상에 대한 가로와 세로 화소의 수를 나타내는데, 본 실험에서는  $64 \times 64$  크기의 영상을 사용하였으므로 각각 64의 값을 가진다.

$$f(n, m) = [0, 255]$$

$$n = 1, 2, \dots, N_r, \quad m = 1, 2, \dots, N_c \quad (1)$$

데이터 양의 축소와 시스템의 고속화를 위해서 식 (2)를 사용하여 그레이 레벨 영상을 이진화한다.  $f_{Th}(n, m)$ 는 이진화된 영상을 나타내며  $Th$ 는 임계치(threshold value)이다. 해당 화소  $f(n, m)$ 의 그레이 레벨이 임계치 이하이면 1로, 임계치를 초과하면 0으로 그 값을 할당한다. 임계치는 보통 128을

사용한다.

$$f_{Th}(n, m) = \begin{cases} 1 & f(n, m) \leq Th \\ 0 & f(n, m) > Th \end{cases} \quad (2)$$

2. 잡음 제거와 윤곽선 조정

입력 영상의 잡음 제거에는 소영역 고립점을 검출하여 제거하는 소영역 제거 필터를 적용하였다. 이 알고리즘은 고립점의 최대 크기를 설정하면 그보다 작은 영역을 모양에 관계없이 제거하는데, 빠르고 간단한 대신 숫자와 접촉해 있는 잡음은 검출하지 못하는 단점이 있다. 제거되지 못한 잡음은 morphological filters를 이용한 숫자 윤곽선 교정 부분에서 수정된다.<sup>(13)</sup>

3. 숫자 영역의 검출과 확대

본 논문에서는 입력 영상에서 숫자가 쓰여진 부분을 검출하기 위해서 사각형 검출법(rectangular detection algorithm)을 제안한다. 사각형 검출법에서는 전체 입력 영상 중에서 숫자 부분만 정보로 검출하여 사용하므로 입력 숫자에 대한 위치 정보를 일반화 할 수 있다. 사각형 검출법은 입력 이진영상에 대하여 1의 값을 갖는 부분, 즉 숫자가 쓰여진 부분 중 상·하·좌·우 각 방향의 가장자리 부분을 계산한 뒤, 계산한 가장자리 4개의 값을 포함하는 사각형을 설정하여 그 사각형 내부를 숫자로 검출하는 알고리즘이다. 이진 영상인  $f_{Th}(n, m)$ 에서 숫자 영역을 결정하는 함수를  $Rect[\cdot]$ 라 하면 숫자 영역  $r(n, m)$ 은 식 (3)과 같이 나타낼 수 있다. 여기서  $n_{min}, n_{max}, m_{min}$  그리고  $m_{max}$ 는  $f_{Th}(n, m)$ 에서 1의 값을 가지는 화소 중  $n$ 과  $m$ 의 최대·최소 위치를 나타낸다.

$$r(n, m) = Rect[f_{Th}(n, m)]$$

$$= \begin{cases} \text{if } (n_{min} \leq n \leq n_{max} \text{ and } m_{min} \leq m \leq m_{max}) \\ r(n - n_{min}, m - m_{min}) \\ = f_{Th}(n, m) \\ \text{else} \\ r(n, m) = 0 \end{cases} \quad (3)$$

$$n = 1, 2, \dots, N_r, \quad m = 1, 2, \dots, N_c$$

본 알고리즘은 영상에 잡음이 있을 때에는 잡음까지 숫자 영역으로 간주하는 심각한 문제를 가지고 있어 잡음제거가 필수 요건이다. 그림 3은 잡음

이 제거된 영상에서 숫자 부분을 검출하는 사각형 검출법을 적용한 것으로, 회색으로 표시한 영역이 숫자 부분으로 결정된 영역이다.



그림 3. 숫자 영역의 검출의 예

식 (4)는 검출된 숫자 영역을 일정한 크기로 확대하는 함수로, 수식 중 (int)는 해당 실수값 이하의 최대 정수로 변환하는 연산자이다. 여기서 문제가 되는 점은 가로로 길게 쓴 1과 같이 가로와 세로의 비가 매우 큰 경우 위의 방법을 이용하면 숫자의 정보를 잃어버릴 가능성이 있다는 것이다. 따라서 위의 알고리즘을 적용하기 전에 식 (5)를 이용하여 가로와 세로의 비( $\xi$ )를 계산하여 2.5보다 큰 경우 세로 방향만 확대하여 숫자의 정보를 보존한다. 그림 4에서는  $\xi$ 의 값에 따른 숫자 영역 확대의 한 예를 보여준다. 수직에 가까운 숫자 1의 경우 상·하로만 확대되었음을 볼 수 있다.

