

# Madaline의 잡음에 대한 성능분석

## (Sensitivity Analysis of Madaline)

吳相勳\*, 李永稷\*

(Sang Hoon Oh and Young Jik Lee)

### 要約

학습이 완료된 신경회로망의 입력잡음에 대한 성능변화는 그 신경회로망의 학습정도에 대한 기준이 된다. 또, h/w로 구현 시에는 가중치가 학습에 의해 얻은 값과 달라지는 것을 반드시 고려해야 한다. 이것은 가중치에 잡음이 섞인 것으로 해석할 수 있다. 이 논문에서는, 학습이 완료된 다층 신경회로망 Madaline에서 가중치가 정규분포 잡음에 의해 변하는 경우와 입력패턴의 일부분이 달라지는 경우에, 출력뉴런의 민감도를 유도하였다. 유도한 수식은 가중치, 입력패턴, 그리고 입력패턴이 달라질 확률 혹은 가중치에 섞인 잡음의 분산에 대한 함수이다. 수식유도 결과는 필기체 숫자 인식 Madaline의 시뮬레이션을 통하여 확인하였다.

### Abstract

Well-trained neural networks have low sensitivity to input errors. Also, the sensitivity to weight errors must be considered when implementing neural networks with hardware of limited precision. In this paper, we derive the sensitivity of the Madaline to weight perturbation or input errors in terms of the trained weights, the input pattern, and the variance of weight perturbation or the probability of input errors. The result is verified with a simulation of the Madaline recognizing handwritten digits.

### 1. 서론

신경회로망이 패턴인식 문제를 학습하였다면, 입력패턴의 일부분이 달라질 경우 출력뉴런 값의 변동이 그 신경회로망의 학습정도에 대한 기준이 된다.<sup>1</sup> 신

경회로망을 H/W로 구현 시에는 가중치 구현의 정확성과 정밀도에 의해서도 출력값이 영향을 받는다.<sup>[2, 3]</sup>

<sup>4)</sup> 신경회로망을 실제 문제에 응용할 경우, 이와 같이 입력패턴 혹은 가중치에 대한 출력뉴런의 민감도를 반드시 고려해야 한다. 이 부분에 대한 연구는 대부분이 시뮬레이션으로 특성을 알아보기나<sup>5)</sup>, 간략화된 가정 하에서 특성 계산이 이루어져 왔다.<sup>[1, 2, 3, 4]</sup> Choi 등은 미분 가능한 활성화 함수를 1차 함수로 근사시켜 다층 퍼셉트론의 민감도를 구하였으나<sup>[1]</sup>, 이 방법은 1차 근사의 한계를 지니고 있다. Xie와

\* 正會員, 韓國電子通信研究所 基礎技術研究部  
(Electronics and Telecommunication  
Research Institute, Research Dept.)  
接受日字 : 1993年 6月 17日

Jabri는 비선형 변환을 준선형 변환으로 근사시키고, 입력, 가중치, 그리고 가중치 적용 합이 균일분포라는 가정 하에 다층 퍼셉트론의 양자화 효과를 분석하였다.<sup>[2]</sup> Stevenson 등은 Madaline(many adaptive linear elements)을 대상으로 하여 입력 패턴의 균일분포 가정 하에 가중치 변동에 의한 민감도를 계산하였고, 가중치의 균일분포 가정 하에 입력 패턴의 변동에 의한 민감도를 계산하였다.<sup>[3]</sup> 그렇지만, 실제 문제에서 이러한 가정들은 거의 맞지 않다.<sup>[4]</sup> Orzechowski 등은 Madaline에서 학습에 의해 결정된 가중치가 정규분포 잡음에 의해 변할 경우의 민감도를 계산하였지만<sup>[4]</sup>, 중간층 뉴런들의 상태가 잡음에 의해 달라지는사건(event)을 고려해야 하므로 오랜 계산 시간이 소요된다.

