

논문 2004-41CI-6-7

교사학습과 비교사학습의 접목에 의한 두뇌방식의 지능 정보 처리 알고리즘 개발: 학습패턴의 생성

(Development of Brain-Style Intelligent Information Processing
Algorithm Through the Merge of Supervised and Unsupervised
Learning: Generation of Exemplar Patterns for Training)

오 상 훈*

(Sang-Hoon Oh)

요 약

시간/경제적 문제 혹은 수집 대상의 제한으로 충분한 수의 학습패턴을 모을 수 없는 경우에 인간의 두뇌를 모방한 교사 학습 및 비교사 학습 모델을 이용하여 새로운 학습패턴을 생성하는 알고리즘을 제안하였다. 비교사 학습은 독립성분분석을 사용하여 패턴의 특성을 분석 후 생성하며, 교사 학습은 다층퍼셉트론 모델을 사용하여 생성된 패턴의 검증에 하는 단계로 적용되었다. 통계학적으로 이와 같은 형태의 패턴 생성을 분석하였으며, 필기체 숫자의 학습 패턴 수를 변동시키면서 패턴 생성의 효과를 시험패턴에 대한 오인식률로 확인한 결과 성능이 향상됨을 보였다.

Abstract

We propose a new algorithm to generate additional training patterns using the brain-style information processing algorithm, that is, supervised and unsupervised learning models. This will be useful in the case that we do not have enough number of training patterns because of limitation such as time consuming, economic problem, and so on. We adopt the independent component analysis as an unsupervised model for generating exemplar patterns and multilayer perceptrons as supervised models for verifying usefulness of the generated patterns. After statistical analysis of the proposed pattern generation algorithm, we verify successful operations of our algorithm through simulation of handwritten digit recognition with various numbers of training patterns.

Keywords : 교사 학습, 비교사 학습, 패턴 생성, 일반화 성능.

I. 서 론

뇌 정보처리 방식의 공학적 응용을 다루는 뇌공학에서 현재 활용 및 개발 중인 학습 알고리즘은 크게 교사 학습(supervised learning)과 비교사 학습(unsupervised learning)으로 대별되며, 그 각각이 뇌의 이해를 기반으로 한다. 대표적인 예로, 가장 널리 활용되는 교사 학습

모델인 퍼셉트론(Perceptron)은 Albus와 Marr의 이론적 고찰과 Masao Ito의 실험적 고찰에 의해 소뇌의 작용과 비슷하다는 것이 알려졌다^{[1][2]}. 또한, 최근에 각광을 받고 있는 비교사 학습 모델인 ICA(Independent Component Analysis)에 의해 Bell과 Sejnowski는 시각 피질의 V1 layer에 존재하는 simple cell의 특성을 재현 하였으며^[3], Apo Hyvarinen은 topographic ICA 방법에 의해 simple feature들의 self organizing 특성도 보였다^{[4][5]}. 이수영 교수팀은 음성신호를 대상으로 ICA 학습을 실시하여 추출된 특징이 귀의 달팽이관에서 추출되는 청각특징과 유사함을 보였다^[6].

이제까지 교사 학습과 비교사 학습은 주로 각각의 고

* 중신회원, 목원대학교 정보통신공학부
(Div. of Information Communication Engineering ,
Mokwon University)
※ 이 논문은 학술진흥재단의 지원에 의하여 연구되었
음 (KRF-2003-003-D00245)
접수일자: 2004년7월1일, 수정완료일: 2004년11월15일

유한 영역 내에서 연구가 수행되었다. 그렇지만, 이와 같은 교사학습과 비교사 학습의 접목은 그 각각이 나타낼 수 없었던 많은 효과를 얻을 수 있을 것이다^{[7][8]}. 그 한가지 예가 ICA와 MLP(Multi-Layer Perceptron)를 접목시켜 선택적 주의 집중(selective attention) 효과를 구현하여 잡음환경에서 음성인식 성능을 향상시킨 것이다^[7].

본 논문은 교사/비교사 학습의 접목에 의해 패턴인식, 음성인식, 의료진단(medical diagnosis) 등에 활용이 가능한 학습패턴 생성 알고리즘을 제안하고자 한다.

