

비선형 주요성분 분석을 위한 신경망의 효율적인 학습알고리즘

조용현, 윤중환, 박창환
 대구효성가톨릭대학교 컴퓨터정보통신공학부

A Efficient Learning Algorithm of Neural Networks for Nonlinear PCA

Yong-Hyun Cho, Joong-Hwan Yun, Chang-Hwan Park
 Dept. of Computer Information Communication, Catholic University of Taegu HyoSung

요 약

본 논문에서는 데이터 내의 비선형 속성을 보다 빠르고 정확하게 추출하기 위한 수정된 학습알고리즘의 비선형 주요 성분분석 신경망을 제안한다. 제안된 학습알고리즘은 신경망의 학습시에 과거의 속성을 반영하기 위한 모멘텀 항이 추가된 학습기법이다. 이는 최적해로의 수렴에 따른 발전을 억제하여 그 수렴성능을 좀 더 개선시키는 모멘텀의 장점을 그대로 살리기 위한 것이다. 제안된 학습알고리즘을 이용한 신경망을 128×128 픽셀의 Lenna와 256×128 픽셀의 차량 번호판 영상들을 대상으로 시뮬레이션 한 결과, 제안된 학습알고리즘이 기존의 비선형 주요성분 분석을 위한 신경망이나 선형속성을 가지는 역전파 알고리즘을 이용한 신경망보다 더욱 우수한 수렴 성능과 특징추출 성능이 있음을 확인하였다.

1. 서론

최근 대규모 분산 병렬처리와 학습 능력을 가진 신경망은 신호와 패턴처리 분야 그리고 데이터 마이닝 분야 등에서 데이터를 압축하는데 널리 이용되고 있다. 데이터 압축에 있어서 일련의 과정은 고차원 공간으로부터 저차원 공간으로의 사상이라고 할 수 있는데, 차원 축소 문제는 특징추출과 밀접한 관계가 있다. 특징추출은 인식이나 분류와 같은 문제에서 데이터의 사용을 용이하게 하기 위해 속성이나 두드러진 현상을 식별하기 위한 것이다. 이러한 데이터의 특징들을 추출하기 위한 많은 기법들이 제시되어 왔다[1-2].

주요성분분석(principal component analysis : PCA)은 주어진 입력데이터의 특징을 추출하는 기법으로 이용되고 있지만, 이는 입력데이터에 대한 공분산행렬(covariance matrix)과 고유치(eigenvalue) 그리고 그에 따른 고유벡터(eigenvector)를 계산하기 위한 수식적인 과정이 필요하다. 그래서 대규모의 데이터를 대상으로 할 때 공분산행렬과 고유치 그리고 고유벡터를 계산하는 과정이 매우 복잡하고 부정확하게 이루어지며 계산시간도 오래 걸린다. 이러한 문제를 보완하기 위해 PCA를 신경망에 적용하였지만, 이 방법은 입력 데이터의 분포가 선형적인 관계에 있을 때는 효율적으로 데이터의 특징을 추출하지만 비선형적인 관계가 존재할 경우 특징을 추출하는 데는 어려움이 존재한다[3-6]. 따라서 입력 데이터의 비선형성이 강할 경우 PCA의 이용은 매우 비효율적이다. 이러한 제약을 해결하기 위해서 비선형 주요성분분석(nonlinear principal component analysis : NLPCA)이 제안되었다[3-9]. 그 중에서 Kramer는 비선형 특징들을 추출하기 위하여 5 계층 구조의 신경망(neural network : NN)을 제안하였다[3]. 이 기법을 기반으로 한 신경망이 여러 분야에 응용되어왔다. Kari Mantere[7] 등은 광학분야에서 반대되는 색의 신호와 무색 신호의 백색광 적용 특성을 시뮬레이션하고 근사화하는데 응용하였으며, Dimitrios Tzovaras[8] 등은 영상처리분야에서 스틸 이미지를 사용하여 부호화하고 복호화하는데 응용하였고 Fabio Del Frate[9]등은 지구과학분야에서 기존과 증기의 대기 중 축면의 방사선 반전을 위해 응용하였다.