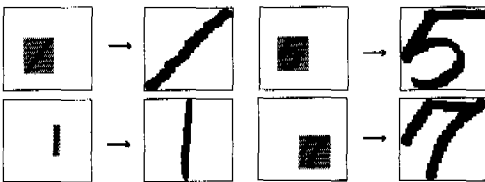


그림 4. 가로와 세로의 비율에 따른 숫자 영역 확대

$$e(n, m) = \text{Enlarge}[r(n, m)] \tag{4}$$

$$= \begin{cases} \eta \leq 2.5 \\ e(n, m) = r(x, y) \\ \text{else} \\ e(n, m) = r(z, y) \end{cases}$$

$$n = 1, 2, \dots, N_r, \quad m = 1, 2, \dots, N_c$$

여기서,

$$x = (\text{int})\left[\frac{n_{\max} - n_{\min}}{N_r} \cdot (n - 1)\right]$$

$$y = (\text{int})\left[\frac{m_{\max} - m_{\min}}{N_c} \cdot (m - 1)\right] \tag{5}$$

$$z = (\text{int})\left(\frac{n_{\max} - n_{\min}}{2}\right) + (n - 1)$$

$$\xi = \frac{m_{\max} - m_{\min}}{n_{\max} - n_{\min}}$$

#### 4. 세선화

세선화(thinning) 방법은 여러 가지 종류가 개발되어 있고 그의 특성도 여러 가지이다.<sup>[6]</sup> 여기서 주목할 점은 세선화 대상이 얼마의 굵기를 가지느냐에 따라 여러 가지 세선화 알고리즘 적용 결과가 달라진다는 것이다. 본 논문에서는 일반적인 세선화 알고리즘을 적용하기 전에 숫자의 굵기를 전체적으로 줄여주는 에지 검출 세선화 알고리즘(edge detection thinning algorithm)을 제안하여 세선화 특성을 개선한다. 제안한 알고리즘은 완전한 세선화는 할 수 없지만 전체적인 숫자의 굵기를 줄임으로서 기존의 세선화 알고리즘을 적용하였을 때 숫자의 굵기에 따른 오차를 줄일 수 있다. 방법은 원 영상에서 에지를 검출하고, 검출한 에지를 원 숫자 영상에서 제거함으로 전체적인 숫자의 굵기를 줄이는 것을 반복하는 것이다.

그림 5는 에지 검출 세선화의 방법과 1회의 에지 검출 세선화 알고리즘을 적용한 결과를 설명한다. 제안한 알고리즘을 3회 정도 반복한 다음 일반적인 세선화 알고리즘인 two-pass thinning을 적용하면 효과적인 세선화 결과를 얻을 수 있다. 에지를 검출하여 제거하는 과정을 반복하는 경우 원 영상에서 굵기가 얇은 곳의 연결 부분이 끊어질 우려가 있다. 이를 방지하기 위해서 교차점 보상(crossing compensation)을 도입하여 가로와 세로가 교차하는 한 화소 굵기의 점을 계산하여 세선화된 영상에 더함으로써 개선하였다.

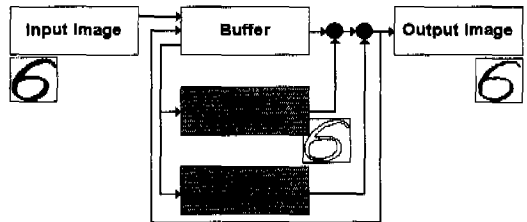


그림 5. 에지 검출 세선화 알고리즘

### III. 특징추출과 인식 알고리즘

#### 1. 입력 필기체 숫자의 형태 파악

모든 형태의 문자에는 일정한 형태가 있다. 비록 쓰는 사람의 글씨체나 글꼴에 따라서 다양한 형태

가 존재하지만 전체적인 모양을 보면 유사한 특성을 가지고 있다. 예를 들면 숫자 1의 경우는 면적이 좁은 긴 직사각형 형태나 대각선 형태를, 숫자 0의 경우는 면적이 넓은 사각형이나 원의 형태를 가진다. 본 연구에서는 이 형태를 추출하기 위해서 여러 가지 기울기의 직선을 발생하고, 그 중 숫자 부분과 접하는 접선을 이용하여 전체적인 형태를 결정한다. 그림 6(a)는 형태결정을 위한 직선 생성을 설명한다.  $L_1(n)$ 은 식 (6)과 같이 나타낼 수 있으며 나머지 직선의 방정식  $L_2(n), L_3(n), L_4(n)$ 도 그와 유사하게 표현할 수 있다.

$$L_1(n) = \Theta(n, m) n + h_1 \quad (6)$$

여기서, 기울기  $\Theta(n, m)$ 는 숫자부분인  $e(n, m)$ 과 접하지 않는 최대의 절대값을 나타낸다.  $h_1$ 을  $m_1$ 부터  $m_2$ 까지 감소시키며 접선을 발생하고, 숫자를 포함하며 발생된 접선들로 둘러싸이는 영역을  $R_1(n, m)$ 라고 한다. 다음으로 그림 6(b)처럼, 상하 좌우의 경계를 설정하기 위한 네 개의 직선을 설정하여 각 방향의 경계  $n_1', n_2', m_1', m_2'$ 을 각각 계산한다. 이 점선들은 각 방향에 대하여 숫자영역  $e(n, m)$ 에 접하지 않는 각 방향의 최대 또는 최소 경계를 나타낸다. 이 네 개의 점선에 의하여 형성되는 영역을  $R_2(n, m)$ 라 한다. 전체적인 숫자의 형태인  $R_3(n, m)$ 은 식 (7)처럼  $R_1(n, m)$ 와  $R_2(n, m)$ 의 교집합으로 표현된다. 그림 7에서 그의 예를 보인다.