인간의 두뇌를 모방한 지능정보처리 모델은 학습패턴을 대상으로 학습을 수행하여 정보처리에 필요한 모델이 완성되도록 한다. 그렇지만, 실세계에서는 시간/경제적 혹은 학습패턴 추출 대상의 제약으로 인해 충분한 수의 학습패턴을 수집할 수 없는 경우가 많다. 만약, 충분한 양의 학습패턴을 만들 수 집한 경우 이를 근거로 학습에 필요한 패턴을 생성시켜 지능정보처리 모델이 학습을 수행토록 한다면, 위와 같은 시간 및 경제적 제약 혹은 추출대상의 제한으로 인한 문제들을 해결할 수 있을 것이다. 즉, 의료 진단분야에서 아주 많은 수의 환자를 대상으로 학습패턴을 수집할 수 없는 응용문제인 경우에도 학습을 수행하여 자동 진단 시스템을 구축할 수 있다. 또한, 음성인식과 같이 학습패턴 수집에 많은 시간과 경제적 지출이 요구되는 경우에 학습패턴 생성을 이용하면 시간적/경제적 지출을 줄일 수 있다.

학습패턴의 생성과 관련된 연구는 크게 세 가지 형태로 진행되어 왔다. 그 첫 번째는 현재의 학습패턴에 임의의 잡음(random noise)을 더하는 형태로 학습패턴을 변형시켜서, 학습에 의해 얻게 되는 일반화 성능을 향상시키는 방법이다^{[9][10][11]}. 그렇지만, 단지 학습패턴에 임의의 잡음을 더하는 것은 학습패턴의 특성을 제대로 반영하여 새로운 학습패턴을 만들어 내는 것이라고 할 수 없다.

두 번째는 이 보다 조금 더 진전된 형태로써 수집된 학습 패턴을 부분적 요소가 없는 missing 패턴과 모든 요소가 제대로 갖추어진 complete 패턴으로 분류하여, missing 패턴의 복원을 시도하는 방법이다^{[12][13]}. 즉, 먼저 완전한 데이터로 MLP를 학습시킨다. 다음에 missing 부분을 데이터들의 평균값으로 대체 시킨 후, MLP에서 입력값을 학습에 의해 찾아가는 형태로 데이터를 복원시킨다. 이 방법은 missing 패턴의 모양을 학습에 의해 찾아가는 방식이지만, 학습에 필요한 새로운

패턴을 생성하는 방법은 아니다.

마지막 세 번째 방법은 학습패턴의 prototype을 설정한 후, RDT(random deformation technique)를 적용하여 특정 class에 속하는 패턴들을 생성시키는 방법이다^[14]. 그렇지만, 이 방법은 RDT의 적용 결과 생성된 패턴이 특정 class에 속하는 지 여부를 사람의 눈으로 판단하므로, 결국 사람이 일일이 개입하는 방식이다.

이와 달리, 본 논문에서는 기본 패턴을 근거로 학습 패턴 자체를 자동으로 생성시키는 알고리즘을 제안하고자 하는 것으로써, 이상의 방법과는 확연히 차이가 있으며, 특히, 비교사학습과 교사학습의 접목에 의한 방법을 찾고자 한다. 즉, 비교사학습에 의해 패턴들의 특성을 분석해내고, 그 특성을 반영시켜 기본 패턴들의 변형을 만들어 내고자 한다. 그리고, 이렇게 생성된 변형 패턴들을 교사학습 모델에 입력시켜 특정 class에 속하는 지 여부를 판정하도록 하고자 한다.

