

기술기하강과 동적터널링에 기반을 둔 학습알고리즘의 신경망을 이용한 영상데이터의 주요특징추출

조 용 현†

요 약

본 논문에서는 새로운 학습알고리즘의 3층 전향 신경망을 이용한 입력데이터의 주요 특징추출에 대해서 제안하였다. 제안된 학습알고리즘에서에서는 빠른 수렴속도의 최적화가 가능하도록 하기 위하여 기술기하강의 역전파 알고리즘을 이용하고, 국소최적해를 만났을 때 이를 벗어난 새로운 연결가중치의 설정을 위하여 동적터널링의 역전파 알고리즘을 이용함으로써 빠른 수렴속도로 전역최적해로 수렴하도록 학습시킬 수 있다. 제안된 학습 알고리즘을 이용한 다층신경망을 12×12 픽셀의 영상데이터들과 128×128 픽셀의 Lenna 영상데이터를 대상으로 시뮬레이션한 결과, 단층신경망을 이용하는 Sanger 방법이나 측면연결을 가지는 단층신경망을 이용하는 Foldiak 방법 및 기술기하강에 기초를 둔 기존의 역전파 알고리즘을 이용한 다층신경망에 의한 결과와 비교할 때 더욱 우수한 수렴성능과 추출성능이 있음을 확인할 수 있었다.

Principal Feature Extraction of Image Data Using Neural Networks of Learning Algorithm Based on Steepest Descent and Dynamic Tunneling

Yong-Hyun Cho†

ABSTRACT

This paper proposes an efficient principal feature extraction of the image data using neural networks of a new learning algorithm. The proposed learning algorithm is a backpropagation(BP) algorithm based on the steepest descent and dynamic tunneling. The BP algorithm based on the steepest descent is applied for high-speed optimization, and the BP algorithm based on the dynamic tunneling is also applied for global optimization. Converging to the local minimum by the BP algorithm of steepest descent, the new initial weights for escaping the local minimum is estimated by the BP algorithm of dynamic tunneling. The proposed algorithm has been applied to the 3 image data of 12×12 pixels and the Lenna image of 128×128 pixels respectively. The simulation results shows that the proposed algorithm has better performances of the convergence and the feature extraction, in comparison with those using the Sanger method and the Foldiak method for single-layer neural networks and the BP algorithm for multilayer neural network.

1. 서 론

주어진 입력데이터에 대한 주요 특징의 효율적인

추출은 패턴인식에서의 데이터 압축 및 부호화, 영상 처리에서의 주파수 추정을 위한 고해상도 공간분석, 그리고 제어 시스템의 감축 등에서 매우 절실하게 요구되는 연구과제이다. 지금까지 이를 위한 많은 연구들이 활발하게 이루어지고 있다^[1-6].

† 정 회 원 : 대구효성가톨릭대학교 공과대학 전자·정보공학부 교수

논문접수 : 1998년 9월 28일, 심사완료 : 1999년 2월 11일

주요성분분석(principal component analysis : PCA)은 주어진 입력데이터의 특징을 추출하는 기법으로 널리 이용되고 있으며, 이는 정상확률과정(stationary stochastic process)의 입력벡터와 특별한 차원의 출력벡터를 위한 최적의 선형변환(linear transform)을 결정하는 통계적 방법이다^{11, 12}. 결국 입력데이터 내에 포함된 정보를 추출하고 압축하여 통계적 규칙들을 찾아 내는 것이다. 따라서 PCA는 대용량의 입력데이터를 통계적으로 독립인 특징들의 집합으로 변환시키는 기법이다.