그러나 Kramer의 NLPCA에 기반한 이러한 방법들은 비선형 속성

의 추출을 위한 신경망을 학습시키는데 매우 오랜 시간이 걸린다는 제약이 있다[4-9]. 이러한 제약을 개선한다면 그 성능을 좀 더 증가시켜 NLPCA를 효율적으로 이용할 수 있을 것이다.

본 논문에서는 모멘텀 항(momentum)을 추가한 학습 알고리즘을 제안하여 NLPCA를 위한 신경망을 학습시킴으로서 입력데이터의 주요 특징을 보다 빠르고 정확하게 추출되도록 하였다. 제안된 알고리즘을 이용한 신경망을 128×128 픽셀의 Lenna와 256×128 픽셀의 차량 번호판 영상들을 대상으로 시뮬레이션하여 선형 속성을 가지는 역전파 알고리즘 및 기존 NLPCA의 방법을 이용한 신경망에 의한 결과와 비교 고찰하였다.

2. 효율적인 비선형 주요성분 분석을 위한 신경망의 학습 알고리즘

NLPCA는 PCA의 방법과 같은 수단으로 다변량 데이터의 분석을 위한 새로운 기법이다. PCA에서와 마찬가지로 NLPCA는 실험 데이터 분석과 가시화, 그리고 차원 축소를 목적으로 함으로서 변수들간의 상호관계를 제거하고 식별하기 위해 사용된다. 하지만 PCA는 단순히 변수들 간의 선형적인 상호관계만을 식별하나, NLPCA는 데이터들 속에 나타나는 특성들을 선형적이거나 비선형적인 상호관계로 구분할 수 있다. NLPCA에서는 입력층의 입력 데이터가 출력층에 재구성되는 동일 사상을 수행도록 신경망을 학습시킨다. 이는 실제 입력 데이터의 분포에 얽은 특징 공간을 생성함으로써 차원을 축소시키기 위함이다.

그림 1은 NLPCA를 위한 일반적인 신경망의 구조를 나타낸 것이다. 그림에서 계층 1과 계층 2로 이루어진 부분장은 입력데이터를 특징공간으로 사상하는 부호화 망이다. 여기서 특징층, 즉 계층 2의 뉴런 p 는 $p < n$ 인 뉴런의 수를 가지게 한다. 계층 2에서 계층 4까지는 입력 데이터의 주요 특징들을 출력층으로 사상하는 복호화망이다. 일반적으로 비선형 사상을 위한 부호화망과 복호화망에서의 각 계층 1과 3은 비선형 함수를 사용해야 한다. 이때 널리 이용되는 함수는 S-자형함수 $f(x) = 1/(1 + e^{-x})$ 이다. 나

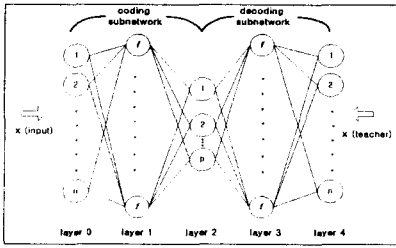


그림 1. NLPCA를 위한 신경망의 구조.

머지 계층들은 선형 함수나 비선형 함수가 이용될 수 있다.

NLPCA를 위한 신경망의 입력-출력 함수관계를 나타내면 다음과 같다. 즉,

$$\sigma(x) = \sum_{i=1}^N \overline{w}_i f(\overline{w}_i x + \overline{\theta}_i) + \overline{\theta} \quad (1)$$

이다. 여기서 \overline{w}_i 와 $\overline{\theta}$ 및 w_i 와 θ_i 는 각각 상위 및 하위계층의 연결가중치와 문턱치들이다. $\sigma(x)$ 는 범용 근사화기(universal approximator)로 N 이 충분히 크다면 임의의 정확성을 가진 어떤 경계함수(bounded function)를 근사화 할 수 있다. 그러나 $\sigma(x)$ 는 선형의 출력장치로 s-자형의 은닉장치를 가진 2-계층망으로 구성된 함수이다.