II. 비교사 및 교사 학습 모델

1. 독립성분분석을 이용한 패턴 생성

학습패턴 생성을 위한 비교사 학습 모델은 최근에 개발된 ICA 알고리즘을 기반으로 한다. ICA란 그림 1과 같이 상호 독립적인 신호원들(s)이 섞여서 측정된 신호들(x)로부터, 상호 독립적이라는 것 이외에는 신호원들에 대한 정보가 전혀 없다는 가정 하에, 신호원을 다시 복원해 내는 학습 방법이다. 즉, ICA가 주어진 데이터 x에서 변환을 거친 뒤의 출력(u)이 지닌 정보량 혹은 엔트로피가 최대가 되도록 학습을 수행하며, 그 결과 얻어지는 변환 벡터들은 독립성분에 대한 정보를 지니게 된다^{[16][17]}. 특히, 입력신호와 분리해내고자 하는 신

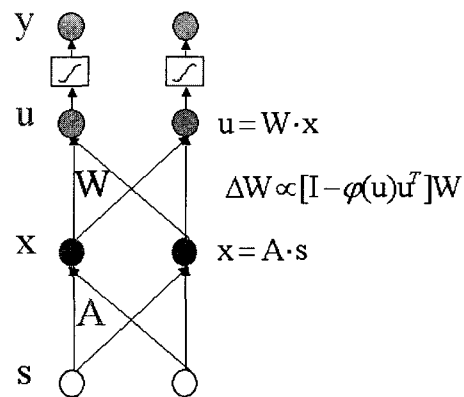


그림 1. ICA 네트워크 및 학습방법

Fig. 1. An ICA network and its training method.

호의 수가 같지 않을 경우에도 이 방법을 적용하여 데이터 내부에 존재하는 구조를 파악할 수 있다^[16].

ICA는 또 다른 측면에서 projection pursuit learning으로도 해석할 수 있으며^[19], 이는 데이터가 지닌 특성을 잘 반영하는 투사(projection) 방향을 찾아내는 것으로 해석할 수 있다. 따라서, 데이터의 특성을 잘 반영하는 여러 개의 투사 방향을 ICA로 찾아낸 후, 이를 기반으로 여러 가지로 변형된 학습패턴을 생성하는 알고리즘을 고안한다. 물론, PCA(Principle Component Analysis)를 이용하여 정보의 변환된 형태를 표시하는 경우도 있으나, PCA는 변환된 정보들 간의 2차 통계적 특성인 상관관계(correlation) 만을 고려한 방법이지만, ICA는 고차 통계적 특성도 고려한 방법이어서 데이터가 지닌 구조적 특성을 보다 더 잘 반영한다고 할 수 있다.

2. 다층퍼셉트론을 이용한 과도변형패턴 제거

ICA에 의해 생성된 패턴 중 과도하게 변형된 것을 학습패턴의 대상에서 제외시키는 방법은 교사 학습 모델 중 가장 뛰어나다고 평가받는 MLP를 기반으로 고안한다. MLP는 그림 2와 같은 구조를 가지며, 학습에 의해 패턴을 분류하는 기능을 수행하는 데, 이때 학습에 의해 얻어진 MLP의 출력(y)은 입력패턴이 특정 클래스(class)에 속할 확률을 나타낸다^[20]. 또한 MLP의 각 계층은 패턴의 분류에 필요한 정보는 유지시키면서, 패턴의 분류에 필요치 않은 정보는 제거시키는 기능도 수행한다. MLP의 동작을 이러한 정보처리 관점에서 조명하여, Shannon 엔트로피의 일반화된 형태인 Renyi 엔트로피를 MLP의 학습에 도입하여 학습을 개선시킬 수도 있다^[21]. 또한 학습 목적에 따라 중간층 뉴런들이 가

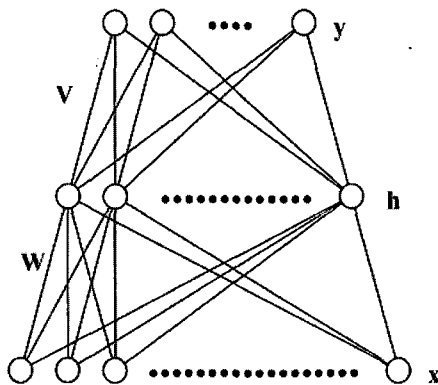


그림 2. 다층퍼셉트론 구조
Fig. 2. A structure of Multilayer Perceptron.

지는 정보량을 조절하는 방법도 있다^[22]. 본 논문에서는 MLP의 학습에 의한 패턴 분류기능 및 클래스 확률 표현 특성을 이용하여, 비교사 학습 방법인 ICA에 의해 생성된 패턴 중 그 변형 정도가 과도하여 특정 클래스에 속하지 않는 패턴들은 제외시키는 기능을 수행한다.