PCA를 위한 기존의 수치적 방법들은 대부분 높은 차원의 입력공간을 더 낮은 차원의 표현공간으로 사상시켜 입력데이터가 가지는 두드러진 특징들을 추출한다. 여기서 사상은 푸리에 변환(Fourier transform), 불연속 역변환(discrete cosine transform), 그리고 KL 변환(Karhunen Loève transform) 등의 변환동작을 통해서 얻어진다. 이들 변환 중 KL 변환은 확률과정에 대해서는 최적의 방법으로 알려져 있다¹¹. 이는 신호 에너지의 대부분을 처음 몇몇 표본들 내에 포함시킬 수 있으며, 동시에 데이터의 완전한 비상관(decorrelation) 관계를 얻을 수 있기 때문이다. 특히 후자의 특성은 효율적인 데이터 압축 및 재구성능을 가능하게 한다. 그러나 이들 방법들에서는 입력데이터에 대한 공분산행렬(covariance matrix)의 계산이 요구되며, 고유치(eigenvalue)와 그에 따른 고유벡터(eigenvector)를 계산하기 위한 수치적 과정이 필요하다. 이때 계산되는 고유치와 고유벡터는 데이터의 주요 특징을 추출하는데 이용된다. 그러나 대규모의 데이터를 대상으로 할 때 공분산행렬의 계산은 매우 복잡하고 부정확하게 이루어지게 된다. 또한 최대 유효고유치에 일치하는 고유벡터만이 변환과정에 이용된다 할지라도 모든 고유치와 고유벡터들이 계산되어야 하는 불합리성이 존재한다. 따라서 기존의 변환에 기초를 둔 기법들은 대규모의 실시간 처리가 요구되는 응용문제에서는 매우 비효율적이다.

최근, 대규모 분산 병렬처리와 학습 능력을 가진 신경망은 신호와 패턴처리 분야에서 특히 많은 관심이 되어왔다^{14, 15}. 이러한 속성을 가진 신경망에 기반을 둔 여러 가지 PCA 기법들이 기존의 변환에 기반을 둔 방법들의 대안으로 제안되어 왔다^{11, 31}. 이때 이용된 신경망은 주로 입력층과 출력층으로 구성된 단층구조이다. Oja¹¹는 정규화된 헤비안규칙(normalized Hebbian rule)의 적용학습 방법을 이용한 간단한 뉴런모델을 제안하

고, 정상입력벡터 계열의 주요 특징으로 연결가중치 벡터(weights vector)가 수렴됨을 증명하였다. 그러나 이 방법은 첫 번째 주요 특징만을 추출할 수 있으며, 임의의 m개의 특징추출로의 확장은 불가능하였다. Sanger 등¹¹은 일반화된(generalized) 헤비안규칙을 이용함으로써 정상과정의 첫 번째 m 개의 주요 특징들을 계산하기 위한 다중 뉴런모델을 제안하였다. 한편, Foldiak^{11, 21}은 망의 입력과 출력사이의 연결가중치 경신에는 정규화된 헤비안규칙을 이용하고, 망의 출력사이의 측면연결 가중치 경신에는 반 헤비안규칙(anti-Hebbian rule)을 함께 이용한 학습알고리즘을 제안하였다. 최근 Bannar 등¹³은 Gram-Schmidt 직교화와 순환최소자승(recursive least square)법을 이용한 학습알고리즘을 제안하였다.

지금까지 제안된 대부분의 방법들은 단층신경망을 이용하고 있으며 헤비안규칙에 기초를 둔 학습알고리즘으로 연결가중치 벡터를 경신하고 있다. 또한 학습율(learning rate)은 고정된 값으로 설정하였다. 한편 Bannar 등의 방법도 적응적 학습율을 이용하여 가중치 벡터를 경신시키나 최적의 주요특징을 찾는 데는 여전히 적절한 초기값 설정과 같은 제약이 있다.

한편 학습알고리즘으로써 역전파(backpropagation) 알고리즘을 이용한 다층신경망이 영상데이터의 압축 등 입력데이터의 주요 특징을 추출하는데 이용되어 왔다^{15, 16}. 이는 n차원 공간에서 m차원 공간으로 패턴집합이 쉽게 사상되도록 학습시킬 수 있기 때문이다. 그러나 기울기하강(steepest descent)의 속성을 이용하는 역전파 알고리즘은 학습파라미터의 설정에 따라 수렴속도와 건설성 중 하나 이상의 문제를 가지며, 전역최소점으로서의 수렴이 보장되어 있지 않다. 이러한 문제를 해결하기 위한 여러 방법들이 많은 연구자들에 의해서 연구되어 왔다^{17, 18}.

본 논문에서는 새로운 학습알고리즘의 3층 전향신경망을 이용하여 입력데이터의 주요 특징을 효율적으로 추출하는 방법에 대해서 제안한다. 제안된 새로운 학습알고리즘은 기울기하강과 동적터널링(dynamic tunneling)의 속성을 조합한 역전파 알고리즘이다. 제안된 학습 알고리즘을 이용한 신경망을 12×12 픽셀(pixel)의 3가지 영상데이터들과 128×128 픽셀의 Lena 영상데이터를 대상으로 시뮬레이션하여 그 타당성을 확인하였으며, 기존의 Sanger 방법과 Foldiak의 방법 및 기울기하강법에 기초를 둔 역전파 알고리즘을 이용

한 신경망에 의한 결과와 비교 고찰하였다.