본 연구팀은 시그모이드 활성화 함수를 가진 다층 퍼셉트론에서 입력패턴에 섞인 잡음에 의한 판별성능의 변화를 유도하였으며<sup>[6]</sup>, 가중치의 변동에 따른 판별성능 변화 역시 같은 방법으로 구할 수 있다. 이 논문에서는, -1, +1 상태 만을 가지는 뉴런으로 구성된 Madaline에서, 학습 후 입력패턴과 가중치의 변동에 의해 출력뉴런 값이 달라질 확률(민감도)을 수식으로 유도하였다. 유도한 수식은 가중치, 입력패턴, 그리고 입력패턴이 달라질 확률 혹은 가중치의 변동 정도를 나타내는 분산의 함수이다. 2장에서 수식유도를 하였으며, 3장에서 필기체 숫자 인식문제를 학습시킨 Madaline을 분석대상으로 하여 민감도의 유도결과를 시뮬레이션 결과와 비교하였다. 그리고, 4장에서 결론을 맺었다.

## II. Madaline의 잡음에 의한 출력 변화

### 1. Madaline의 구성

Madaline은 다층퍼셉트론과 마찬가지로 뉴런들의 층으로 구성되며, 각 뉴런들은 -1, +1 상태 만을 가진다.<sup>[7]</sup>  $l$ 층의 뉴런상태 벡터를  $s^l$  이라 하고,  $l$ 과  $l-1$ 층 사이의 가중치 행렬을  $W^l$  이라 하면

$$s_j^l = \begin{cases} +1 & \text{if } w_{j0}^l + \sum_i w_{ji}^l s_i^{l-1} \geq 0 \\ -1 & \text{otherwise} \end{cases} \quad (1)$$

와 같이  $s_j^l$ 이 결정된다. 즉, Madaline에서 각 층의 뉴런들은 가중치에 의해 아래 층의 뉴런들로 구성된 공간을 반으로 분할하며, 여러 층으로 구성된 뉴런들이 적절한 가중치들로 입력공간을 임의로 분할한다. Madaline은 입력공간을 임의의 2진수에 해당하는 분할 공간으로 나누는 기능을 수행한다.<sup>[7]</sup>

위에서 설명한 Madaline은 다층퍼셉트론의 시그

모이드 활성화 함수를 식 (1)과 같은 임계함수로 대체시킨 것이다. 따라서, Madaline과 같은 구조(뉴런 수, 연결모양)의 다층퍼셉트론을 학습알고리즘에 따라 학습시켜, 학습완료 후 뉴런활성화 함수를 시그모이드에서 임계함수로 변경하면 Madaline의 학습이 이루어졌다고 볼 수 있다.<sup>[8]</sup> Madaline의 학습알고리즘으로는 MR(Madaline Rule)도 있다. MR은 제일 아래층(입력층 제외)부터 윗층으로 차례로 선택된 뉴런의 가중치를 error가 줄어들도록 변형시키는데, 최소방해원칙에 따라 가중치 합이 가장 작은 뉴런부터 차례로 선택한다. 선택된 뉴런의 출력 반전이 error를 감소시키는 방향이면  $\alpha$ -LMS알고리즘에 따라 그 뉴런에 연결된 가중치를 변경한다. 만약, 선택된 뉴런의 출력 반전이 error를 감소시키지 않으면 그 뉴런에 연결된 가중치들은 변경시키지 않는다.

### 2. 잡음 섞인 가중치에 의한 출력의 변화

학습이 완료된 Madaline을 H/W로 구현할 때, 학습된 가중치는 구현방법에 따라 다른 값으로 변형된다. 예를 들어, 디지털 H/W로 Madaline이 구현된다면, 학습완료된 가중치는 정량화되어 저장되므로 균일분포 잡음이 섞인 것으로 볼 수 있다.<sup>[2]</sup> 아날로그 H/W로 Madaline을 구현한다면, 아날로그 곱셈기가 가중치와 뉴런 값의 곱셈기능을 수행할 것이다. 이때, 아날로그 곱셈기의 이득의 변동, 가중치를 저장 디바이스에 저장할 때의 열잡음 등 때문에 곱셈결과는 마치 Gaussian 잡음이 섞인 가중치와 뉴런 값을 곱한 형태일 것이다.<sup>[4, 9, 10, 11]</sup> 여기서는, 입력층, 하나의 중간층, 그리고 출력층으로 구성된 Madaline에서 학습된 가중치가 정규분포 잡음에 의해 변할 경우, 한 입력패턴에 대한 출력뉴런 값이 잡음에 의해 변할 확률을 구하였다

