

# 모멘트를 이용한 비선형 주요성분분석 신경망의 효율적인 학습알고리즘

## An efficient learning algorithm of nonlinear PCA neural networks using momentum

조용현\*  
Yong-Hyun Cho

### < Abstract >

This paper proposes an efficient feature extraction of the image data using nonlinear principal component analysis neural networks of a new learning algorithm. The proposed method is a learning algorithm with momentum for reflecting the past trends. It is to get the better performance by restraining an oscillation due to converge the global optimum. The proposed algorithm has been applied to the cancer image of  $256 \times 256$  pixels and the coin image of  $128 \times 128$  pixels respectively. The simulation results show that the proposed algorithm has better performances of the convergence and the nonlinear feature extraction, in comparison with those using the backpropagation and the conventional nonlinear PCA neural networks.

**Key Words :** *Nonlinear principal component analysis(NLPCA) neural networks, Feature extraction, Momentum, Backpropagation neural network*

### 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 있으며, 이는 특징 추출과 밀접한 관계가 있다. 특징추출은 인식이

나 분류와 같은 문제에서 데이터의 사용을 용이하게 하기 위해 속성이나 두르려진 형상을 식별하기 위한 것이다. 이러한 데이터의 특징들을 추출하기 위한 많은 기법들이 제시되어 왔다<sup>[1-2]</sup>.

주요성분분석(principal component analysis: PCA)은 주어진 입력데이터의 특징을 추출하는 기법으로 이용되고 있지만, 이는 입력데이터에 대한 공분산행렬(covariance matrix)과 고유치

\* 정회원 대구가톨릭대학교 공과대학  
컴퓨터정보통신공학부, 교수, 工博  
712-702 경북 경산시 하양읍 금락리 330

Professor, School of Computer & Information Comm.  
Eng., Catholic Univ. of Daegu  
E-mail : yhcho@cuth.cataegu.ac.kr



(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 신경망에 기반을 둔 PCA가 이용되지만, 이 방법은 입력 데이터의 분포가 선형적인 관계에 있을 때는 효율적으로 데이터의 특징을 추출하나 비선형적인 관계가 존재할 경우 특징을 추출하는데는 매우 비효율적이다<sup>[3-6]</sup>. 이러한 제약을 해결하기 위해서 비선형 주요성분분석(nonlinear principal component analysis : NLPCA) 기법이 제안되었다<sup>[3-9]</sup>. 그중에서 Kramer는 비선형 특징들을 추출하기 위하여 5 계층 구조의 신경망(neural network : NN)을 제안<sup>[3]</sup>하였으며, 이 기법을 기반으로 한 신경망이 여러 분야에 응용되어 왔다. Mantere 등<sup>[7]</sup>은 광학분야에서 반대되는 색의 신호와 무색 신호의 백색광 적응 특성을 시뮬레이션하고 근사화하는데 응용하였으며, Tzovaras 등<sup>[8]</sup>은 영상처리분야에서 정지영상을 사용하여 부호화하고 복호화하는데 응용하였고, Frate 등<sup>[9]</sup>은 지구과학분야에서 기온과 증기의 대기 중 측면의 방사선 반전을 위해 응용하였다. 그러나 Kramer의 NLPCA에 기반한 이러한 방법들은 비선형 속성의 추출을 위한 신경망을 학습시키는데 매우 오랜 시간이 걸린다는 제약이 있다<sup>[4-9]</sup>. 이러한 제약을 개선한다면 NLPCA가 가지는 우수한 성능을 좀 더 효율적으로 이용할 수 있을 것이다.

본 논문에서는 모멘트(momentum)를 추가한 학습 알고리즘을 제안하여 NLPCA를 위한 신경망을 학습시킴으로써 입력데이터의 주요 특징을 보다 빠르고 정확하게 추출되도록 하였다. 제안된 학습 알고리즘을 이용한 신경망을 256\*256 픽셀의 간암영상과 128\*128 픽셀의 동전영상들을 대상으로 시뮬레이션하여 역전파 알고리즘 및 기존 NLPCA의 방법을 이용한 신경망에 의한 결과와 비교 고찰하였다.

2. 비선형 주요성분 분석을 위한 신경망의 개선된 학습알고리즘

NLPCA는 PCA의 방법과 같은 수단으로 다변량 데이터의 분석을 위한 새로운 기법이다. PCA에서와 마찬가지로 NLPCA는 실험 데이터의 분석 및 가시화, 그리고 차원 축소를 목적으로 함으로서 변수들간의 상호관계를 제거하고 식별하기 위해 사용된다. PCA는 단순히 변수들 간의 선형적인 상호관계만을 식별하지만, NLPCA는 데이터들 속에 나타나는 특성들을 선형적인 관계뿐만 아니라 비선형적인 상호관계도 구분할 수 있다. NLPCA에서는 입력층의 입력 데이터가 출력층에 재구성되는 동일 사상을 수행하도록 신경망을 학습시킨다. 이는 실제 입력 데이터의 분포에 닮은 특징 공간을 생성함으로써 차원을 축소시키기 위함이다.