III. 학습패턴 생성

주어진 패턴들에 ICA 알고리즘을 적용하여 얻어진 basis 벡터들을 이용하여 패턴들을 생성시키는 과정을 정리하면 다음과 같다.

먼저, 주어진 학습 패턴 x 는 PCA 분석을 수행하여

$$z = Vx \tag{1}$$

를 얻어낸다. 여기서 V 는 eigenvectors로 이루어진 행렬이다. 그 다음 z 에서 eigenvalues가 작은 성분은 제거하여 차원이 축소된 vectors를 \hat{z} 이라 두자. 이 \hat{z} 에 대해 ICA를 적용시켜

$$u = W\hat{z} \tag{2}$$

를 얻게 된다. 한편, u 는 그 요소들이 독립이 되도록 ICA 학습이 이루어졌으므로 u 의 확률밀도함수(p.d.f.: probability density function)는

$$p(u) = \prod_i p(u_i) \tag{3}$$

이다. 따라서, PCA 절차를 거친 후 ICA 학습에 의해 u 를 얻어낸 다음, $p(u_i)$ 에 따라 u 들을 발생시켜 역의 과정을 거쳐 x 를 만들어 낸다. 이것이 ICA를 이용하여 학습패턴을 생성하는 과정이다.

이제, u 의 p.d.f. $p(u)$ 를 이용하여 생성된 패턴의 p.d.f. $p_x(x)$ 를 유도해보자. 먼저,

$$\hat{z} = W^{-1}u = Au \tag{4}$$

이므로

$$p_z(\hat{z}) = \frac{p(W\hat{z})}{|A|} \tag{5}$$

이다. 여기서, $A = W^{-1}$ 이다. 이제 $z = [\hat{z}0]^T$ 라고 두면 z 는 x 와 같은 차원을 지닌 생성된 벡터이며,

$$p_z(\mathbf{z}) = p_z(\hat{\mathbf{z}}, \mathbf{0}) = p_z(\hat{\mathbf{z}})p_z(\mathbf{0}) = p_z(\hat{\mathbf{z}}) \quad (6)$$

로 주어진다. 여기서, $\tilde{\mathbf{z}}$ 는 \mathbf{z} 에서 $\hat{\mathbf{z}}$ 를 제외한 나머지 요소들을 나타내며, $\tilde{\mathbf{z}} = \mathbf{0}$ 일 확률이 1이기 때문에 식 (6)의 마지막 부분이 성립된다. 그러면, 생성된 학습패턴 \mathbf{x} 의 p.d.f.는 식 (1)의 관계를 이용하면

$$p_x(\mathbf{x}) = \frac{p_z(\mathbf{V}\mathbf{x})}{|\mathbf{V}^{-1}|} \quad (7)$$

에서 유도된다. $\mathbf{V}\mathbf{x}$ 에서 eigenvalues가 큰 순서로 PCA에 따라 차원 축소 후 얻어진 차원의 벡터를 $(\mathbf{V}\mathbf{x})_z$ 라 두면

$$p_x(\mathbf{x}) = \frac{p_z(\mathbf{V}\mathbf{x})}{|\mathbf{V}^{-1}|} = \frac{p_z([\mathbf{V}\mathbf{x}]_z)p_z(\mathbf{0})}{|\mathbf{V}^{-1}|} \quad (8)$$

이므로, 식 (5)와 (6)을 이용하면

$$p_x(\mathbf{x}) = \frac{p_z([\mathbf{V}\mathbf{x}]_z)}{|\mathbf{V}^{-1}|} = \frac{p(\mathbf{W}[\mathbf{V}\mathbf{x}]_z)}{|\mathbf{V}^{-1}| |\mathbf{A}|} \quad (9)$$

로 얻어진다. 만약 $p(u_i)$ 를 정확히 알고 있다면 생성해 내는 패턴들 역시 정확한 학습패턴일 것이다. 그렇지만 현실적으로 이를 제대로 알아내는 것은 힘들므로 이 논문에서는 $p(u_i)$ 가 가우시안(Gaussian) 분포를 지니도록 한다.