입력패턴을  $\mathbf{x} = [x_1, x_2, \dots, x_N]$  중간층의 상태벡터를  $\mathbf{h} = [h_1, h_2, \dots, h_H]$  라 하자. 여기서,  $N$ 은 입력 패턴 벡터의 차원 수이고,  $H$ 는 중간층 뉴런 수이다. 그리고, 중간층과 입력층 사이의 가중치 행렬을  $\mathbf{W}$ , 출력층과 중간층 사이의 가중치 행렬을  $\mathbf{V}$ 라 하자. 그러면,  $k$ 번째 중간층 뉴런의 상태는

$$h_j = \begin{cases} +1 & \text{if } a_j = w_{j0} + \sum_{i=1}^N w_{ji} x_i \geq 0 \\ -1 & \text{otherwise} \end{cases} \quad (2)$$

와 같이 결정된다. 여기서,  $a_j$ 는  $h_j$ 에 입력되는 가중치 적용 합이다.  $k$ 번째 출력뉴런에 입력되는 값은

$$b_k = v_{k0} + \sum_{j=1}^H v_{kj} h_j \quad (3)$$

이며,  $b_k$ 에 의해 출력뉴런의 상태가 결정된다.

학습에 의해 결정된 가중치 행렬  $\mathbf{W}$ 와  $\mathbf{V}$ 의 각 요소에  $N(0, \sigma_w)$  분포 잡음이 섞이면,  $a_i$ 는 정규분포 확률 변수들의 합이므로 정규분포이다. 그 평균과 분산은

$$\eta_{a_i} = w_{j_0} + \sum_{i=1}^N w_{ji} x_i \quad (4)$$

$$\sigma_{a_i}^2 = \sigma_w^2 \left( 1 + \sum_{i=1}^N x_i^2 \right) \quad (5)$$

이며,  $a_i$ 와  $a_j$  간의 상관계수는 0이다. 따라서, 식 (2)와 같이 변환 후 중간층 뉴런값의 평균과 분산은

$$E[h_j] = 1 - 2G\left(-\frac{\eta_{a_i}}{\sigma_{a_i}}\right) \quad (6)$$

$$\text{Var}[h_j] = 4G\left(-\frac{\eta_{a_i}}{\sigma_{a_i}}\right) \left[ 1 - G\left(-\frac{\eta_{a_i}}{\sigma_{a_i}}\right) \right] \quad (7)$$

이다. 여기서  $G(\cdot)$ 는 Gaussian 분포함수이다.  $b_k$ 는, 서로 독립이며 분산이 유한한 확률변수들의 합이므로,  $H$ 가 충분히 크면 ( $H > 8$ ) central limit theorem에 의해 정규분포이다.<sup>[12]</sup> 그 평균과 분산은

$$\eta_{b_k} = v_{k_0} + \sum_{j=1}^H v_{kj} E[h_j] \quad (8)$$

$$\sigma_{b_k}^2 = (1 + H)\sigma_w^2 + \sum_{j=1}^H v_{kj}^2 \text{Var}[h_j] \quad (9)$$

로 계산된다. 따라서,  $k$ 번째 출력뉴런의 상태가 학습된 값( $t_k$ )과 달라질 확률은

$$P_t = \begin{cases} G\left(-\frac{\eta_{b_k}}{\sigma_{b_k}}\right) & \text{if } t_k = 1 \\ G\left(\frac{\eta_{b_k}}{\sigma_{b_k}}\right) & \text{if } t_k = -1 \end{cases} \quad (10)$$

이다.

### 3. 입력패턴의 변동에 의한 출력의 변화

Madaline이 주어진 입력패턴을 제대로 학습하였다면, 입력패턴의 일부변이 바뀌더라도 출력패턴은 큰 영향을 받지 않을 것이다. 따라서, 입력패턴의 변동에 의한 출력패턴의 변화는 학습정도의 기준이 될 수 있다.<sup>[11]</sup> 여기서, 입력패턴의 일부변이 변할 경우 출력뉴런 값이 달라질 확률을 구하였다. -1과 +1로 이루어진  $\mathbf{x}$ 의 각 요소  $x_i$ 가

$$x_i' = \begin{cases} -x_i & \text{with probability } p \\ x_i & \text{with probability } 1-p \end{cases} \quad (11)$$