$a_i(l)$ 는 계층 l 의 뉴런 i 번째의 활성화 함수이고, $u_i(l)$ 는 l 계층의 연결강도와 $l-1$ 계층의 입력 값의 곱으로 이루어진 가중된 합이며, $w_{ij}(l)$ 는 계층 l 의 i 번째 뉴런과 $l-1$ 계층의 j 번째 뉴런사이의 연결강도, $\theta_i(l)$ 를 계층 l 의 뉴런 i 를 위한 바이어스(bias)라 할 때 $a_i(l)$ 과 $u_i(l)$ 는 각각 다음과 같이 나타낼 수 있다. 즉,

$$a_i(l) = \begin{cases} f(u_i(l)) & \text{if } l=1 \text{ or } l=3, \\ u_i(l) & \text{if } l=2 \text{ or } l=4. \end{cases} \quad (2)$$

$$u_i(l) = \sum_{j=1}^{N_{l-1}} w_{ij}(l) a_j(l-1) + \theta_i(l). \quad (3)$$

이다. N_l 은 계층 l 의 뉴런의 수이며, 0번째 계층의 활성화함수는 입력 $a_i(0) = x_i$ 로 정의한다. 이때 학습은 입력데이터의 값과 출력층의 뉴런 출력값 사이의 평균 자승 오차를 정의하여 이를 최소화하는 것이다. 평균 자승 오차 J_{NET} 을 수식으로 나타내면 다음과 같다. 즉,

$$J_{NET} = \frac{1}{NP} \sum_{i=1}^N \sum_{k=1}^P (a_{i,k}(4) - x_{i,k})^2 \quad (4)$$

이다. 여기서 n 은 출력층의 뉴런 수이고, P 는 학습 데이터 벡터 x_1, \dots, x_p 이며, $a_{i,k}(4)$ 는 계층 4에서 k 번째 패턴의 i 번째 뉴런의 활성화함수이다. J_{NET} 을 최소화하기 위한 가중치의 경신은 역전파 알고리즘을 이용하며, 그 가중치 경신 식은 다음과 같다. 즉,

$$w_{ij}(t+1) = w_{ij}(t) + \beta \delta_i(l) a_j(l-1) \quad (5)$$

이다. 여기서 β 는 학습률이다. 또한 식 (5)에서 $\delta_i(l)$ 의 값을 구하기 위한 계산은 다음 식을 이용하여 구할 수 있다. 즉,

$$\delta_i(4) = a_i(4) - x_i \quad (6)$$

$$\delta_i(l) = d_i(l) \sum_{j=1}^{N_{l+1}} \beta_j(l+1) w_{ij}(l+1), \quad l=1,2,3 \quad (7)$$

$$d_i(l) = \begin{cases} 1 & \text{if } l=2, \\ f'(u_i(l)) & \text{otherwise.} \end{cases}$$

이며, 여기서 f' 은 비선형 활성화 함수 f 의 미분이다.

이상에서의 NLPCA를 위한 신경망의 학습은 오차함수 J_{NET} 을 가중치 w_{ij} 에 대한 미분으로 기울기하강(gradient descent)법을 이용한 뉴런사이의 연결가중치를 구하는 과정이다. 그러나 이 기법에서는 식 (5)에서처럼 미분값만 고려되었으며, 과거의 속성을 반영하는 모멘트 항은 추가되지 않았다. 따라서 최소화과정에서 발생하는 발진을 없애기 위해 학습 알고리즘에 모멘트 항을 추가하면 모멘트가 가지는 우수한 속성을 그대로 얻을 수 있을 것이다. 즉, 식 (5)를 이용하여 가중치를 경신하면 전역 최소점에서의 수렴속도가 느려지고 또 수렴하는 과정에서 발진이 일어나는 경우도 발생한다. 이러한 단점을 보완하기 위해 모멘트 항을 추가하였다. 모멘트 항이 추가된 제안된 연결가중치 경신식은 다음과 같다. 즉,

$$w_{ij}(t+1) = w_{ij}(t) + \beta \delta_i(l) a_j(l-1) + \alpha \Delta w_{ij}(t-1) \quad (8)$$

이다. 여기서, α 는 모멘트(moment)이고, $\Delta w_{ij}(t-1)$ 는 이전의 연결가중치 변화량을 나타낸 것이다.