$$R_3(n, m) = R_1(n, m) \cap R_2(n, m) \quad (7)$$

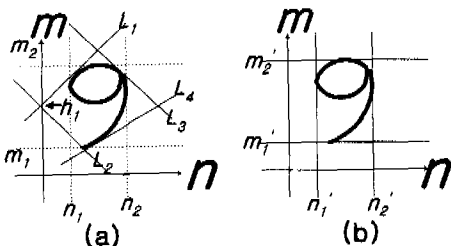


그림 6. 숫자 9의 형태 파악

$$R_3(n, m) = R_1(n, m) \cap R_2(n, m)$$

그림 7. 숫자 9의 형태 결정의 예

## 2. 템플레이트 매칭에 의한 패턴 분류

본 연구에서는 Concordia 대학의 2000개의 학습 데이터에 대하여 형태 파악 알고리즘을 적용하고 전체적인 모양을 설정한 다음, 그림 8에 나타난 16개의 적절한 템플레이트를 설정하였다. 숫자 입력은 16개의 패턴 중 하나로 분류되어서 특징이 추출되고 각각의 신경회로망에 학습된다. 유사한 패턴으로 분류하는 방법으로 유클리드 거리 계산법(Euclidean distance method)을 이용하였다<sup>[13]</sup>. 유클리드 거리의 계산식은 식 (8)과 같다. 여기서  $P_i(n, m)$ 은 입력 영상에서  $n, m$  위치의 화소값이며,  $P_j(n, m)$ 은 템플레이트 영상에서  $n, m$  위치의 화소 값으로 0 또는 1의 값을 가진다.

$$EUD = \sum_{n=1}^N \sum_{m=1}^M (P_i(n, m) - P_j(n, m))^2 \quad (8)$$

## 3. 특징 추출과 FSOM의 학습

### 가. 특징 추출

입력 필기 숫자는 16개의 템플레이트로 분류된다. 각 템플레이트 영상을 살펴보면 숫자의 정보를 포함할 가능성이 높은 부분과 낮은 부분으로 대별될 수 있다. 그림 8은 필기 숫자 형태에 적용적인 특징 추출을 위해 설정한 템플레이트를 나타내는데, 좌측 상단의 템플레이트 0번과 유사한 모양의 필기 숫자인 경우는 중앙의 좁은 영역에 숫자 정보가 존재할 가능성이 높으며 나머지 부분은 가능성이 매우 낮다. 따라서 특징 추출을 위한 메시(mesh)를 설정할 때 중요한 부분은 조밀하게 그리고 정보가 존재할 가능성이 희박한 부분은 간격을 넓게함으로써 특징 점의 수를 줄이면서 그 정보를 효과적으로 표현할 수 있다.

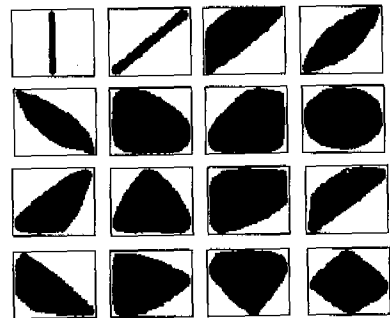


그림 8. 실험에 사용한 16개의 템플레이트

그림 9에서는 각 템플레이트의 특징 추출을 위한 메시지를 나타내는데, 특징점의 개수가 24에서 56까지 존재한다. 템플레이트 16개를 기준으로 할 때, 기존의 64개로 일정한 특징점을 추출하는 방식에 비하여 26.87% 정도의 메모리가 절약된다.

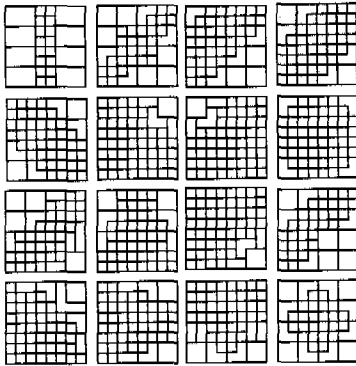


그림 9. 특징 추출을 위한 메시지들

나. FSOM의 학습

16개의 각 템플레이트에 대하여 각각 SOM을 적용한다. 본 논문에서는 학습률을 적용할 때 퍼지 알고리즘을 이용하여서 학습의 효율을 높인 FSOM을 사용한다. 각 FSOM에 대한 노드의 개수는 식 (9)로 설정한다.

$$Node = [(int)\sqrt{0.6 \times N_{template}}]^2 \quad (9)$$

여기서, Node는 각 FSOM의 노드 개수이며  $N_{template}$ 는 각 템플레이트에 분류된 입력 필기숫자의 개수이다. 그리고 (int)는 계산 결과에 대한 최대의 정수 값을 나타낸다. 제안한 FSOM이 일반적인 SOM 알고리즘과 다른 점은 승리노드 중심으로 갱신 반경내의 노드를 갱신할 때 퍼지 알고리즘을 이용하여 승리노드와의 오차에 따라 다른 학습률을 부여함으로써 SOM의 학습성능을 개선한 것이다. FSOM의 알고리즘은 다음과 같다.