와 같이 변한다고 하자. 그러면,  $x_i'$ 와  $x_j'$ 는 서로 독립이며 분산이 유한하므로,  $N$ 이 충분히 클 경우  $a_i$ 는 central limit theorem에 의해 정규분포이다. 그 평균과 분산은

$$\eta_{a_i} = w_{j_0} + (1-2p)\sum_{k=1}^N w_{jk} x_k \quad (12)$$

$$\sigma_{a_i}^2 = 4p(1-p)\sum_{k=1}^N w_{jk}^2 \quad (13)$$

이며,  $a_i$ 와  $a_j$  사이의 상관계수는

$$r_{ij} = \frac{\sum_{k=1}^N w_{jk} w_{ik}}{\sqrt{\sum_{k=1}^N w_{jk}^2} \sqrt{\sum_{k=1}^N w_{ik}^2}} \quad (14)$$

이다. 식 (2)와 같이 변환 후,  $h_j$ 의 평균과 분산은 식 (6), (7)과 같다.  $h_i$ 와  $h_j$  사이의 상관계수는

$$r_{ij}^{(h)} = 2 \frac{\int_{-\eta_{a_i}}^{\eta_{a_j}} f(v)dv - \int_{-\eta_{a_j}}^{-\eta_{a_i}} f(v)dv + G\left(-\frac{\eta_{a_i}}{\sigma_{a_i}}\right) \left[ 1 - 2G\left(-\frac{\eta_{a_j}}{\sigma_{a_j}}\right) \right]}{\sigma_{a_i} \sigma_{a_j}} \quad (15)$$

이며, 여기서

$$f(v) = \frac{1}{\sigma_{a_i} \sqrt{2\pi}} \exp\left[-\frac{v^2}{2\sigma_{a_i}^2}\right] G\left(-\frac{1}{\sqrt{1-r_{ij}^2}} \left(\frac{\eta_{a_i}}{\sigma_{a_i}} + \frac{r_{ij} v}{\sigma_{a_j}}\right)\right) \quad (16)$$

이다. 다층퍼셉트론이나 Madaline 처럼 가중치와 비선형 변환의 층별 구조에 의해 패턴을 인식하는 신경회로망은 중간층 뉴런 간의 상관계수가 줄어드는 형태로 학습이 진행되므로<sup>[13, 8]</sup>,  $|r_{ij}^{(h)}|$ 가 1에 가까울 확률은 거의 없다. 따라서, 식 (3)과 같이 많은  $h_j$ 의 합으로 결정되는  $b_k$ 역시 앞 절에서와 마찬가지로 정규분포이다.  $b_k$ 의 평균  $\eta_{b_k}$ 는 식 (8)과 같으며 분산은

$$\sigma_{b_k}^2 = \sum_{j=1}^H v_{kj}^2 \text{Var}[h_j] + 2 \sum_{j=1}^{H-1} \sum_{i=j+1}^H v_{ki} v_{kj} C_{ij}^{(h)} \quad (17)$$

로 주어진다. 여기서,  $C_{ij}^{(h)}$ 는  $h_i$ 와  $h_j$ 사이의 covariance이다. 결국,  $k$ 번째 출력뉴런의 상태가  $p$ 에 의해 달라질 확률은 위에서 구한  $\eta_{b_k}$ 와  $\sigma_{b_k}$ 를 식 (10)에 대입하면 된다.

### III. 시뮬레이션

학습된 Madaline에서 가중치 혹은 입력패턴의 변동에 의해 출력뉴런 값이 달라질 확률의 계산 결과식

을 확인하기 위해, 40명으로 부터 수집한 필기체 숫자 3,000자를 Madaline에 학습시켰다. 필기체 숫자 패턴은  $16 \times 16$  크기이며, 각 요소는 -1 혹은 +1 값을 가진다. 256개의 입력뉴런, 35개의 중간층 뉴런, 그리고 10개의 출력뉴런으로 구성된 Madaline에 필기체 숫자를 학습시, 10개의 출력뉴런 중 하나만 +1이고 나머지는 -1이 되도록 하였다. 가령, 숫자 "4"는 다섯번째 출력뉴런이 +1이 되도록 학습하였으며, 이 다섯번째 출력뉴런을 숫자 "4"의 목표뉴런이라 한다. 학습완료 후 숫자 "4"에 해당하는 학습패턴 70자를 임의로 선택하여, 각 패턴에 대한 분석결과의 평균치로 "4" 부류(class) 학습패턴에 대한 특성을 추정하였다. 그림 1은 각 패턴에 대하여 가중치에 정규분포 잡음이 섞일 경우 목표뉴런의 상태가 변할 확률  $P_i$ 를 계산 및 시뮬레이션하여, 70개의 분석대상 패턴에 대한 평균치를 그린 것이다. 각 패턴에 대한  $N(0, \sigma_w)$  분포 잡음이 섞인 가중치의 시뮬레이션 횟수는 100번이다. 그림 1에서 계산 결과가 시뮬레이션 결과와 잘 일치함을 볼 수 있다.