3. 시뮬레이션 결과 및 분석

제한된 학습알고리즘의 NLPCA를 위한 신경망의 성능을 평가하기 위해 5층과 3층으로 구성된 전향 신경망을 각각 구성하였다. 이 신경망에서 각 뉴런간의 초기 연결가중치는 랜덤시드(random seed)의 변화에 의해서 각각 -0.5에서 +0.5사이의 임의의 값으로 설정하였다. 학습은 전체 반복회수가 10,000회 이상이거나 전체 오차함수 값이 설정치 이하일 때 종료되도록 하였다.

제한된 학습알고리즘의 5층 전향신경망을 128×128 픽셀의 Lenna 영상과 256×128 픽셀의 차량번호판을 대상으로 시뮬레이션하여 그 타당성을 확인하였으며, 기존의 NLPCA를 위한 학습알고리즘의 5층 전향신경망과 기울기하강법에 기초를 둔 역전파 알고리즘을 이용한 3층 전향신경망에 의한 결과들과 비교하였다.

실험에서 추출된 특징들의 질을 평가하기 위해 다음의 PSNR(peak signal to noise ratio)을 이용하였다. 즉,

$$PSNR = 20 \cdot \log_{10}(255/RMSE) \quad (9)$$

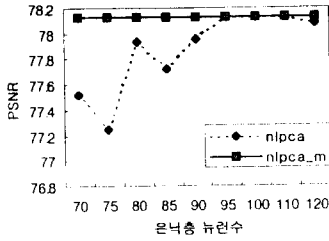
이다. 여기서 RMSE는 $(\sum (x_i - x_o)^2 / \text{전체픽셀수})^{1/2}$ 이다. x_i 는 입력 픽셀의 값이며, x_o 는 재구성된 픽셀의 값이다.

3.1 Lenna 영상

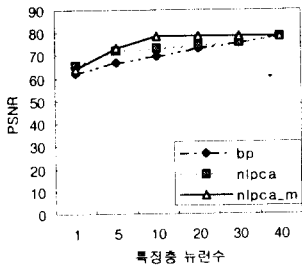
데이터의 주요 특징추출을 위해 128×128 픽셀의 Lenna 영상을 이용하였다. 실험에 이용된 영상은 한 개의 학습표본으로 64개 픽셀을 가지는 전체 256개의 표본을 연속으로 정렬하였다. 실험에서 전체 신경망의 입력뉴런과 출력뉴런 수는 각각 64개로 하였으며, 종료를 위한 허용 오차 함수 설정치는 10^{-3} 으로 하였다.

그림 2의 (a),(b)는 각각 은닉층과 특징층의 뉴런수에 따른 PSNR을 나타낸 것이다. 먼저 그림 2(a)에서는 기존의 NLPCA 경우 은닉층 개수에 따른 PSNR의 변화가 심하여 은닉층의 개수에 대한 성능의 의존도가 높음을 알 수 있다. 그러나 제안된 NLPCA의 경우, 상대적으로 거의 일정한 값을 가지며 은닉층의 뉴런수에 덜 민감함을 알 수 있다. 특히 기존의 NLPCA에서는 뉴런의 수가 110개 이상일 경우 PSNR의 값이 떨어져 뉴런수의 증가에 따른 학습 계산량만 증가함을 알 수 있다. 한편, 그림 2(b)에서는 특징층 뉴런의 변화에 따른 3가지 신경망의 성능을 보여준 것이다. 여기서도 제안된 NLPCA가 상대적으로 안정된 값을 가지며 다른 2가지 신경망에 비해 특징층 뉴런의 수에도 덜 의존함을 알 수 있다. 특히, 특징층 뉴런의 수가 10개부터는 거의 성능의 변화가 없음을 확인할 수 있다.

그림 3은 Lenna 영상을 대상으로 역전파 알고리즘, NLPCA, 제안된 알고리즘을 각각 이용한 오차값의 변화를 나타낸 것이다. 이때 학습률은 0.001로 은닉층 뉴런수는 110개, 특징층 뉴런수는 40개로 설정하였다. 그림에서



(a) 은닉층 개수에 따른 PSNR의 변화.