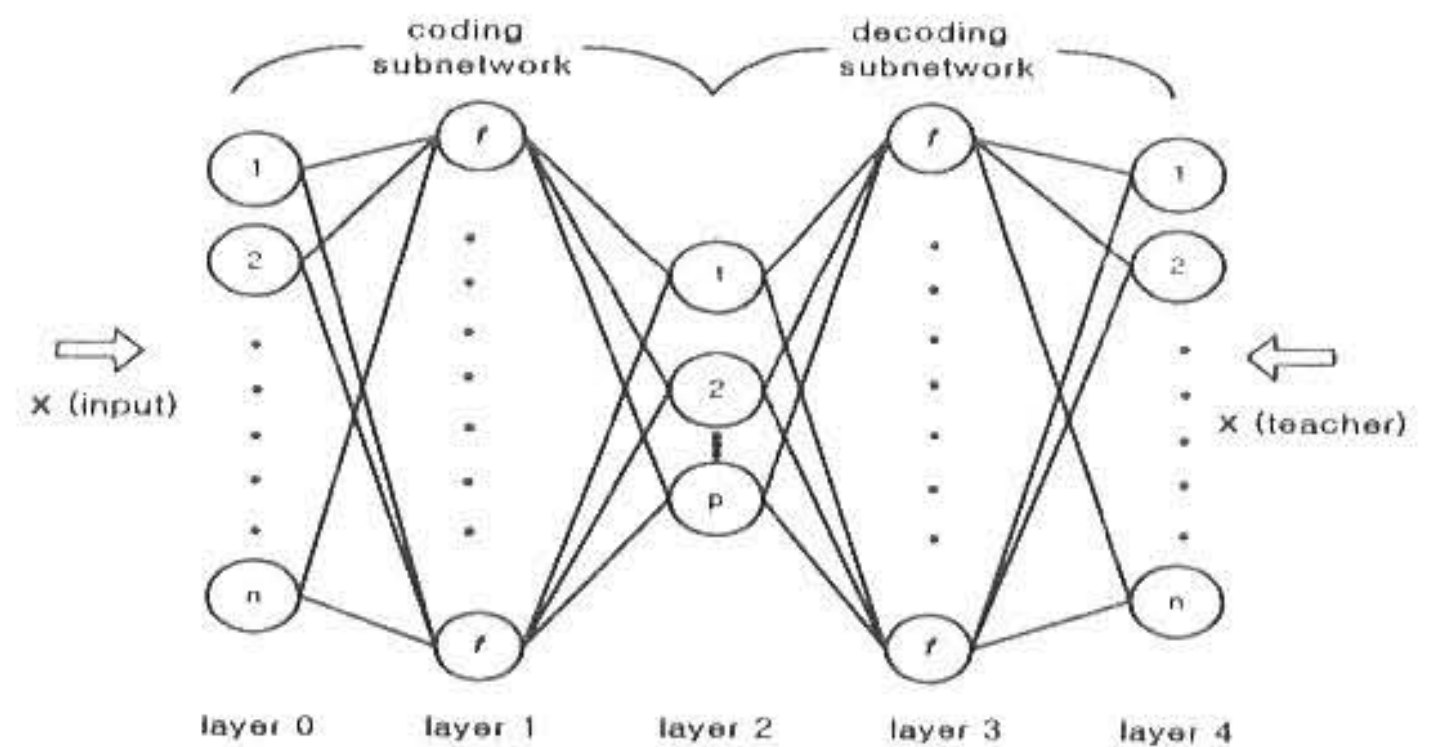


Fig.1 Architecture of NLPCA neural network

Fig.1은 NLPCA를 위한 일반적인 신경망의 구조를 나타낸 것이다. 그림에서 계층 1과 계층 2로 이루어진 부분망은 입력데이터를 특징공간으로 사상하는 부호화망이다. 여기서 특징층(feature-layer), 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-자형(sigmoidal) 함수  $f(x) = 1/(1 + e^{-x})$ 이다. 나머지 계층들은 선형 함수나 비선형 함수가 이용될 수 있다.

NLPCA를 위한 신경망의 입·출력 함수관계를 나타내면 다음과 같다. 즉,

$$\alpha(x) = \sum_{i=1}^n \overline{w}_i f(\underline{w}_i x + \underline{\theta}_i) + \overline{\theta} \quad (1)$$



이다. 여기서  $\bar{w}_i$ 와  $\bar{\theta}$  및  $\underline{w}_i$ 와  $\underline{\theta}_i$ 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)도 근사화 할 수 있다.

$a_i(l)$ 가 계층  $l$ 의  $i$ 번째 뉴런의 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, 0 번째 계층의 활성화함수  $a_i(0) = x_i$ 로 정의된다. 이때 학습은 입력데이터의 값과 출력층의 뉴런 출력값 사이의 평균 자승 오차를 정의하여 이를 최소화하는 것이다. 평균 자승오차  $J_{NET}$ 을 수식으로 나타내면 다음과 같다.

즉,

$$J_{NET} = \frac{1}{np} \sum_{k=1}^p \sum_{i=1}^n (a_{i,k}(4) - x_{i,k})^2 \quad (4)$$

이다. 여기서  $a_{i,k}(4)$ 는 계층 4에서  $k$  번째 패턴의  $i$  번째 뉴런의 활성화함수이다.  $J_{NET}$ 을 최소화하기 위한 가중치의 갱신은 역전파 알고리즘을 이용하며, 그 가중치 갱신식은 다음과 같다.

즉,

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

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

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

$$\delta_i(l) = d_i(l) \sum_{j=1}^{n_{l+1}} \delta_j(l+1) w_{ji}(l+1), \quad l=1, 2, 3$$

$$d_i(l) = \begin{cases} 1 & \text{if } l=2 \\ f'(u_i(l)) & \text{otherwise.} \end{cases} \quad (7)$$

이며, 여기서  $f'$ 은 비선형 활성화함수  $f$ 의 미분이다.

이상에서의 NLPCA를 위한 신경망의 학습은 오차함수  $J_{NET}$ 을 가중치  $w_{ij}$ 에 대해 미분하는 것으로 기울기하강(gradient descent)법을 이용하여 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 속성을 그대로 얻을 수 있을 것이다. 즉, 식 (5)를 이용하여 가중치를 갱신하면 전역최소점으로서의 수렴속도가 느려지고 또 수렴하는 과정에서 발진이 일어나는 경우도 발생한다.

이러한 단점을 보완하기 위해 모멘트 항이 추가된 수정된 갱신식을 제안하였다. 따라서 모멘트 항이 추가된 연결가중치 갱신식은 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,  $\alpha$ 는 모멘트이고,  $\Delta w_{ij}(t-1)$ 는 이전의 연결가중치 변화량이다.

### 3. 시뮬레이션 결과 및 분석

제안된 학습알고리즘의 NLPCA 신경망의 성능을 평가하기 위해 5 층과 3 층으로 구성된 전향신경망을 각각 구성하였다. 이 신경망에서 각 뉴런간의 초기 연결가중치는 랜덤시드(random seed)의 변화에 의해서 각각 -0.5에서 +0.5사이의 임의의 값으로 설정하였다. 팬티엄 III-700 시스템 환경하에서 학습은 전체 반복회수가 20,000 회 이상이거나 전체 오차함수 값이 설정치  $10^{-3}$  이하일 때 종료되도록 하였다.