[Step 0] 초기화

- a. 연결 강도  $v_{ij}(1 \leq i \leq N_f, 1 \leq j \leq N_n)$ 를 0과 1사이의 랜덤값으로 초기화한다. 여기서  $N_f$ 는 특징추출의 개수를,  $N_n$ 은 노드의 개수를 나타낸다.
- b. 초기 학습률과 갱신반경을 설정한다

[Step 1] 정지 조건이 만족될 때까지 Step 2에서

Step 5를 반복

본 논문에서 정지조건은 초기에 설정한 반복회수 (iteration)이며, 반복회수가 증가함에 따라 학습률과 노드의 갱신반경을 감소시킨다.

[Step 2] 유클리드 거리 계산

식 (10)에 의해서 입력패턴  $x_i(t)$ 와 연결강도의 유클리드 거리를 계산한다.

$$EUD_j = \sum_{i=1}^{N_f} (x_i(t) - v_{ij}(t))^2 \quad (10)$$

[Step 3] 승리 노드(winner node) 선정

최소 유클리드 거리를 갖는 노드  $j^*$ 와 유클리드 거리를 다음 식 (11)로 구한다.

$$EUD_{j^*} = \min_j [EUD_j] \quad (11)$$

[Step 4] 퍼지 학습률(fuzzy learning rate :  $u_{ij}$ ) 계산

식 (12)를 이용하여 일반적인 학습률에 퍼지 계수를 곱하여 퍼지 학습률을 구한다.

$$u_{ij}(t) = f(t) \times r(t) \quad (12)$$

$$f(t) = \left( \frac{EUD_{j^*}}{EUD_j} \right)^2$$

여기서,  $r(t)$ 는 학습률을 나타내며,  $f(t)$ 는 퍼지계수이다.

[Step 5] 연결 강도의 갱신

승리 노드를 중심으로 갱신반경 내의 노드 연결강도를 식 (13)을 이용하여 갱신한다.

$$v_{ij}(t+1) = v_{ij}(t) + u_{ij}(t)[x_i(t) - v_{ij}(t)] \quad (13)$$

IV. 실험 및 고찰

제안된 알고리즘의 성능을 평가하기 위하여 Concordia 대학의 필기 숫자 4000개를 이용하였다. 2000개는 학습을 위한 데이터이며 나머지 2000개는 테스트를 위한 데이터이다.

1. 전처리 필터

본 실험에서 사용된 입력 필기체 숫자 영상의 크기는 64×64이며, 입력 영상을 이진화 할 때 150의

임계값을 사용하였다. 전처리 과정은 먼저 이진 영상의 잡음을 제거하고 숫자 부분을 검출한 뒤 숫자 부분을 일정한 크기인 64×64로 확대하였다. 다음으로 숫자 윤곽선 및 공백을 보정하기 위해서 dilation filter와 erosion filter를 각각 2회씩 교대로 적용하였다. 마지막으로 숫자 굵기에 대한 정규화를 위해서 에지 검출 세션화 알고리즘과 two-pass thinning 알고리즘을 사용하여 숫자의 굵기를 한 화소의 일정한 굵기로 재조정하였다<sup>[14]</sup>. 그림 10은 두 숫자의 전처리 결과이다. 형태는 유사하지만 숫자의 크기, 굵기, 위치, 진하기가 다르고 잡음이 섞인 두 숫자를 전처리하여 유사한 패턴으로 묶을 수 있음을 볼 수 있다.

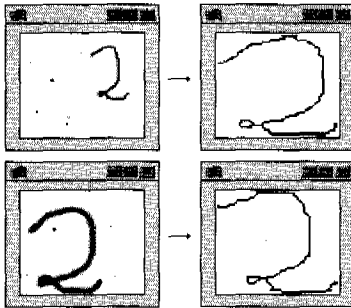


그림 10. 입력 숫자(좌)와 전처리 결과(우)

그림 11는 입력 숫자 6에 대하여 잡음을 제거하고 morphological filters에 의하여 윤곽선 및 공백을 보정한 이후의 전처리 과정을 설명하는 그림이다. 먼저 입력 영상에 대하여 숫자 부분을 검출한 뒤 일정한 크기로 확대하고 확대한 숫자에 대하여 에지 검출 세션화 알고리즘을 3회 적용한 다음 two-pass thinning을 적용한다. 이상의 작업들에 의하여 입력 숫자 6은 일정 크기와 굵기의 형태로 변환된다.