ICA에 의해 생성된 패턴들 중 과도한 변형 패턴의 제거는 MLP를 이용한다. 먼저, MSE(Mean Squared Error)를 이용한 MLP의 학습은 부적절한 포화(incorrect saturation)를 유발하는 성질을 이용한다. 즉, MLP의 학습 시 목표값과 다른 포화영역에 출력값이 위치하게 되는 경우, 이는 다른 패턴들과 달리 학습이 아주 힘든 패턴이거나 혹은 잘못된 패턴으로 볼 수 있으므로 제거시킨다^[20]. 마지막으로, 기존의 학습패턴들과 ICA에 의해 생성되었지만 MSE를 이용한 MLP의 학습에서 제거되지 않은 패턴들을 모아서 패턴 인식기를 다시 학습 시킨 후 일반화 성능을 검사하게 된다. 이 단계에서는 nCE (nth order extension of Cross-Entropy) 오차함수를 이용한 MLP 학습^[20]을 실시하여 패턴인식기의 일반화 성능을 조사한다.

IV. 시뮬레이션

제안한 패턴생성 방법의 효용성을 확인하기 위하여 충분한 학습패턴 및 시험패턴을 확보하고 있는 CEDAR

데이터베이스를 사용하였다^[23]. 이 데이터베이스에 있는 18468개의 패턴으로 제안한 방법에 따라 학습패턴을 생성한 후, MLP의 MSE(mean squared error)에 따른 학습 시 부적절한 포화가 발생하는 패턴들을 제외시킨 다음, MLP를 nCE 오차함수^[20]를 최소화 시키는 방법으로 학습시켜 일반화 성능이 향상되는 지 확인하였다. 또한, 학습 패턴 수를 9234개와 4617개인 경우도 시뮬레이션 하여 보았다. 이때, PCA에 의한 차원 축소 후 얻어지는 벡터는 필기체 숫자의 eigenvalue를 조사하여 70 차원이 되도록 하였다.

패턴 생성은 이 논문에서 제안한 것과 같이 ICA의 출력단에 잡음을 첨가한 방법과 PCA 출력단에 잡음을 첨가하는 방법을 실시하여 보았다. PCA 출력단에 잡음을 첨가하여 패턴을 생성하는 방법이란 식 (1)과 같이 PCA를 수행 후 차원이 축소된 벡터 $\tilde{\mathbf{z}}$ 에 잡음을 더한 후 역으로 \mathbf{x} 를 생성한 것을 말한다. 패턴 생성 결과 PCA 출력단에 잡음을 첨가한 경우는 그림 3에서 보는 바와 같이 생성된 패턴이 기존 패턴에 잡음이 섞인 형태로 보인다. 그림 3에서 첫 열은 입력패턴, 2번째 열은 가우시안 잡음의 표준편차가 0인 경우 생성된 패턴, 4번째부터는 차례로 가우시안 잡음의 표준편차가 0.1씩

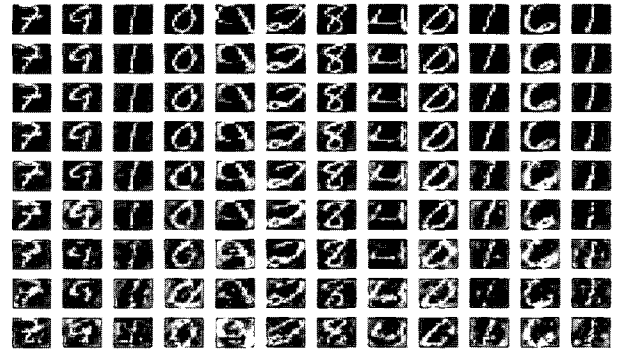


그림 3. PCA에 의한 학습패턴 생성

Fig. 3. Generated patterns using PCA.



그림 4. 제안한 방법에 의해 생성된 패턴

Fig. 4. Generated patterns using the proposed algorithms.

증가하면서 생성된 패턴들이다.