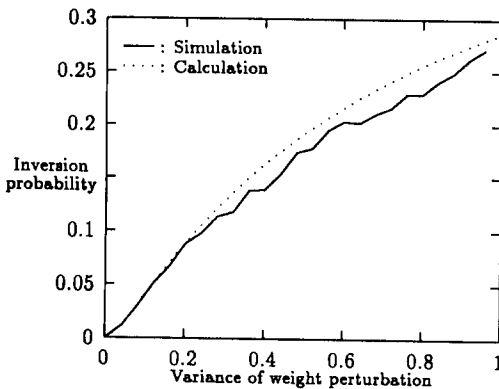


그림 1. 가중치 변동에 의한 목표뉴런의 상태변화 확률

Fig. 1. Inversion probability of target neuron to weight perturbation.

그림 2는 각 패턴의 요소들이 확률  $p$ 로 변할 경우 목표뉴런 상태가 변할 확률의 시뮬레이션 및 계산 결과를 70개의 패턴에 대하여 구한 후 평균한 것이다. 한 패턴의 특징  $p$ 에 대한 시뮬레이션 횟수는 200번이다. 역시, 시뮬레이션과 계산 결과가 잘 일치함을 볼 수 있다. 여기서,  $p$ 가 증가할수록  $P_i$ 가 단조증가하지 않고,  $p$ 가 0.5 근처에서 일시적으로 감소하다가 증가하는 현상이 나타나는데, 이것은 식 (12), (13)에서 보는 바와 같이  $\eta_{ax}$ 와  $\sigma_{ax}$ 가  $p$ 에 대하여 받는 영

향이 다르기 때문이다. 그 영향이  $\eta_{ax}$ 와  $\sigma_{ax}$ 에 나타나서  $g(-\frac{\eta_{ax}}{\sigma_{ax}})$ 가 3차 함수 같은 특성을 보인다. 패턴인식 Madaline은  $r_{ij}^{(h)}$ 가 줄어드는 형태로 학습이 진행되므로<sup>[13]</sup>, 식 (17)에서  $r_{ij}^{(h)}$ 에 관련된 두번째 항을 무시하면, 계산 결과가 다소 틀려짐을 그림 2에서 볼 수 있다. 이 정도의 계산오차를 허용한다면  $r_{ij}^{(h)}$ 를 계산할 필요가 없으므로 계산시간이 단축된다. 따라서, Orzechowski 등이 중간층 뉴런 간의 상관관계를 무시한 것은 어느 정도 타당하다그렇지만, Orzechowski 등이 제안한 방법은 중간층 뉴런의 상태가 잡음에 의해 적어도 3개 까지 변하는 사건을 고려한 것이므로<sup>[4]</sup>, 35개의 중간층 뉴런이 있으면 631가지 사건을 고려해야 한다. Stevenson 등은 모든 학습패턴이 균일한 분포로 존재한다는 가정에 기초하여 가중치의 변동에 의해 출력이 달라질 확률을 구하였다.<sup>[14]</sup> 그렇지만, 실제적인 문제에서 학습패턴의 분포를 알 수 없으므로, 전체 학습패턴을 대상으로 가중치 변동에 의한 출력의 영향을 구할 수는 없다. 학습패턴들 중에서 임의로 선택된 여러 패턴들에 대한 각각의 분석결과를 평균하여, 학습패턴들에 대한 특성을 추정해야 한다.

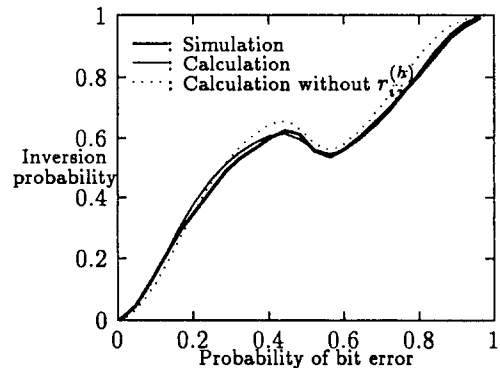


그림 2. 입력뉴런의 변동에 의한 목표뉴런의 상태변화 확률

Fig. 2. Inversion probability of target neuron to input errors.