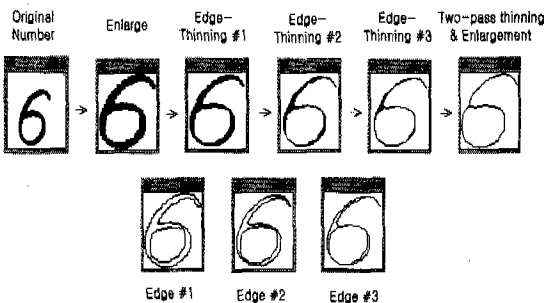


그림 11. 숫자 6 영상에 대한 전처리 과정

그림 12는 숫자 0에서 9까지 숫자에 대하여 잡음 제거, 윤곽선 보정, 일정 크기로 확대, 세션화 과정을 나타낸 그림으로, 일반적인 two-pass thinning 알고리즘에 대하여 에지 검출 세션화를 적용하였을 때와 적용하지 않았을 때 세션화 결과를 보여준다. 일반적인 two-pass thinning 알고리즘 결과를 보면 숫자의 가장자리 부분에 필요치 않은 잡음이 생겨서 최종 전처리 결과의 숫자 모양이 좋지 않고 숫자 부분의 크기가 일정하지 않게 되므로 전체 시스템의 효율을 감소시키게 된다. 가장 아랫부분은 two-pass thinning 알고리즘 적용하기 전에 에지 검출 세션화 알고리즘을 3회 적용한 결과로 숫자 가장자리의 잡음이 제거되어 숫자의 모양과 크기가 일정하게 정규화됨을 볼 수 있다.

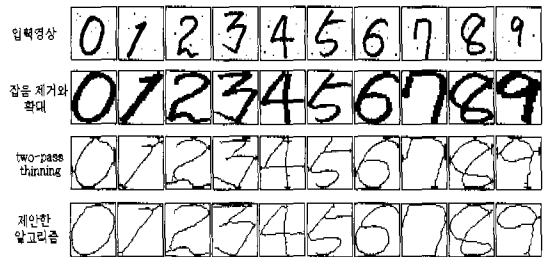


그림 12. 전처리 결과 비교

### 2. 형태 결정과 템플레이트 분류

그림 13은 학습 숫자의 일부에 대하여 확대, 세션화, 형태 결정, 분류에 의한 특징 추출 결과를 보여준다. 0의 경우는 템플레이트 3으로, 1의 경우는 템플레이트 1로, 2의 경우는 템플레이트 5로, 3의 경우는 템플레이트 4로, 그리고 6의 경우는 템플레이트 15로 분류되었다. 따라서 각각의 메시지를 사용하여 특징값을 추출한다. 그림 13에서 (a)는 원 입력영상, (b)는 확대된 결과, (c)는 세션화 결과, (d)는 형태 결정 결과, (e)는 각각의 메시에 의한 특징 추출을 나타낸다. 특징 추출에서 각각 숫자의 정보가 있는 부분은 세밀하게 추출됨을 볼 수 있다.

### 3. FSOM의 성능실험

입력 필기 숫자 데이터 중 1000개의 학습 데이터와 테스트 데이터에 대하여 15×15 크기의 SOM 알고리즘과 FSOM 알고리즘을 적용하여 그 결과를 비교하였다. 반복적인 실험에 의하여 가장 적절한 학습률인 0.4를 사용하였으며 반복회수는 50000회로 설정하였다. 그림 14는 학습이 진행됨에 따라 퍼

지 학습률이 변해가는 과정을 나타낸 그림이다. 반복회수 4000번 이후부터는 일반적인 SOM과 동일한 방법으로 갱신된다. (a)는 반복회수가 1일 때이고 (b)는 100, (c)는 500, (d)는 4000일 때이다. 반복회수가 늘어날수록 승리노드의 주변의 퍼지 학습률이 줄어들어 볼 수 있었다. 일반적인 SOM의 경우 50000회 반복 회수에서 학습률 91.0% 인식을 90.3% 이었고, FSOM의 경우 20000회 반복 회수에서 학습률 99.3% 인식을 97.9%을 나타내었다.