(b) 특징층 개수에 따른 PSNR의 변화.

그림 2. 은닉층과 특징층 개수에 따른 PSNR의 변화.

보면 제안된 알고리즘의 반복수가 100정도에서부터 오차값이 거의 일정하게 유지되며 가장 빠른 수렴속성을 가진다는 것을 알 수가 있다.

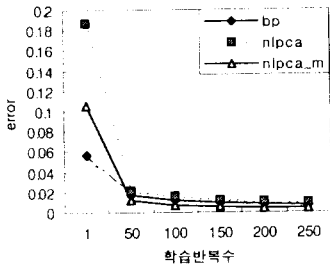


그림 3. 학습반복수에 따른 error변화.

표 1은 전체 오차 함수의 값이 10^{-3} 에 수렴할 때까지의 수렴 시간과 PSNR을 나타낸 것이다. 여기서 3가지 알고리즘 모두 다 최적해로 수렴하며 비슷한 PSNR값을 가진다. 그러나 최적해로의 수렴속도를 보면, 제안된 알고리즘은 역전파 알고리즘보다 약 2.4배, NLPCA를 이용한 신경망보다 약 7.8배 정도 빠르다는 것을 알 수 있다. 이는 제안된 NLPCA가 학습속도 측면에서 더욱 우수한 성능이 있음을 나타낸 것이다.

표 1. 전체 오차값이 10^{-3} 에 수렴하도록 학습시킨 결과.

	time	PSNR
BP알고리즘	2228	78.130806
NLPCA를 이용한 신경망	7259	78.130898
제안된 알고리즘	933	78.131691

그림 4는 Lenna 원영상과 3가지 알고리즘의 방법으로 복원된

영상을 나타내고 있다. (a)는 원영상이고 (b), (c), (d)는 각각 학습반복수 1000번에 대해 40개의 주요특징추출 후, 이들을 이용하여 재구성된 영상을 보여준다. 그림에서 보면, 역전파 알고리즘에 의해 재구성된 결과인 영상 (b)가 상대적으로 NLPCA를 이용한 (c),(d)영상보다 심한 왜곡이 있음을 알 수 있다. 이는 역전파에 이용한 신경망은 3층의 구조로 비선형 속성을 추출된 특징에 충분히 반영되지 않았기 때문이다. 또한 영상 (d)는 우수한 비선형 사상의 속성을 가진 특징들을 추출할 수 있음을 확인할 수 있다.

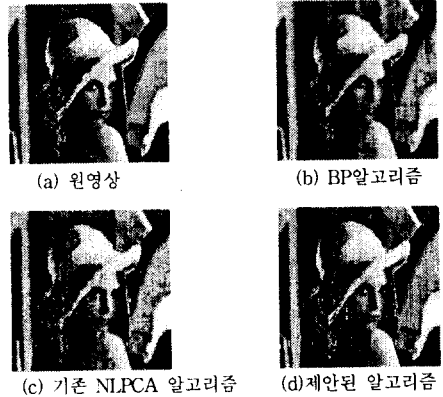


그림 4. 3가지 학습알고리즘에 따른 복원된 128×128 픽셀의 Lenna 영상들.