#### IV. 결 론

신경회로망을 패턴인식에 응용할 때, 입력패턴의 변동에 따라 출력뉴런 값이 달라질 확률은 신경회로망 학습 정도의 기준이 될 수 있다. 또, 신경회로망의 H/W 구현 시 가중치 구현의 정밀도에 따른 출력

뉴런의 영향도 고려해야 한다. 구현에 따른 가중치의 변동은 가중치에 잡음이 섞인 것으로 해석할 수 있다. 이 논문에서는, 다층 신경회로망 모델 중의 하나인 Madaline에서 입력패턴의 변동과 가중치에 잡음이 섞일 경우에, 출력뉴런 값이 학습된 값과 달라질 확률을 수식으로 구하고, 필기체 숫자 인식 Madaline의 시뮬레이션으로 확인하였다. 계산된 수식을 이용하면, 패턴인식 문제를 학습한 Madaline의 학습정도를 판단하거나 가중치의 변동에 의한 영향을 분석할 수 있다.

### 감사의 글

\*이 논문은 한국통신의 출연금에 의해 수행한 연구의 결과입니다. 이 연구를 수행하는 데 많은 도움을 주신 이일항 박사께 감사드립니다.

### 参 考 文 献

- [1] J. Y. Choi and C. -H. Choi, "Sensitivity analysis of multilayer perceptron with differentiable activation functions," *IEEE Trans. Neural Networks*, vol. 3, pp. 101-107, 1992.
- [2] Y. Xie and M. A. Jabri, "Analysis of the effects of quantization in multilayer neural networks using a statistical model," *IEEE Trans. Neural Networks*, vol. 3, pp. 334-338, Mar. 1992.
- [3] M. Stevenson, R. Winter, and B. Widrow, "Sensitivity of feedforward neural networks to weight errors," *IEEE Trans. Neural Networks*, vol. 1, pp. 71-90, Mar. 1990.
- [4] N. S. Orzechowski, S. R. T. Kumara, and C. R. Das, "Performance of multilayer neural networks in binary-to-binary mappings under weight errors," *Proc. ICNN'93 San Francisco*, pp. 1684-1689, Mar. 1993.
- [5] A. von Lehmen, E. G. Paek, P.F. Liao, A. Marrakchi, and J. S. Patel, "Influence of interconnection weight discretization and noise in an optoelectronic neural networks," *Opt. Lett.*, vol. 14, pp. 928-930, Sept. 1989.
- [6] S-H. Oh and Y. Lee, "An analysis on the classification performance of multilayer perceptrons," *Proc. IJCNN '92 Beijing*, vol. II, pp. 787-792, Nov. 1992.
- [7] B. Widrow, R. G. Winter, and R. A. Baxter, "Layered neural nets for pattern recognition," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 36, pp. 1109-1118, July 1988.
- [8] B. Widrow and M. A. Lehr, "30 years of adaptive neural networks: Perceptron, Madaline and backpropation," *Proc. IEEE*, vol. 78, pp. 1415-1422, Oct. 1990.
- [9] S. M. Gowda, B. J. Sheu, and J. Choi, "Testing of programmable analog neural network processors," *Proc. IEEE 1992 Custom Integrated Circuits Conf.*, pp. 17.1.1-17.1.4, 1992.
- [10] D. B. Mundie and L. W. Massengill, "Weight decay and resolution effects in feedforward artificial neural networks," *IEEE Trans. Neural Networks*, vol. 2, pp. 168-170, Jan. 1991.
- [11] J. B. Lont and W. Guggenbuhl, "Analog CMOS implementation of a multilayer perceptron with nonlinear synapses," *IEEE Trans. Neural Networks*, vol. 3, pp. 457-465, May 1992.
- [12] A. Papoulis, *Probability, Random variables, and stochastic Processes*. McGraw-Hill, New York, 2nd edition, 1984.
- [13] S-H. Oh and Y. Lee, "Effect of nonlinear functions on correlation between weighted sums in multilayer perceptrons," *IEEE Trans. Neural Networks*, accepted for publication.

---

著 者 紹 介

吳 相 勳(正會員) 第 28卷 B編 第 4號 參照

李 永 稷(正會員) 第 28卷 B編 第 3號 參照