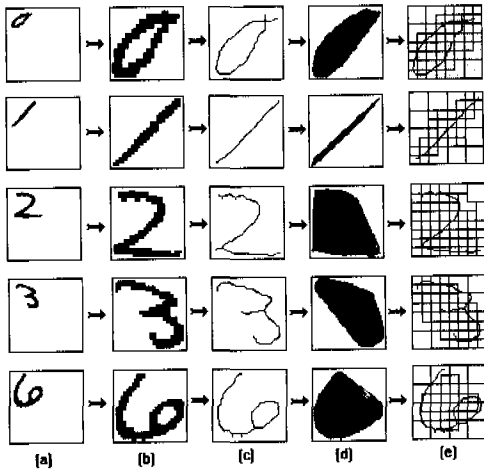


그림 13. 형태 결정과 특징 추출

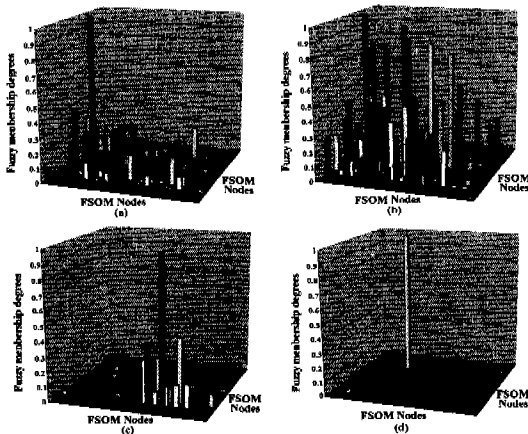


그림 14. 학습과정의 퍼지 학습률 변화

그림 15는 SOM과 FSOM의 성능을 비교한 그림이며, 16은 일반적인 SOM과 FSOM의 오차의 수렴에 대한 비교이다. 일반적인 SOM에 비하여 FSOM에서 오차 수렴 속도가 빠르며 수렴하는 값도 작아서 높은 인식률을 가짐을 볼 수 있었다.

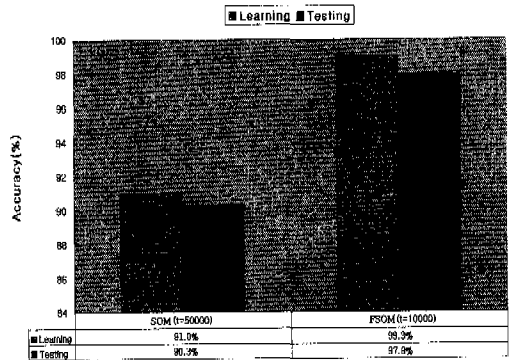


그림 15. SOM과 FSOM의 성능비교

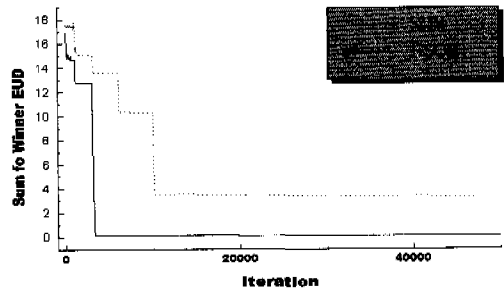


그림 16. SOM과 FSOM의 오차 수렴특성 비교

#### 4. 각 그룹별 인식률

전체 필기 숫자 데이터의 분류 결과와 각각 FSOM의 크기, 그리고 각 그룹의 인식률을 표 1에 나타내었다. 전체적으로 고른 인식률을 보였으나 그룹 13은 저조한 인식률은 나타났다. 본 실험에서 설정한 16개의 템플레이트 중에서 그룹 13으로 분류하는 템플레이트가 적당하지 않아서 학습시 전체 2000개의 데이터 중 24개의 데이터만 그룹 13으로 분류되었다. 그룹 13은 75%의 낮은 인식률을 보이지만 전체 데이터의 2% 미만의 개수이므로 전체 인식률에는 커다란 영향을 미치지 못한다. 반면 그룹 0, 2, 3, 4, 6, 8, 10, 11, 12, 14, 15는 100%의 인식률은 보였다.

#### 5. 기존의 인식 알고리즘과 비교

표 2에는 기존의 알고리즘의 성능과 비교하였다. 여기서 신뢰도(reliability)는 식 (14)에 의하여 계산되는데 이것은 인식률(recognition ratio)에 reject율(reject ratio)을 고려한 것이다. 제안한 알고리즘의 평균인식률은 바르게 인식된 테스트 데이터 값을 전체 테스트 데이터 개수로 나눈 값을 백분율 취하여 계산하였다. 제안한 알고리즘에 의하여 특징점의



수를 줄이고 학습시간을 단축시키면서도 다른 알고리즘과 유사한 인식성능을 얻을 수 있었다.