2. 적응적 특징추출을 위한 신경망의 학습알고리즘

PCA는 n 차원 입력공간의 데이터를 m 차원 출력공간의 데이터로 투영시키는 것이다. 여기서 m (n 이더, 이는 입력데이터 벡터의 대부분 내부정보를 유지하도록 차원의 감소를 얻는 것이다. 자기상관행렬(auto-correlation matrix) $\mathbf{R}_{xx} = \langle \mathbf{x}\mathbf{x}^T \rangle$ 를 가진 평균이 영인 입력벡터 $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ 에 대해서 생각해 보자. 여기서 $\langle \mathbf{x} \rangle$ 는 기대치를 나타낸다. 또한 $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ 이 \mathbf{R}_{xx} 의 고유벡터와 직교되는 연결가중치 벡터라 할 때, $\mathbf{w}_1 = [w_{11}, w_{12}, \dots, w_{1n}]^T$ 는 가장 큰 고유치 λ_1 과 일치하며, $\mathbf{w}_2 = [w_{21}, w_{22}, \dots, w_{2n}]^T$ 는 두 번째로 큰 고유치 λ_2 와 $\mathbf{w}_n = [w_{n1}, w_{n2}, \dots, w_{nn}]^T$ 는 가장 작은 고유치 λ_n 과 일치한다. 이상의 관계를 행렬방정식으로 나타내면 다음과 같다. 즉,

$$\mathbf{R}_{xx} \mathbf{w}_j = \lambda_j \mathbf{w}_j \quad (j = 1, 2, \dots, n) \quad (1)$$

이다. 여기서 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ 이다.

주어진 입력벡터 \mathbf{x} 를 위한 첫 번째 m 개의 주요 특징벡터 \mathbf{y} 는 다음의 선형변환식으로 나타낼 수 있다. 즉,

$$\mathbf{y} = \mathbf{W} \mathbf{x} \quad (2)$$

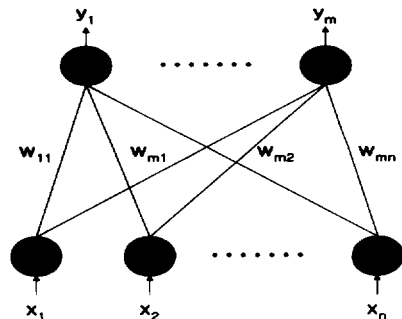
이다. 여기서 $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]^T \in \mathbf{R}^{m \times n}$ 이다. 즉, 식 (2)는 행렬 \mathbf{W} 의 열들이 그것의 가장 큰 고유치들과 일치하는 상관행렬 \mathbf{R}_{xx} 의 고유벡터들임을 의미한다. 다시 말하면, 입력데이터 공간에서 m 차원의 주요 특징을 나타내는 부공간은 \mathbf{R}_{xx} 의 m 개 주요 고유벡터에 의해 구성된 부공간으로 정의된다. 결국 PCA는 $\langle \|\mathbf{w}_j^T \mathbf{x}\|_2^2 \rangle$ 가 최대인 고유벡터 $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ 의 방향을 찾는 것이다.

따라서 입력되는 데이터의 주요특징들을 추출하기 위한 기존의 방법들을 살펴보면, 수치적 방법에서는 먼저 자기상관행렬 \mathbf{R}_{xx} 를 계산한 다음 그것과 일치하는 고유벡터를 계산하여야 한다. 그러나 입력데이터 벡터가 대규모의 차원을 가진다면, 상관행렬은 대단히 크게 되며 또한 요구되는 고유벡터를 찾는 것도 매우 복잡한 일이다. 이를 해결하기 위한 대안으로 상관행렬의 고유벡터를 실시간으로 추정하기 위하여 적응학

습 알고리즘을 이용하는 방법들이 제안되었다. 여기서는 입력데이터 벡터로부터 직접 고유벡터를 추정할 수 있기 때문에 상관행렬의 추정과정이 요구되지 않는다. 이때 망의 구조로는 단층신경망과 다층신경망이 이용된다. 지금까지 이들 망을 학습시키기 위한 많은 적응 학습 알고리즘들이 제안되었으나, 본 연구에서는 단층 신경망을 대상으로 Sanger에 의해 제안된 학습알고리즘과 Foldiak에 의해 제안된 학습 알고리즘에 대해서 알아본다. 또한 다층신경망을 이용한 입력데이터의 주요특징 추출에 대해서도 살펴본다.