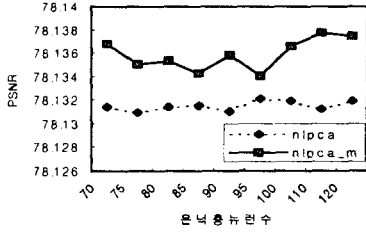
3.2 차량변호판

128×128 픽셀의 Lenna 영상에 비해 큰 규모의 영상으로 256×128 픽셀의 차량번호판을 이용하여 실험하였다. Lenna 영상의 실험에서처럼 64픽셀을 한 개의 학습표본으로 하는 전체 512개의 표본을 가지도록 연속으로 정렬하였다. 실험에서도 입력뉴런과 출력뉴런 수는 각각 64 개로, 허용 오차 함수를 10^{-3} 으로 하였다. 그림 5의 (a),(b)는 그림 2의 (a),(b)에서처럼 각각 은닉층과 특징층의 뉴런수에 따른 PSNR을 나타낸 것이다. 먼저 그림 5 (a)에서 제안된 NLPCA는 은닉층 개수에 따른 PSNR의 변화가 심하나 PSNR의 값이 기존의 NLPCA보다 상대적으로 높음을 알 수 있고, 또 뉴런의 수가 110개 이상일 경우 PSNR의 값이 안정되어 진다는 것을 알 수 있다. 한편, 그림 5(b)에서는 특징층 뉴런의 변화에 따른 3가지 신경망의 성능을 보여준 것이다. 여기서도 제안된 NLPCA가 상대적으로 안정된 값을 가지며 다른 2가지 신경망에 비해 특징층 뉴런의 수에도 덜 의존함을 알 수 있다.

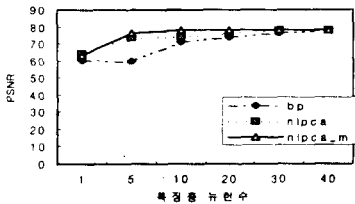
표 2는 허용 오차 함수가 10^{-3} 에 수렴할 때 수렴시간과 PSNR을 나타낸 것이다. 여기서도 그 최적해로의 수렴속도를 보면, 제안된 알고리즘이 역전파 알고리즘보다 약 4.4배 NLPCA를 이용한 신경망보다 약 48배정도 빠르다는 것을 알 수 있다. 이는 제안된 NLPCA의 학습속도가 다른 방법에 비해 더욱 우수한 성능이 있음을 재확인할 수 있다. 또한 표 1과 비교하여 볼 때, 최적해로의 수렴속도와 PSNR에서 문제의 규모가 커질수록 제안된 학습 알고리즘의 성능 개선의 결과가 더 뚜렷한 것을 알 수 있다. 이는 문제의 규모가 증가할수록 초기 연결가중치의 설정은 더욱 어려워지며 이에 따른 학습성능의 의존도도 더욱 증가한다는 것을 알 수 있다.

그림 6은 차량번호판을 대상으로 3가지 학습알고리즘 각각의 오차값의 변화를 나타낸 것이다. 여기서도 학습률과 은닉층 및 특징층 뉴런수는 각각 0.001과 110개 및 40개로 하였다. 그림에서 보면 제안된 알고리즘의 반복수가 100정도에서부터 오차값이 거의 일정하게 유지되며, 그림 3에서처럼 가장 빠른 수렴속성이 있음을 확인할 수 있다.

그림 7은 차량 번호판의 원영상과 3가지 알고리즘의 방법으로 복원된 영상을 나타내고 있다. (a)는 원영상이고, (b), (c), (d)는



(a) 은닉층 개수에 따른 PSNR의 변화.



(b) 특징층 개수에 따른 PSNR의 변화.

그림 5. 은닉층과 특징층 개수에 따른 PSNR의 변화.

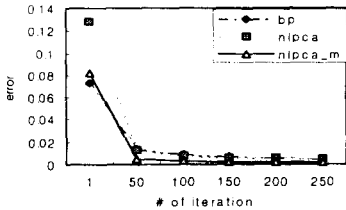


그림 6. 학습 알고리즘들에 대한 error 변화.

표 2. 전체 오차값이 10^{-3} 에 수렴하도록 학습시킨 결과.

	time	PSNR
BP알고리즘	3176	78.130913
NLPCA를 이용한 알고리즘	3406	78.131935
제안된 알고리즘	717	78.136589

각각 학습반복수 300번에 대해 40개의 주요특징추출 후, 이들을 이용하여 재구성된 영상을 보여준다. 그림에서도 보면, 역전파 알고리즘에 의해 재구성된 결과인 영상 (b)가 상대적으로 NLPCA를 이용한 (c),(d)영상보다 심한 왜곡이 있음을 알 수 있다. 또한 제안된 NLPCA는 기존의 NLPCA보다 더욱 우수한 비선형 속성을 가진 특징들을 추출할 수 있음을 알 수 있다. 이는 제안된 NLPCA가 인식이나 분류 등에서 좀 더 우수한 특성이 있음을 나타내는 것이다. 결국 제안된 학습알고리즘의 5층 신경망을 이용하여 효과적으로 입력영상의 주요 특징들을 추출할 수 있으며, 또한 영상의 압축효과도 우수한 결과를 얻을 수 있다.