제안된 학습알고리즘의 5층 전향신경망을 256×256 픽셀의 간암영상과 128×128 픽셀의 동전영상을 대상으로 시뮬레이션하고 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,

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

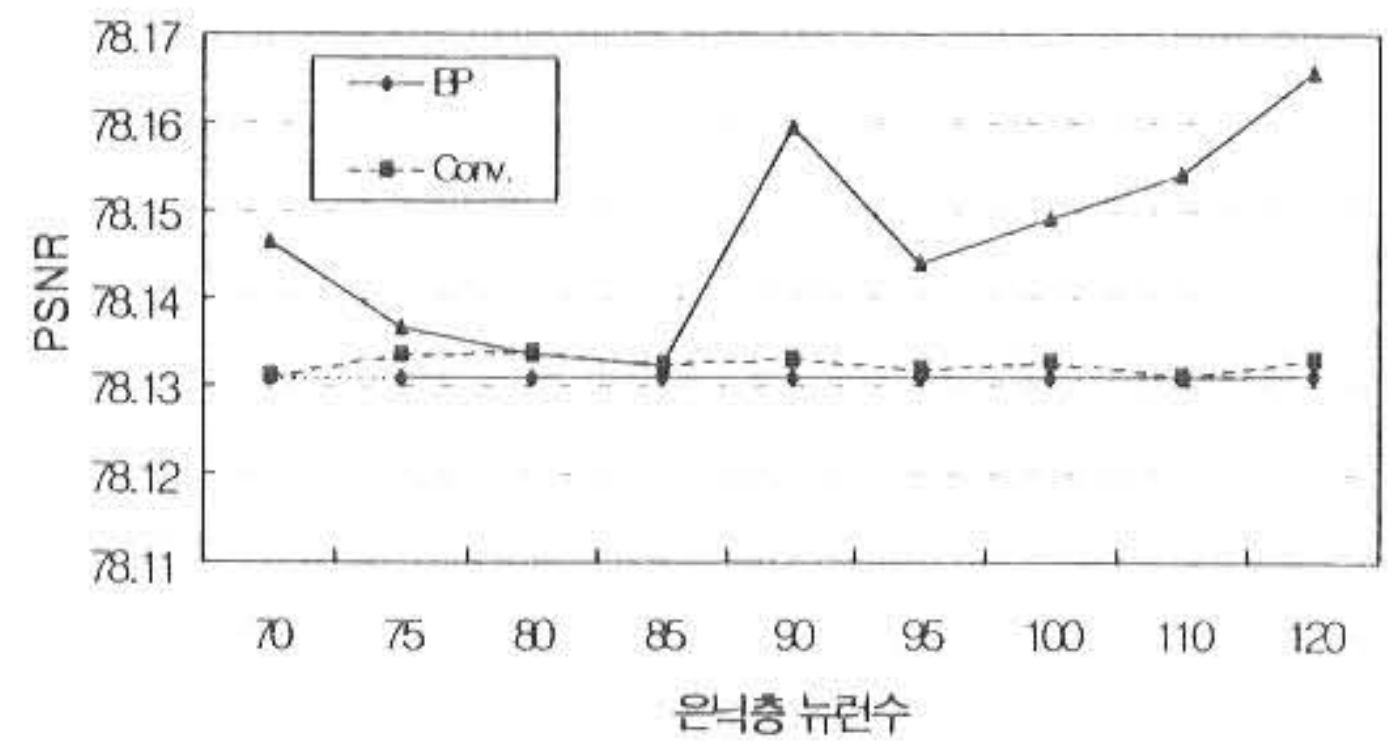
이다. 이때 RMSE(root mean square error)는  $(\sum(x_i - x_o)^2/n^2)^{1/2}$ 이다.  $x_i$ 는 입력 픽셀의 값이며  $x_o$ 는 재구성된 픽셀의 값이고,  $n$ 은 영상의 전체 픽셀 수이다.

### 3.1 CT로 촬영된 간암영상

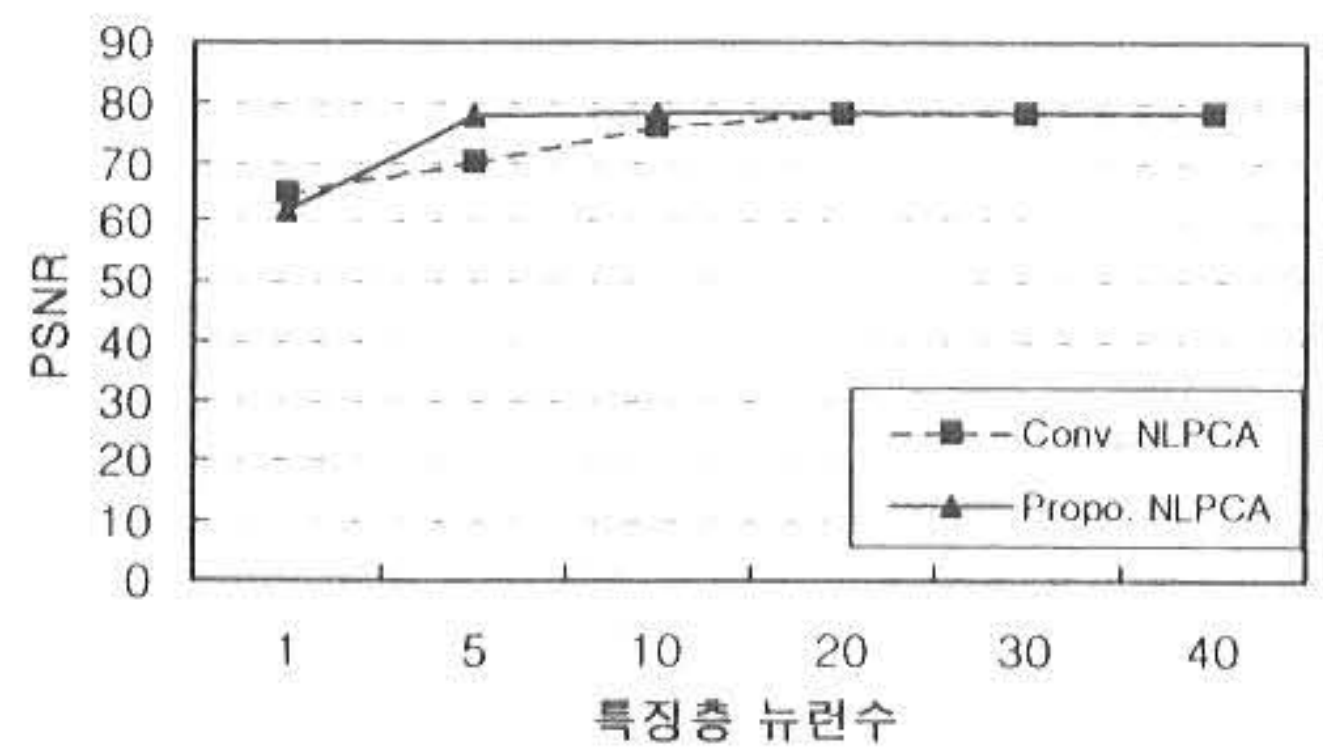
데이터의 주요 특징추출을 위해 256×256 픽셀의 간암영상을 이용하였다. 실험에 이용된 영상은 한 개의 학습표본으로 64 개 픽셀을 가지는 전체 1024 개의 표본을 연속으로 정렬하였다. 실험에서 전체 신경망의 입력뉴런과 출력뉴런 수는 각각 64 개로 하였으며, 종료를 위한 허용 오차 함수의 설정치는  $10^{-3}$ 으로 하였다.