2.1 단층신경망을 이용한 적응적 특징추출

주어진 입력벡터에 대한 첫 번째로 주요한 특징을 추출하기 위해서는 출력장치로 한 개의 선형 처리장치를 가진 망이 이용된다. 이를 위한 많은 학습알고리즘들이 제안되었으며, 이때의 학습은 자율학습(unsupervised learning)이다. m 개의 주요특징을 추출하기 위해서는 m 개의 선형 처리장치로 구성되는 망이 요구되나 한 개의 선형 처리장치를 가진 망을 위한 학습알고리즘을 확장하여 여러개의 선형 처리장치를 가진 망을 학습시키는 것은 불가능하다. 따라서 주어진 입력벡터에 대한 첫 번째로 중요한 m 개의 주요특징들을 추출하기 위해서는 m 개의 선형 처리장치를 출력층으로 하는 단층신경망이 구성되어야 하며, 이 망을 위한 별도의 학습알고리즘이 요구된다. (그림 1)은 첫 번째로 중요한 m 개의 주요특징들을 추출하기 위해 n 개의 입력뉴런과 m 개의 출력뉴런으로 구성된 단층신경망의 구조이다. 그림에서 입력과 출력의 관계를 행렬방정식으로 나타내면 식 (2)와 동일하다. 여기서 각 뉴런의 활성화함수는 선형함수이고, $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]^T$ 로



(그림 1) m 개의 첫 번째 주요특징 추출을 위한 단층신경망

연결가중치 행렬이며 $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ 는 연결가중치 벡터이다. 따라서 신경망을 학습시켜 연결가중치를 구함으로써 주어진 n 개의 입력벡터 x 로부터 m 개의 주요 특징벡터 즉 출력벡터 y 를 구할 수 있다.

일반화된 헤비안규칙에 기반을 둔 (그림 1)의 단층신경망을 학습시키는 알고리즘을 살펴보면 다음과 같다. 즉,

$$w_{ij}(t+1) = w_{ij}(t) + \eta_i y_i(t) [x_j(t) - \sum_{h=1}^m w_{hj}(t) y_h(t)]$$

$$(i = 1, 2, \dots, m, j = 1, 2, \dots, n) \quad (3)$$

이다. 여기서 η_i 는 학습율로서 학습 반복수 t 가 증가함에 따라 보통 지수함수적으로 감소되나 실험에서는 고정된 값으로 설정하였다. 식 (3)을 이용하여 연결가중치를 구한 다음 입력벡터와 함께 식 (2)에 대입하면 주요특징 벡터를 구할 수 있다. 즉, 학습은 상관행렬의 고유벡터를 추정하기 위하여 임의의 초기연결 가중치로부터 주요특징을 내포하는 연결 가중치를 구하는 과정이다. 결국 신경망은 입력데이터 내에 포함된 상호 독립인 정보를 부호화시키는 검출기이다.

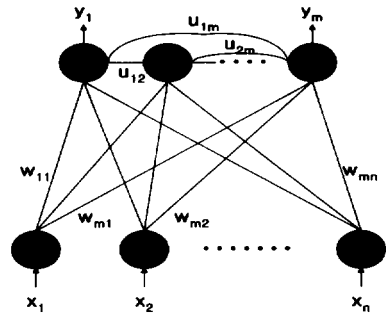
한편, 입력뉴런과 출력뉴런 사이의 연결 및 출력뉴런 상호간에도 연결을 가진 새로운 단층신경망이 m 개의 첫 번째 주요특징을 추출하기 위해 제안되었다^[2]. (그림 2)는 n 개의 입력뉴런과 m 개의 출력뉴런으로 구성된 입력과 출력뉴런간 및 출력뉴런 상호간의 측면연결을 가진 단층신경망의 구조이다. 그림에서 입력과 출력의 관계를 나타내면 식 (4)와 같다. 즉,

$$y_i = \sum_{j=1}^n w_{ij} x_j + \sum_{h=1}^m u_{ih} y_h$$

$$(i = 1, 2, \dots, m) \quad (4)$$

이다. 여기서 w_{ij} 는 입력뉴런과 출력뉴런을 연결하는 연결가중치이고, u_{ih} 는 출력뉴런 상호간의 측면연결 가중치이다. 이때 출력뉴런 간의 상호연결을 보면 뉴런 i 는 다만 $h < i$ 인 뉴런에만 연결된다. 따라서 신경망을 학습시켜 입출력 뉴런간 및 출력뉴런 상호간의 연결가중치를 각각 구함으로써 주어진 n 개의 입력벡터 x 로부터 m 개의 주요특징벡터 즉 출력벡터 y 를 구할 수 있다.