이와 반면에, 본 논문에서 제안한 바와 같이 ICA 출력단에 잡음을 더한 경우는 그림 4에서 보는 것처럼 생성된 패턴이 기존 패턴의 모양을 많이 유지하고 있는 상태에서 패턴이 흐리게 변형되는 형태임을 볼 수 있다. 새로운 패턴 생성 효과를 확연히 보일 수 없는 이유는 앞 절에서 설명한 바와 같이 현실적으로 $p(u_i)$ 를 제대로 알아내는 것은 힘들므로 이 논문에서는 가우시안(Gaussian) 분포를 지니도록 하였기 때문이다. 여기서 첫 열은 입력패턴, 2번째 열부터는 가우시안 잡음의 표준편차가 0부터 0.1씩 증가하면서 생성된 패턴들이다.

그림 5는 그림 4와 같이 생성된 패턴들에 대하여 MSE를 이용한 MLP의 학습으로 과도 변형 패턴들을 제거한 후, nCE를 이용한 MLP의 학습으로 시험패턴에 대한 오인식률을 조사한 그래프이다. 이 그림에서 x축은 패턴 생성 시 발생시키는 가우시안 잡음의 표준편차가 72번째 eigenvalue의 몇 배인지를 나타낸다. 학습패턴이 18468개, 9234개, 그리고 4617개인 경우에 각각 제안한 알고리즘을 적용시켜 얻은 결과이며, MLP의 초기 가중치를 10번 다르게 적용시켜 얻은 결과의 평균치를 그림으로 그렸다. 또한, 수치의 명확한 값을 알아보기 쉽도록 표 1에 오인식률을 적어두었다. 먼저, 그림 5의 x축이 0인 경우는 패턴 미생성에 해당하는 경우로써, 오인식률이 각각 4.60%(학습패턴 수 18468개), 5.38%(학습패턴 수 9234개), 7.17%(학습패턴 수 4617개)이다. 한편, 패턴 생성 시 얻어진 시험패턴에 대한 오인식률의 평균값들 중, 학습 패턴이 18468개인 경우는 가우시안 잡음의 표준편차가 6배(n_sig=6)일 때 가장 좋은 결과인 오인식률 3.87%임을 표1에서 볼 수 있으며 이는 패턴 미생성인 경우보다 15.9%의 성능 향상이 있음을 나타낸다. 학습 패턴수 9234개인 경우는 n_sig가 7일 때 11.9%의 성능향상이 있으며, 학습 패턴 수 4617인 경우도 비슷한 성능 향상이 있다. 그렇지만, 비록 패턴 생성을 하더라도 기본적으로 학습패턴을 많이 가진 경우 보다 좋은 성능을 얻을 수 없음을 알 수 있다. 이는 앞에서 이야기 한 바와 같이 $p(u_i)$ 를 정확히 알고 있지 못하여 Gaussian 분포를 지닌다는 가정 하에 패턴 생성을 하였기 때문이다. 만약, $p(u_i)$ 를 정확히 안다면 훨씬 더 좋은 결과를 얻을 수 있을 것이다.

한편, 제안한 방법의 성능을 검증하기 위하여 입력패턴에 임의의 잡음을 더하는 형태로 패턴을 생성한 경우 MLP를 이용한 일반화 성능도 조사하여 보았다. 학습

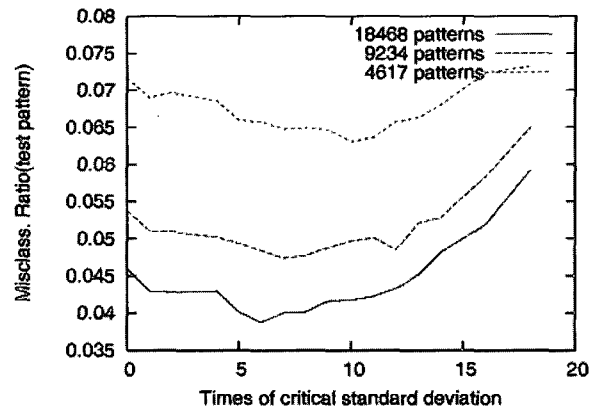


그림 5. 제안한 패턴 생성 후 일반화 성능 조사결과
Fig. 5. Misclassification ratio of test patterns using the proposed pattern generation algorithm.