Fig. 2의 (a)와 (b)는 각각 은닉층과 특징층의 뉴런수에 따른 PSNR값의 변화를 나타낸 것이다. 실험에서 모멘트는 0.9로 하였다. 먼저 Fig. 2(a)에서 기존의 NLPCA와 역전파 알고리즘의 경우는 유사한 값을 가지며, 은닉층의 뉴런수에 따른 PSNR의 변화도 거의 없음을 알 수 있다. 그러나 제안된 NLPCA의 경우는 다른 2 가지 알고리즘보다 높은 PSNR값을 가져 우수한 특징 추출능력이 있음을 알 수 있다. 한편, Fig 2(b)에서는 특징층 뉴런수의 변화에 따른 기존의 NLPCA와 제안된 NLPCA사이의 PSNR 값을 보여준 것이다. 여기에서도 제안된 NLPCA가 기존의 NLPCA보다 안정된 값을 가지며 특징층 뉴런수에 덜 의존함을 알 수 있다. 특히, 제안된 알고리즘의 경우 특징층 뉴런의

수가 5 개부터는 거의 성능의 변화가 없음을 확인할 수 있다.



(a) Hidden-layer



(b) Feature-layer

Fig.2 PSNR characteristics by the hidden-layer and the feature-layer neurons

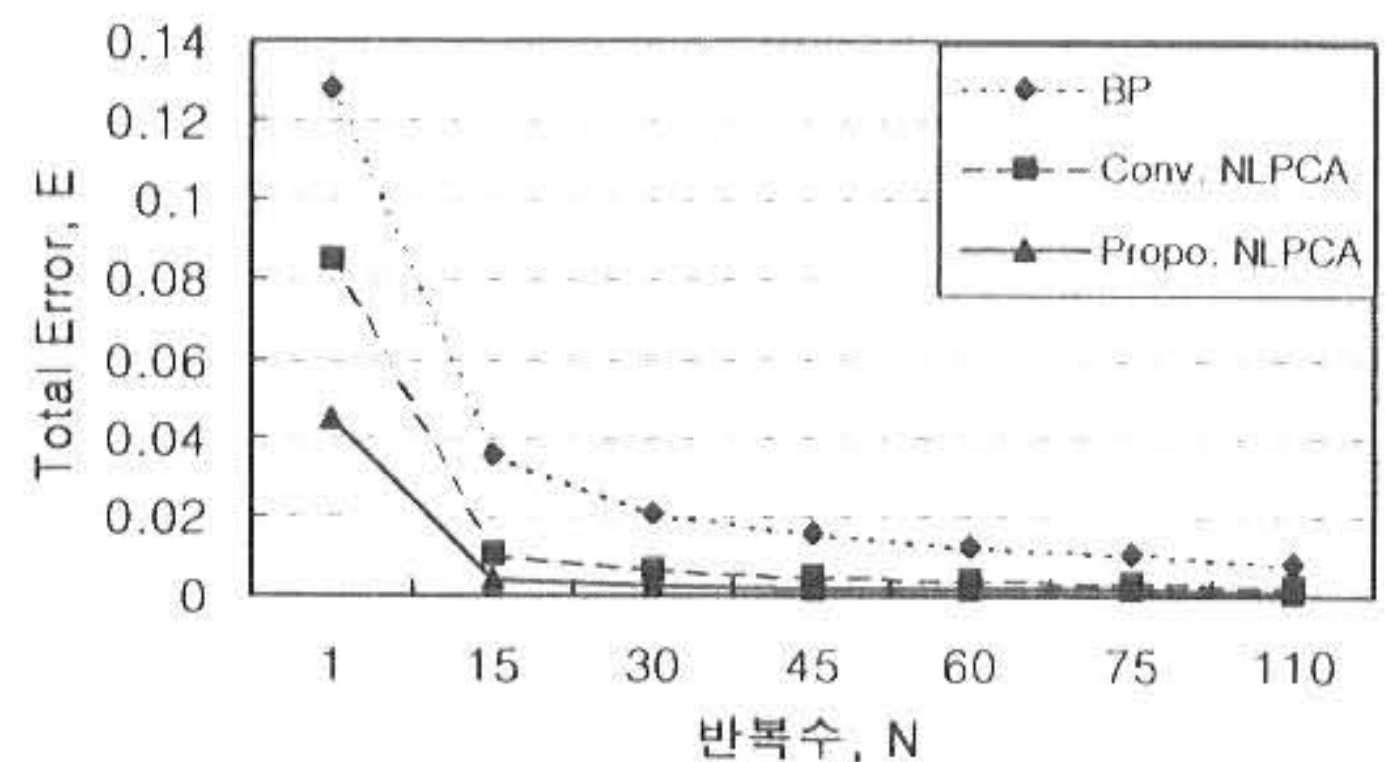


Fig. 3 Error characteristics by iterations

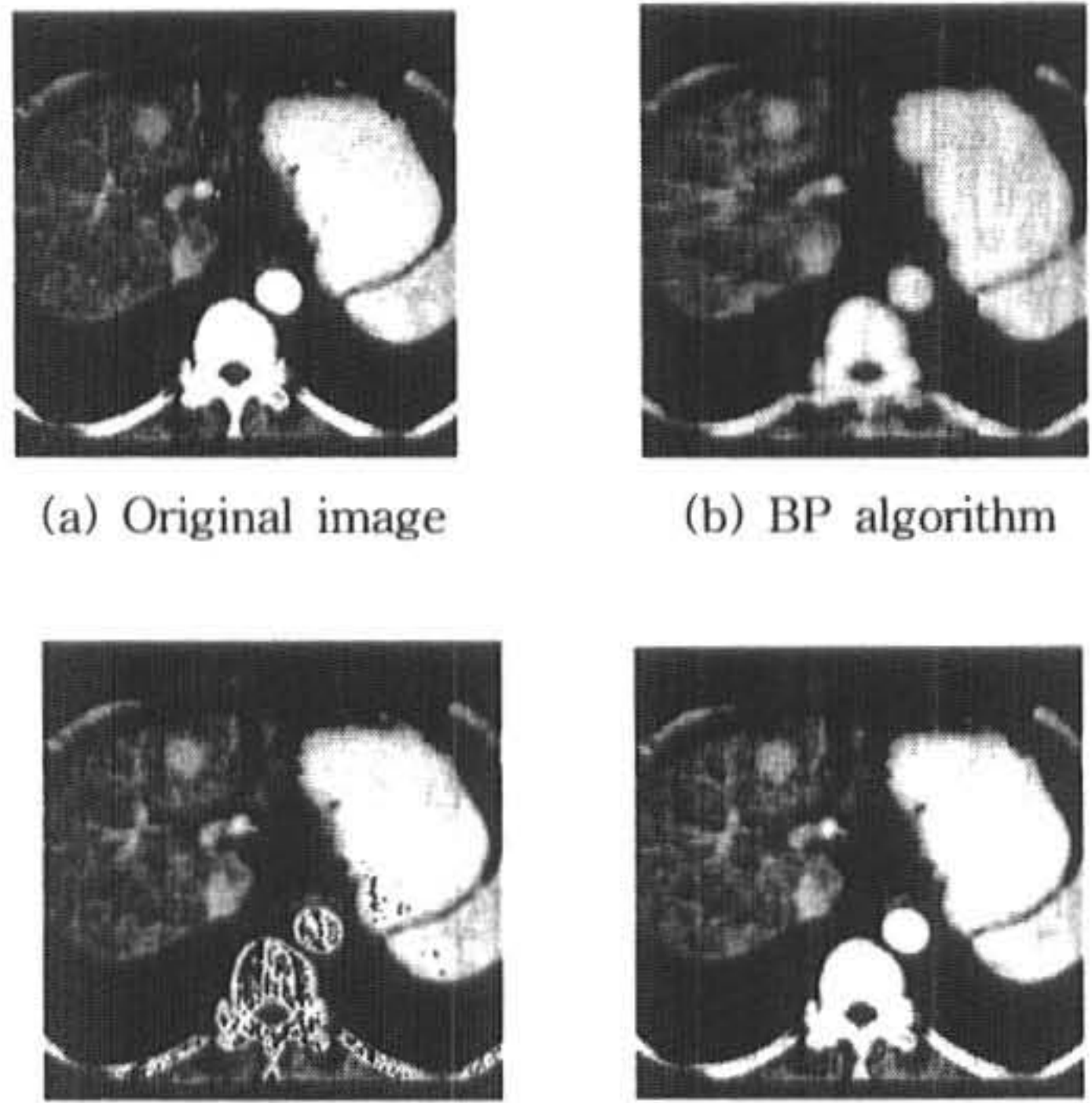
Fig. 3은 간암영상을 대상으로 하여 역전파 알고리즘, 기존 NLPCA, 제안된 NLPCA를 각각 이용한 오차값의 변화를 나타낸 것이다. 이때 학습률은 0.001로 은닉층 뉴런수는 110 개, 특징층 뉴런수는 40개로 설정하였다. Fig



에서 제안된 알고리즘은 반복수가 15 회 정도 부터 오차값이 거의 일정하게 유지되며 다른 2 가지 알고리즘에 보다 빠른 수렴속성을 가진다는 것을 알 수가 있다.

Table 1은 전체 오차값이  $10^{-3}$ 에 수렴할 때 까지 수렴시간 T와 PSNR값을 나타낸 것이다. 여기서 3 가지 알고리즘 모두 다 최적해로 수렴을 하지만 최적해로의 수렴속도를 보면, 제안된 NLPCA는 역전파 알고리즘보다 약 39 배정도, 기존의 NLPCA를 이용한 신경망보다는 약 5 배정도 빠르다는 것을 알 수 있다. 이는 제안된 NLPCA가 학습속도 측면에서 더욱 우수한 성능이 있음을 나타낸 것이다.