$$Reliability = \frac{Recog.}{Recog. + Error} \times 100 [\%] \quad (14)$$

표 1. 각 그룹에 대한 정보

Group	각 숫자의 개수									FSOM 노드개수	오인식 개수	인식률 [%]	
	0	1	2	3	4	5	6	7	8				9
0 학습	0	16	0	0	0	0	0	0	1	0	9	0	100.0
0 테스트	0	45	0	0	0	0	0	0	0	0	0	0	0
1 학습	0	99	0	0	2	0	0	1	2	0	49	7	95.55
1 테스트	0	77	0	0	8	1	1	0	1	2	0	0	0
2 학습	0	0	0	0	13	8	0	49	1	27	49	0	100.0
2 테스트	0	0	0	0	14	7	0	68	3	24	0	0	0
3 학습	13	76	3	18	50	39	18	7	66	9	169	0	100.0
3 테스트	20	73	3	19	42	54	8	7	87	14	0	0	0
4 학습	0	4	2	2	3	0	2	2	2	1	9	1	100.0
4 테스트	0	2	0	1	3	0	1	1	1	2	0	0	0
5 학습	6	1	43	27	0	0	15	0	2	0	49	8	98.86
5 테스트	7	0	35	21	0	6	22	0	3	0	0	0	0
6 학습	3	0	34	59	1	23	35	0	9	0	121	3	100.0
6 테스트	43	0	35	48	2	27	51	0	5	1	0	0	0
7 학습	53	0	11	33	6	21	10	0	19	1	81	5	99.38
7 테스트	66	0	12	33	7	16	11	1	15	0	0	0	0
8 학습	29	0	10	32	5	36	83	0	11	0	121	4	100.0
8 테스트	29	0	32	39	15	27	73	0	12	0	0	0	0
9 학습	1	0	46	4	0	31	18	1	0	0	36	3	98.94
9 테스트	3	0	57	3	1	4	26	0	0	0	0	0	0
10 학습	35	0	7	19	24	20	2	26	45	34	121	0	100.0
10 테스트	36	0	6	27	10	19	1	11	28	24	0	0	0
11 학습	0	3	0	2	35	45	0	41	28	78	144	4	100.0
11 테스트	0	3	0	0	52	36	0	46	27	79	0	0	0
12 학습	0	1	5	0	0	0	1	46	1	0	4	0	100.0
12 테스트	0	0	7	0	0	0	2	0	0	0	0	0	0
13 학습	0	0	15	0	2	0	2	3	2	0	9	5	75.00
13 테스트	1	0	10	4	2	0	0	0	0	3	0	0	0
14 학습	0	0	2	3	26	3	1	70	8	48	81	2	100.0
14 테스트	0	0	2	5	20	1	0	29	18	53	0	0	0
15 학습	3	0	2	1	13	2	3	0	0	1	16	1	100.0
15 테스트	0	0	2	0	24	2	5	0	0	0	0	0	0

표 2. 기존의 알고리즘과 비교평가

Methods	Recog. [%]	Error [%]	Reject [%]	Reliability [%]
Legault <sup>(15)</sup>	93.90	1.60	4.50	98.32
Krzyzak <sup>(16)</sup>	94.85	5.15	0.00	94.85
Suen <sup>(17)</sup>	93.05	0.00	6.95	100.0
Mai <sup>(18)</sup>	92.95	2.15	4.90	97.74
제안한 방법	97.85	2.15	0.00	97.85

### V. 결론

본 연구에서는 전처리 필터, 필기 숫자 형태를 이용한 특징추출과 분류학습, 그리고 FSOM을 이용하여 필기숫자를 인식하는 새로운 알고리즘을 제안하였다. 외접선에 의한 필기 숫자의 형태 파악은 숫자의 형태에 따른 적응적인 특징추출과 데이터를 몇 개의 그룹으로 분할하는 것을 가능하게 하였으며 FSOM을 분할하여 학습할 수 있어 학습의 효율을 높이며 학습시간을 단축할 수 있었다. 여기서 학습 시간의 의미는 1회 학습시간이 아니라 원하는 학습률에 도달할 때까지의 반복에 의한 학습 시간의 총합을 말한다. 즉, 제안한 알고리즘에서는 성능이 저조한 그룹만 학습 파라미터를 달리하여 재학습하고, 기존의 잘 학습된 그룹의 학습결과는 그대로 사용

할 수 있다. 따라서 본 알고리즘은 학습데이터를 추가될 경우 추가되는 데이터가 분류되는 그룹만 학습하면 되므로 학습 데이터가 자주 갱신되는 경우 매우 효과적인 알고리즘이라 할 수 있다. 그리고 분류의 기준을 숫자의 특성을 더 잘 표현할 수 있는 다른 대상으로 수정하면 성능을 더 높일 수 있으며, 그룹별 학습방법은 학습이 필요한 신경회로망의 여러 분야에 적용이 가능하다.

실험결과를 살펴보면 각 그룹의 FSOM에 대하여 학습할 때 고려되지 않은 숫자가 전체 인식률을 나쁘게 하는 요인으로 작용하였고, 특히 그룹 13의 경우는 다른 그룹에 비하여 인식률이 상대적으로 저조하였다. 더 좋은 성능을 위해서는 그룹화를 위한 기준 템플레이트를 임의로 설정하지 않고 전체 데이터의 형태에 대한 통계적 특성을 이용하여 각 그룹에 데이터가 고르게 분포하게 하여 전체적인 인식률을 높이는 방법이 바람직할 것이다.