표 1. 패턴 생성에 의해 얻어진 최고 성능의 시험패턴 오인식률
Table 1. The best of misclassification ratio of test patterns in each simulation with various numbers of training patterns.

학습패턴수	18468	9234	4617
패턴미생성	4.60%	5.38%	7.17%
제안한 패턴생성	3.87% (n_sig=6)	4.74% (n_sig=7)	6.31% (n_sig=10)
입력잡음 패턴생성	3.91% (n_sig=5)	4.91% (n_sig=4)	6.67% (n_sig=3,4)

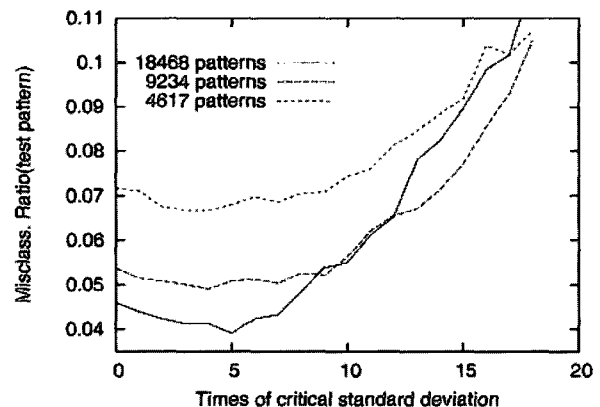


그림 6. 입력 잡음 패턴 생성 후 일반화 성능 조사 결과
Fig. 6. Misclassification ratio of test patterns using the random input noise method.

대상 및 패턴 생성을 위한 잡음의 표준편차 및 MLP의 학습방법은 그림 5의 경우와 동일하도록 하여 시뮬레이션 결과를 그림 6에 그렸다. 입력잡음에 의해 얻어질 수 있는 향상된 오인식률은 표 1에 정리된 바와 같이 3.91%(학습패턴 수 18468개), 4.91%(학습패턴 수 9234

개), 6.67%(학습패턴 수 4617개)이다. 이는 표 1의 제안한 패턴생성 방법 결과보다 성능이 나쁨을 알 수 있다. 또한, 그림 6에서 보는 바와 같이 패턴 생성 잡음을 크게 할수록 오인식률이 아주 급격하게 나빠짐도 확인된다.

V. 결 론

이 논문에서는 비교사학습인 ICA를 이용하여 패턴의 숨겨진 구조를 분석한 후 이를 바탕으로 새로운 학습패턴들을 생성시켰다. 이 생성된 패턴들은 MSE를 최소화시키는 MLP의 학습 후, incorrect saturation을 일으키는 패턴을 제외시켰다. 마지막으로, 기존의 학습패턴과 생성된 패턴을 대상으로 nCE를 최소화시키는 MLP를 수행하여 시험패턴에 대한 오인식률을 점검한다. 이상의 방법을 CEDAR 데이터 베이스에 적용하여 효용성을 입증하였다. 제안한 방법은 충분한 수의 학습패턴을 수집하기 힘든 상황에 적용하여 시스템의 성능을 향상시킬 수 있을 것이다.