Fig.4는 간암의 원 영상과 3 가지 알고리



(a) Original image (b) BP algorithm (c) Conventional NLPCA (d) Proposed NLPCA

Fig. 4 Reconstructed liver-cancer images by 3 algorithms

Table 1. Simulation results in Total error= $10^{-3}$

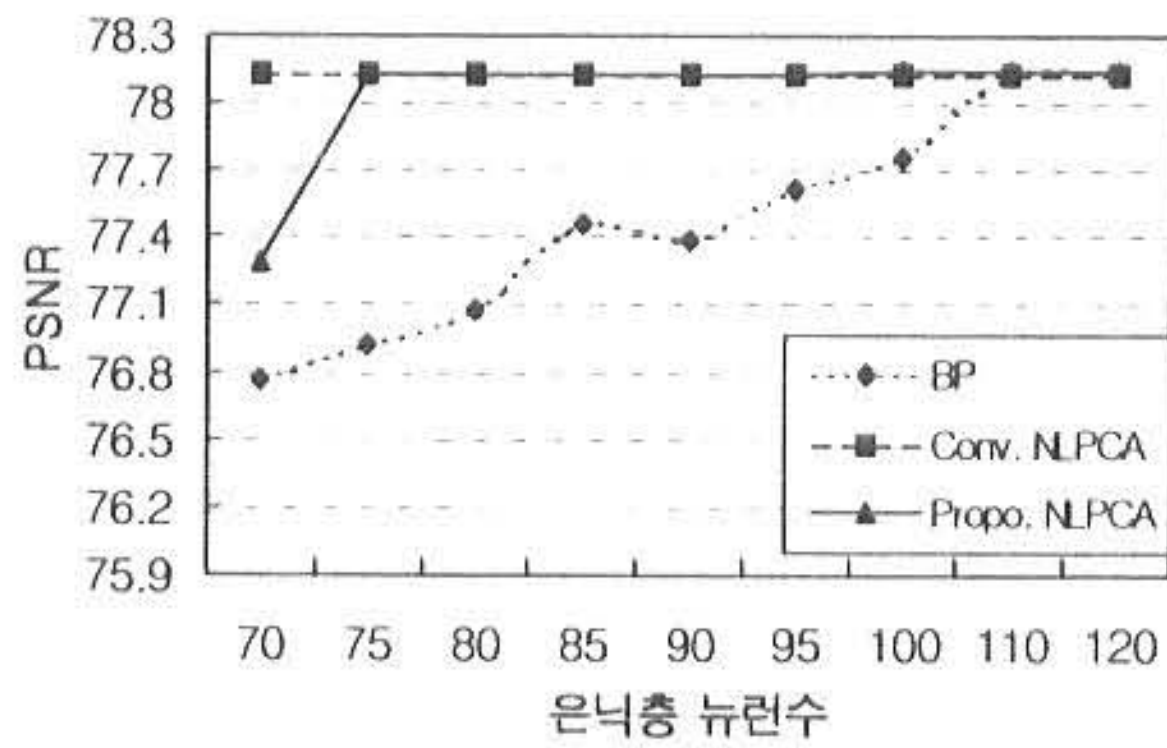
	T	PSNR
BP 알고리즘	15530	78.130806
기존의 NLPCA	1961	78.131195
제안된 NLPCA	397	78.153900

즘의 방법으로 복원된 영상을 나타내고 있다. (a)는 원 영상이고, (b), (c), (d)는 각각 학습 반복수 300 번에 대해 40 개의 주요특징추출 후, 이들을 이용하여 복원된 영상을 보여준다. Fig에서 보면, 역전파 알고리즘에 의해 복원된 영상 (b)가 상대적으로 NLPCA를 이용한 (c)와 (d)영상보다 심한 왜곡이 있음을 알 수 있다. 이는 역전파 알고리즘을 이용한 신경망은 3 층의 구조로 입력층과 특징을 추출하는 은닉층 사이가 단층이므로 입력영상 내의 비선형 속성을 충분히 추출하지 못하기 때문이다. 한편 영상(c)는 영상의 비선형 속성을 추출할 수 있는 기존의 NLPCA 방법으로 영상 (d)에 비해 잡음이 많음을 알 수 있다. 하지만 제안된 학습알고리즘의 NLPCA 신경망을 이용한 영상 (d)는 역전파 알고리즘이나 기존의 NLPCA를 이용한 신경망에 비해 우수한 결과 영상을 나타내어 데이터의 비선형 속성을 효과적으로 추출할 수 있음을 확인할 수 있다.