(그림 2)와 같은 단층신경망을 학습시키기 위해서 정규화된 헤비안규칙과 반 헤비안규칙에 기반을 둔 알고리즘이 이용된다. 입력과 출력뉴런 간의 연결가중치 경신에는 정규화된 헤비안규칙을 이용하고 출력뉴런



(그림 2) m 개의 첫 번째 주요특징 추출을 위한 측면연결의 단층신경망

간의 측면 연결가중치 경신에는 반 헤비안규칙을 이용한다. 먼저 입출력 뉴런간의 연결가중치 w_{ij} 의 경신규칙을 살펴보면 다음과 같다. 즉,

$$w_{ij}(t+1) = w_{ij}(t) + \eta [y_i(t) x_j(t) - w_{ij}(t) y_i(t)^2]$$

$$(i = 1, 2, \dots, m, j = 1, 2, \dots, n) \quad (5)$$

이다. 또한 출력뉴런 간의 측면 연결가중치 u_{ih} 의 경신규칙은 다음과 같다. 즉,

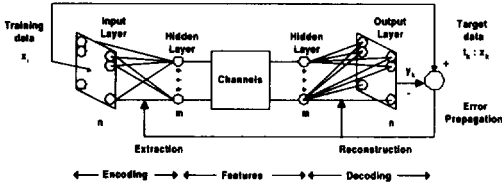
$$u_{ih}(t+1) = u_{ih}(t) + \rho y_i(t) y_h(t), (i > h) \quad (6)$$

이다. 여기서 ρ 는 식 (5)의 η 와 일반적으로 동일한 값으로 설정되며, 둘 다 양의 값을 가지는 학습율로 실험에서는 고정된 값으로 설정하였다. 따라서 식 (5)와 (6)을 이용하여 입력과 출력뉴런 간의 연결가중치 및 출력뉴런 상호간의 측면 연결가중치 경신을 경신시켜 식 (4)에 대입하면 m 개의 주요특징들을 추출할 수 있다.

2.2 다층신경망을 이용한 적응적 특징추출

입력데이터에 존재하는 주요특징 추출은 입력의 개수를 감소시키는데 사용된다. 즉, 주요 특징의 추출은 원 데이터를 감소된 차원의 특징공간(feature space)에 사상시키는 것이다. 이러한 관점에서 보면, PCA는 출력을 입력과 같게 만들고 은닉층으로부터 주요 특징을 가진 신호를 얻는 지도학습(supervised learning) 문제로 공식화할 수 있다. 즉, 신경망에서 입력층과 은닉층간의 연결은 데이터를 부호화(encoding)하는 것이고, 은닉층과 출력층간의 연결은 부호화된 데이터를 복호화(decoding)하는 작업으로 볼 수 있다. 따라서 다층신

경망을 입력데이터의 주요 특징추출에 적용할 때, 은닉층 뉴런의 개수는 (그림 3)에서 보는 바와 같이 입력층과 출력층 뉴런의 개수보다 더 적게 구성한다. 일반적으로 3층 전향 신경망은 가장 간단한 다층신경망이다.



(그림 3) m개의 첫 번째 주요특징 추출을 위한 다층신경망

그림에서, 학습하는 동안에 입력데이터들은 학습데이터 및 원하는 출력데이터로써 각각 입력층과 출력층에 동시에 이용된다. 신경망의 첫 번째 단계는 데이터의 중복성(redundancy)을 감소하기 위한 변형규칙을 학습시키는 것이고, 두 번째 단계는 변형된 규칙을 되돌리기 위해 학습하는 과정이다. 결국 신경망은 학습을 통하여 입력데이터가 은닉층 뉴런의 출력에서 원 크기의 m/n 으로 압축되도록 내부연결 가중치를 경신시킨다. 그리고 나서 은닉층 뉴런들에서 생성된 값을 읽어 m 개의 주요 특징들을 추출하고, 추출된 m 개의 특징들을 동일 신경망의 출력층 뉴런으로 전송시킴으로써 원래의 입력데이터를 다시 복구시킬 수 있다.