참 고 문 헌

- [1] M. Tatsuno, "Computational Neuroscience-Methods in Neuronal modeling," Educational Lecture, BSI Retreat, Oiso, Japan, Sept. 18-20, 2000.
- [2] J. Albus, "A theory of cerebellar function," *Mathematical Biosciences*, vol. 10, pp. 25-61, 1971.
- [3] A. J. Bell and T. J. Sejnowski, "The independent components of natural scenes are edge filters," *Vision Research*, vol. 37, pp. 3327-3338, 1997.
- [4] A. Hyvarinen, P. O. Hoyer, and M. Inki, "Topographic independent component analysis," *Neural Computation*, vol. 13, no. 7, pp. 1527-1558, 2001.
- [5] A. Hyvarinen, P. O. Hoyer, and J. Hurri, "Extensions of ICA as models of natural images and visual processing," *Proceedings of ICA*, April 1-4, 2003, Nara, Japan.
- [6] J. H. Lee and S.-Y. Lee, "On the efficient speech feature extraction based on independent component analysis," *Neural Processing Letters*, vol. 15, no. 3, pp. 235-245, June 2002.
- [7] U.-M. Bae and S.-Y. Lee, "A complementary approach to blind signal separation for real-world speech recognition," *Advances in Neural Information Processing Systems 13*, pp. 765-771, Nov. 2000.
- [8] I. V. Tetko and A. E. P. Villa, "Efficient partition of learning data sets for neural network training," *Neural Networks*, vol. 10, pp. 1361-1374, 1997.
- [9] C. Wang and J. C. Principe, "Training neural networks with additive noise in the desired signal," *IEEE Trans. Neural Networks*, vol. 10, pp. 1511-1517, 1999.
- [10] L. Holmstrom and P. Koistinen, "Using additive noise in back-propagation training," *IEEE Trans. Neural Networks*, vol. 3, pp. 24-38, 1992.
- [11] K. Matsuoka, "Noise injection into inputs in back-propagation learning," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 22, no. 3, pp. 436-440, 1992.
- [12] S.-Y. Yoon and S.-Y. Lee, "Training algorithm with incomplete data for feed-forward neural networks," *Neural Processing Letters*, vol. 10, pp. 171-179, 1999.
- [13] Z. Ghahramani and M. Jordan, "Supervised learning from incomplete data via an EM approach," *Advances in Neural Information Processing Systems 6*, pp. 120-127, Nov. 1994.
- [14] B. Kamgar-Parsi, B. Kamgar-Parsi, J. E. Dayhoff, and A. K. Jain, "Improving classification boundaries by exemplar generation for visual pattern discrimination," *Proc. IJCNN2001*, vol. 4, pp. 2969-2974.
- [15] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, pp. 607-609, 13 June, 1996.
- [16] T.-W. Lee, et al., "A unifying information-theoretic framework for independent component analysis," *Computers & Mathematics with Applications*, vol. 31, no. 11, pp. 1-21, March 2000.
- [17] S.-H. Oh, A. Cichocki, S. Choi, S.-I. Amari, and S.-Y. Lee, "Comparison of ICA/BSS algorithms in noisy environment," *Proc. ICONIP*, vol. 2, pp. 1192-1197, Nov. 2000.
- [18] M. Girolami, A. Cichocki, and S.-I. Amari, "A common neural-network model for unsupervised exploratory data analysis and independent component analysis," *IEEE Trans. Neural Networks*, vol. 9, no. 6, Nov. 1998.
- [19] J. Karhunen, et al., "A class of neural networks for independent component analysis," *IEEE Trans. Neural Networks*, vol. 8, pp. 486-504, 1997.
- [20] S.-H. Oh, "Improving the error back-propagation algorithm with a modified error function," *IEEE Trans. Neural Networks*, vol. 8, pp. 799-803,

1997.

- [21] J. C. Principe, et al., "Learning from examples with information theoretic criteria," *The Journal of VLSI Signal Proc., special issue on neural networks*, vol. 26, pp. 61-77, Aug. 2000.
- [22] R. Kamimura and S. Nakanishi, "Hidden information maximization for feature detection and rule discovery," *Network: Computation in Neural Systems*, vol. 6, pp. 577-602, 1995.
- [23] J. J. Hull, "A database for handwritten text recognition research," *IEEE Trans. Pat. Ana. Mach. Int.*, vol. 15, no. 5, pp. 550-554, May 1994.

저 자 소 개



오 상 훈(중신회원)

1986년 2월 부산대학교 전자공학과 학사
 1988년 2월 부산대학교 대학원 전자공학과 석사
 1999년 8월 한국과학기술원 전기 및 전자공학과 박사
 1988년~1989년 금성반도체(주) 사원
 1990년~1998년 한국전자통신연구원 선임연구원
 1999년~2000년 KAIST 뇌과학연구센터 연구원

2000년 일본 RIKEN Brain Science Institute 연구원
 2000년~2001년 (주)엑스텔테크놀로지 연구소장
 2001년~2002년 2월 한국과학기술원 초빙교수
 2002년 3월~현재 목원대학교 정보통신공학부 조교수
 <주관심분야 : 지능신호처리, ICA, BSS>