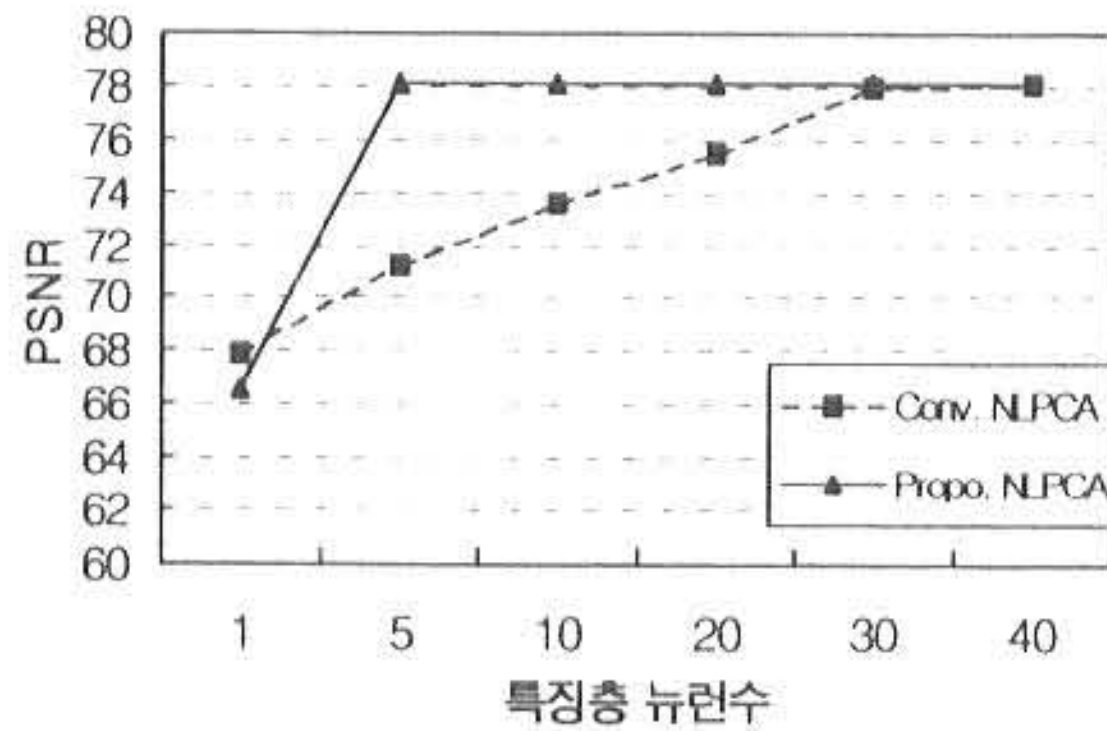
### 3.2 동전영상

256×256 픽셀의 간암영상에 비해 규모가 작은 128×128 픽셀의 동전영상을 이용하여 실험하였다. 간암영상의 실험과 마찬가지로 64 픽셀을 한 개의 학습표본으로 하는 전체 256 개의 표본을 가지도록 연속으로 정렬하였으며, 입력뉴런과 출력뉴런 수는 각각 64 개로 모멘트는 0.9로 설정하였고 허용 오차함수의 설정치는  $10^{-3}$ 으로 하였다. Fig.5의 (a)와 (b)는 Fig.2의 (a)와 (b)에서처럼 각각 은닉층과 특징층의 뉴런수에 따른 PSNR값의 변화를 나타낸 것이다. 먼저 Fig.5(a)에서 기존의 NLPCA와 제안된 NLPCA는 은닉층 뉴런수에 따른 PSNR의 변화가 거의 일정하지만 역전파 알고리즘은 은닉층의 뉴런수가 증가함에 따라 PSNR의 값이 증가됨을 볼 수 있다. 한편, Fig.5(b)에서는 특징층 뉴런의 변화에 따른 2 가지 신경망의 성능을 보여주고 있으며, 여기서도 Fig.2(b)에서와 마찬가지로 특징층 뉴런수가 5 개부터는 제안된 NLPCA 신경망이 기존의 NLPCA 신경망보다 안정된 PSNR값을 가지며 특징층 뉴런의 수에도 덜 의존함을 알 수 있다.





(a) Hidden-layer



(b) Feature-layer

Fig. 5 PSNR characteristics by the hidden-layer and the feature-layer neurons

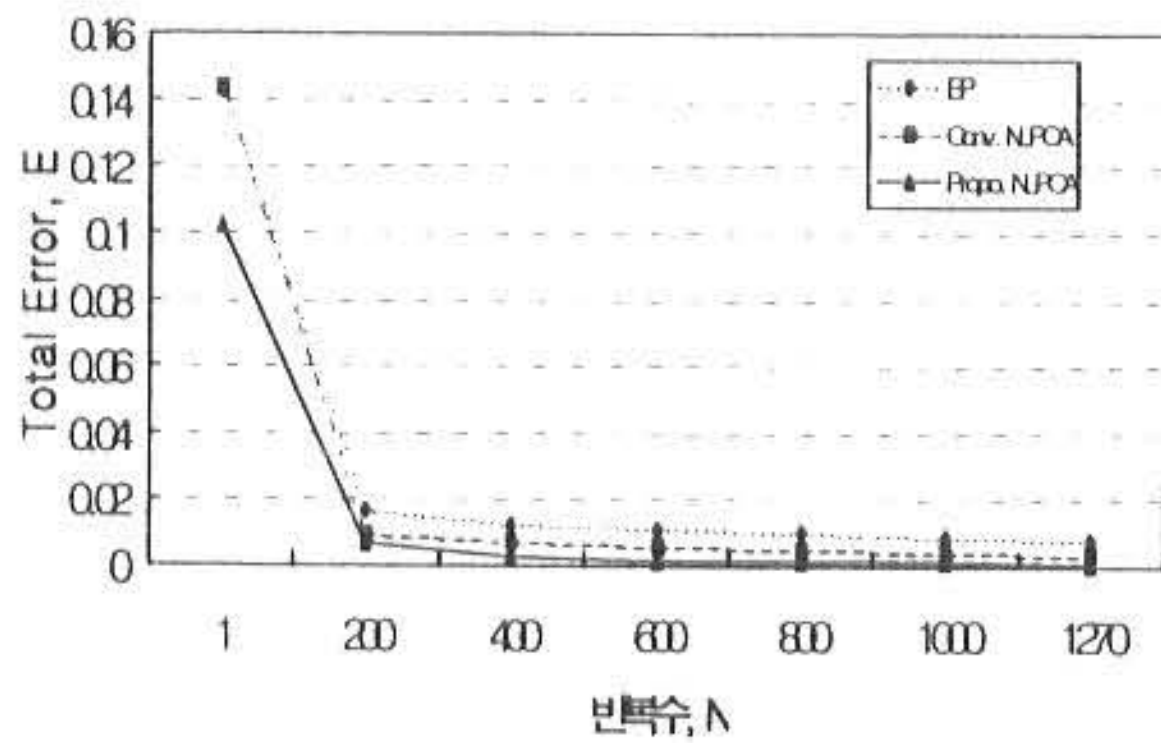


Fig. 6 Error characteristics by iterations

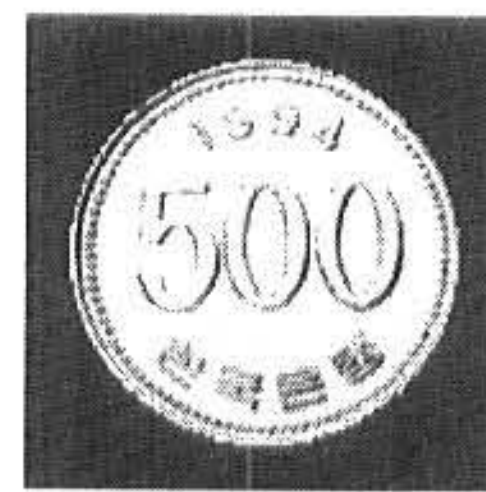
Fig.6은 동전영상을 대상으로 3 가지 학습 알고리즘의 오차값 변화를 나타낸 것이다. 여기서도 학습률과 은닉층 및 특징층 뉴런수는 각각 0.001과 110 개 및 40 개로 하였다. Fig 에서 보면 제안된 알고리즘의 반복수가 600 정도에서부터 오차값이 거의 일정하게 유지되고 Fig.3에서처럼 가장 빠른 수렴속성이 있음을 확인할 수 있다.

Table 2는 허용 오차함수가  $10^{-3}$ 에 수렴할 때 수렴시간 T와 PSNR의 값을 나타낸 것이다. 여기에서도 그 최적해로의 수렴속도를 보면, 제안된 알고리즘이 역전파 알고리즘보다 약 13 배정도 빠르고 NLPCA를 이용한 신경망보다 약 6 배정도 빠르다는 것을 알 수 있다. 이는 제안된 NLPCA의 학습속도가 다른 방법에 비해 더욱 개선됨을 재확인할 수 있다.