(a) 원영상



(b) BP 알고리즘



(c) 기존 NLPCA 알고리즘



(d) 제안된 알고리즘

그림 7. 학습알고리즘에 따른 복원된 256×128 픽셀의 차량 번호판

4. 결 론

본 논문에서는 모멘트를 추가한 새로운 학습 알고리즘의 5층 전향 신경망을 이용한 입력 데이터의 비선형 주요 특징 추출에 대해서 제안하였다. 제안된 학습알고리즘은 신경망의 학습시에 과거의 속성을 반영하기 위한 모멘트 항이 추가된 학습기법이다. 이는 최적해로의 수렴에 따른 발진을 억제하여 그 수렴속도를 좀 더 개선시키는 모멘텀의 장점을 그대로 살리기 위함이다.

제안된 학습알고리즘을 이용한 신경망을 128×128 픽셀의 Lenna와 256×128 픽셀의 차량 번호판 영상들을 대상으로 시뮬레이션 한 결과, 기존의 역전파 알고리즘을 이용한 3층 신경망이나 NLPCA를 이용한 5층 신경망보다 더욱 우수한 수렴성능과 특징 추출성능이 있음을 확인할 수 있었다.

향후 좀더 빠른 최적화 기법을 이용한 학습 알고리즘에 대한 연구가 계속 진행되어야 하며, 영상 인식 등과 같은 다양한 분야의 적용에 대한 연구도 남아 있다.

5. 참고문헌

- [1] K. I. Diamantaras and S. Y. Kung, 'Principal Component Neural Networks : Theory and Applications, Adaptive and Learning Systems for Signal Processing, Communication, and Control,' John Wiley & Sons, Inc., New York, 1996.
- [2] 조용현, "기울기하강과 동적터널링에 기반을 둔 학습알고리즘의 신경망을 이용한 영상데이터의 주요특징추출," 한국정보처리학회 논문지, 제 6권, 제 4호, pp.1393-1402, 1999.
- [3] M. A. Kramer, "Nonlinear Principal Component Analysis Using Autoassociative Neural Networks," AIChE Journal, Vol.37, No.2, pp.233-243, Feb. 1991.
- [4] E. C. Malthouse, "Limitations of Nonlinear PCA as Performed with Generic Neural Networks," IEEE Trans. on Neural Networks, Vol.9, No.1, pp.165-173, Jan. 1998.
- [5] D. Fotheringham and R. Baddeley, "Nonlinear principal components analysis of neural spike train data," Biol. Cybern. 77, pp.283-288, 1997.
- [6] B. Schölkopf, A. Smola, and K. R. Müller, "Nonlinear Component Analysis as a Kernel Eigenvalue Problem," Neural Computation, Technical Report 44, Max-Planck-Institut für biologische Kybernetik, 1996.
- [7] K. Mantere, J. Parkkinen, and T. Jaaskelainen, "Simulation of white-light adaptation characteristics with use of nonlinear neural principal component analysis," Optical Society of America, Vol.14, No.9, pp.2049-2056, Sept. 1997.
- [8] D. Tzovaras and M. G. Strintzis, "Use of Nonlinear Principal Component Analysis and Vector Quantization for Image Coding," IEEE Trans. on Image Processing, Vol.7, No.8, pp.1218-1223, Aug. 1998.
- [9] F. D. Frate and G. Schiavon, "Nonlinear Principal Component Analysis for the Radiometric Inversion of Atmospheric Profiles by Using Neural Networks," IEEE Trans. on Geoscience and Remote Sensing, Vol.37, No.5, pp.2335-2342, Sept. 1999.
- [10] N. Qian, "On the momentum term in gradient descent learning algorithms," Neural Networks 12, pp.145-151, 1999.