### 참고 문헌

- [1] Yoshihiko Hamamoto, Shunji Uchimura, Masanori Watanabe, Tetsuya Yasuda and Shingo Tomita, "Recognition of Handwritten Numerals Using Gabor Features," Proceedings of ICPR'96, pp. 250-253, 1996.
- [2] Dahai Cheng and Hong Yan, "Recognition of Handwritten Numeral Base on Contour Information," Pattern Recognition, Vol. 3, No. 1, pp. 235-255, 1998.
- [3] Sven Behnke and Marcus Pfister, "Recognition of Handwritten Digits Using Structural Information," Proceedings of ICNN'97, Vol. 4, pp. 1391-1396, 1997.
- [4] N. V. Subba Reddy and P. Nagabhushan, "A Three-dimensional Neural Network Model for Unconstrained Handwritten Numeral Recognition," Pattern Recognition, Vol. 31, No. 5, pp. 511-516, 1998.
- [5] Zheru Chi, Jing Wu and Hong Yan, "Handwritten Numeral Recognition Using Self-Organizing Maps and Fuzzy Rules," Pattern Recognition, Vol. 28, No. 1, pp. 59-66, 1995.
- [6] Sung-Bae Cho, "Recognition of Unconstrained Handwritten Numerals by Self-Organizing Neural Network," Proceedings of ICPR'96,

pp. 426-430, 1996.

[7] Y. S. Huang and C. Y. Suen, "A Methode Combining Multiple Experts for the Recognition of Unconstrained Handwritten Numerals," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 17, No. 1, pp. 90-94, January 1995.

[8] L. Xu, A. Krzyzak, and C. Y. Suen, "Method of Combining Multiple Classifiers and Their Application to Handwritten Numeral Recognition," IEEE Transaction on Systems, Man and Cybernetics, Vol. SMC-22, No. 3, pp. 418-435 1992.

[9] T. K. Suen, R. Legault, C. Nadal, M. Cheriet, and L. Lam, "Building a New Generation of Handwriting Recognition Systems," Pattern Recognition Letters, Vol. 14, No.4, pp. 305-315, 1993.

[10] J. Pearl, Probabilistic Reasoning in Intelligent System: Network of Plausible Inference, Morgan Kaufmann Publishers Inc., San Mateo, California, 1998.

[11] M. D. McLeish, P. Yao, and T. Stirtzinger, "A Study on the Use of Belief Functions for Medical Expert Systems," Journal of Applied Statistics, Vol. 18, No.1, pp155-174, 1991.

[12] Chin-Teng Lin and C. S. George Lee, Neural Fuzzy Systems, Prentice-Hall Inc., ISBN 0-13-261413-8, 1996.

[13] R. C. Gonzalez and R. E. Woods, Digital image processing, Addison-Wesley publishing Co., ISBN 0-201-11026-1, 1993.

[14] Maher A, Sid-Ahmed, "Image processing theory, algorithms, and architectures," Mc Graw-Hill Inc., ISBN 0-07-057240-2, 1993.

[15] R. Legault and C. Y. Suen, "Contour Tracking and Parametric Approximations for the Digitized Patterns," Computer Vision and Shape Recognition : Singapore, pp. 225-240, 1989.

[16] A. Krzyzak, W. Dai and C. Y. Suen, "Unconstrained Handwritten Character Classification using Modified Backpropagation Model," In Proc. 1st Int. Workshop on Frontiers in Handwriting Recognition, pp. 155-166, Montreal,

Canada, 1990.

[17] C. Y. Seun, C. Nadal, R. Legault, T. A. Mai and L. Lam, "Computer Recognition of Unconstrained Handwritten Numerals," Proceeding of the IEEE, Vol. 80, No. 7, pp. 1162-1180, 1992.

[18] T. Mai and C. Y. Suen, "A General Knowledge-based System for the Recognition of Unconstrained Handwritten Numerals," IEEE Trans. on Systems, Man and Cybernetics, Vol. 20, NO. 4, pp.835-848, 1990.

서 석 배(Seok-Bae Seo)

정회원

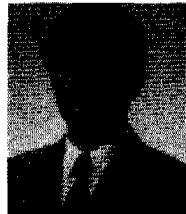


1997년 2월 : 부경대학교 제어계  
측공학과 졸업(학사)  
1999년 2월 : 동아대학교 전자공  
학과 졸업(석사)  
1999년~현재 : 동아대학교 컴퓨  
터공학과 박사과정

<주관심 분야> 신경망, 영상압축, 패턴인식 등

강 대 성(Dae-Seong Kang)

정회원



1984년 2월 : 경북대학교 전자공  
학과(학사)  
1991년 8월 : Texas A&M  
Univ., Electrical Eng.  
(석사)

1994년 5월 : Texas A&M Univ., Electrical Eng.  
(박사)

1984년 3월~1989년 6월 : 국방과학연구소 연구원  
1994년 7월~1995년 8월 : 한국전자통신연구소 선임  
연구원

1995년 9월~현재 : 동아대학교 전기전자컴퓨터공학  
부 조교수

<주관심 분야> 영상처리, 패턴인식, 영상코딩, 통신  
시스템 등

김 대 진(Daijin Kim)

정회원

현재 포항공과대학교 컴퓨터공학과 부교수