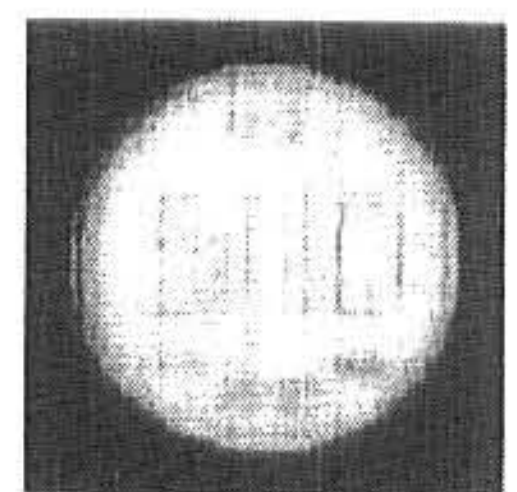
따라서 최적해로의 수렴속도 측면에서 제안된 학습 알고리즘의 성능 개선을 재확인할 수 있다. 이는 문제의 규모가 증가할수록 초기 연결가중치의 설정은 더욱 어려워지며 이에 따른 학습성능의 의존도도 더욱 증가한다는 것을 알 수 있다.

Table 2. Simulation results in Total error= $10^{-3}$

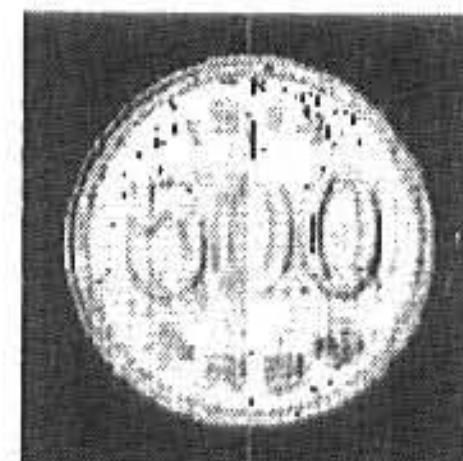
	T	PSNR
BP 알고리즘	15088	78.130981
기존의 NLPCA	6542	78.130852
제안된 NLPCA	1136	78.141144



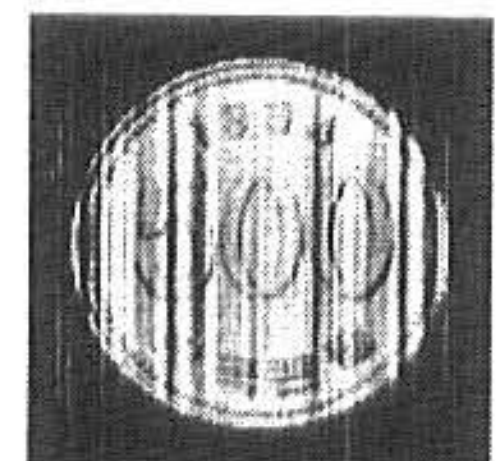
(a) Original image



(b) BP algorithm



(c) Conventional NLPCA



(d) Proposed NLPCA

Fig. 7 Reconstructed coin images by 3 algorithms

Fig.7은 동전의 원 영상과 3 가지 알고리즘으로 복원된 영상을 나타내고 있다. Fig.7(a)는 원 영상이고, (b), (c), (d)는 각각 학습반복수 1000 번에 대해 40 개의 주요특징추출 후, 이들을 이용하여 재구성된 영상을 보여준다. Fig 에서도 역전파 알고리즘에 의해 재구성된 결



과인 영상 (b)가 상대적으로 NLPCA를 이용한 (c)와 (d)영상보다 심한 왜곡이 있음을 알 수 있으며, 제안된 NLPCA가 기존의 NLPCA보다 우수한 비선형 속성을 가진 특징들을 추출할 수 있음을 알 수 있다. 따라서 제안된 학습알고리즘의 5층 신경망을 이용하면 효과적으로 입력영상의 주요 특징들을 추출할 수 있으며, 또한 영상의 압축효과도 우수함을 알 수 있다.

#### 4. 결론 및 향후과제

본 논문에서는 모멘트를 추가한 학습알고리즘을 5층 전향 신경망에 적용하여 입력 데이터들의 효율적인 비선형 주요 특징추출을 위한 방법을 제안하였다. 제안된 학습알고리즘은 과거의 속성을 반영하여 보다 능률적으로 신경망을 학습하기 위한 모멘트 항이 추가된 기법이다. 이는 최적해로의 수렴에 따른 발진을 억제하여 그 수렴속도를 개선시키는 모멘트의 장점을 살리기 위함이다.

제안된 학습알고리즘을 이용한 NLPCA 신경망을 256×256 픽셀의 간암영상과 128×128 픽셀의 동전영상들을 대상으로 시뮬레이션 한 결과, 역전파 알고리즘을 이용한 3층 신경망이나 기존의 NLPCA 신경망보다 우수한 수렴 성능과 특징 추출성능이 있음을 확인할 수 있었다.

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

#### 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 영상 데이터의 주요특징추출. 한국정보처리학회 논문지, 제6권, 제4호, 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, February 1991.
- 4) E.C. Malthouse : Limitations of Nonlinear PCA as Performed with Generic Neural Networks. *IEEE Transactions on Neural Networks*, Vol.9, No.1, pp.165-173, January 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. Jaskelainen : 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, September 1997.
- 8) D. Tzovaras and M. G. Strintzis : Use of Nonlinear Principal Component Analysis and Vector Quantization for Image Coding. *IEEE Transactions on Image Processing*, Vol.7, No.8, pp.1218-1223, August 1998.
- 9) F. D. Frate and G. Schiavon : Nonlinear Principal Component Analysis for the Radiometric Inversion of Atmospheric Profiles by Using Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, Vol.37, No.5, pp.2335-2342, September 1999.
- 10) N. Qian : On the Momentum Term in Gradient Descent Learning Algorithms. *Neural Networks* 12, pp.145-151, 1999.

(2000년 8월 20일 접수, 2000년 11월 25일 채택)