3. 기울기하강과 동적터널링 속성을 조합한 학습알고리즘

다층신경망은 입력층과 출력층 간에 하나 이상의 은닉층을 가진 구조이다. 은닉층은 단층 신경망이 가지는 여러 가지 한계점을 극복하기 위해서 제안되었다. 다층신경망의 학습알고리즘으로 역전파 알고리즘이 제안되어 최근에 널리 이용되고 있다.

제안된 학습알고리즘은 기울기하강과 동적터널링에 기초를 둔 역전파 알고리즘이다. 먼저 기존의 역전파 알고리즘에 대하여 간략하게 설명한다.

i_{jp} 와 y_{kp} 를 각각 p번째 입력 신호에 대한 은닉층과 출력층의 j번째와 k번째 뉴런 출력으로 하자. 이 때 신경망의 입력출력 관계는 다음과 같다.

$$y_{kp} = f_k(\sum_j w_{kj} i_{jp})$$

$$i_{jp} = f_j(\sum_i w_{ji} x_{ip}) \tag{7}$$

여기서 $f(\cdot)$ 와 w_{kj} 는 뉴런의 활성화함수와 연결가중치를 나타낸다. 또한 전체 오차함수 $E(\mathbf{w})$ 와 p번째 입력에 대한 오차함수 $E_p(\mathbf{w})$ 는 다음과 같이 각각 정의된다.

$$E(\mathbf{w}) = \sum_p E_p(\mathbf{w})$$

$$E_p(\mathbf{w}) = (1/2) \sum_k (d_{kp} - y_{kp})^2 \tag{8}$$

식에서 d_{kp} 는 출력층의 k번째 뉴런에서의 원하는 출력이다. 연결가중치 \mathbf{w} 에 대하여 $E(\mathbf{w})$ 를 최소화하기 위해, 기존의 역전파 알고리즘에서는 $(\partial E_p(\mathbf{w}) / \partial \mathbf{w})$ 를 계산한 후 $E_p(\mathbf{w})$ 의 기울기 하강 방향으로 가중치를 경신시킨다. 이때 출력층에 대한 연결가중치는

$$w_{kj}(t+1) = w_{kj}(t) + \eta \delta_{kp} i_{jp} + \alpha \Delta_p w_{kj}(t-1)$$

$$\delta_{kp} = (y_{kp} - d_{kp}) f'_k(\sum_j w_{kj} i_{jp}) \tag{9}$$

와 같이 경신되며, 또한 은닉층의 연결가중치는

$$w_{ji}(t+1) = w_{ji}(t) + \eta \delta_{jp} x_{ip} + \alpha \Delta_p w_{ji}(t-1)$$

$$\delta_{jp} = f'_j(\sum_i w_{ji} x_{ip}) \sum_k \delta_{kp} w_{kj} \tag{10}$$

와 같이 경신된다. 여기서 η 는 학습율이고 α 는 모멘텀(momentum)이다. 즉, 연결가중치의 경신은 위의 식 (9)와 (10)에 따라 후향으로 이루어진다.

한편, 동적터널링 알고리즘은 Yao에 의해서 제안되었다¹⁾. 이는 동적최적화와 동적터널링으로 구성된다. 여기서 동적최적화는 초기 상태에서부터 국소최소값 중의 한 값을 찾는 과정이며, 동적터널링은 찾은 국소최소값보다 하위유역에 위치한 새로운 초기 상태를 찾는 과정이다. 즉, 동적터널링은 동적최적화가 국소최소점에 빠지는 것을 막아준다.

따라서, 신경망에서 연결가중치 \mathbf{w} 에 대한 오차함수 $E(\mathbf{w})$ 를 최소화하는 문제를 살펴보자. 이 문제는 $(\partial E_p(\mathbf{w}) / \partial \mathbf{w})$ 를 이용하여 $E(\mathbf{w})$ 의 국소최소값 \mathbf{w}^* 를 찾는 것으로 변형될 수 있다. 즉, $\Delta_p \mathbf{w} = -\alpha (\partial E_p(\mathbf{w}) / \partial \mathbf{w})$ 는 동적최적화와 동일하다. 동적터널링은 동적최적화에서 구해진 \mathbf{w}^* 에다 결정론적으로나 임의로 설정되는 교란 벡터 ϵ 이 추가된 새로운 시작점에서 출발하는 것으로 다음과 같은 미분 방정식의 형태로 주어진다.

$$\Delta_p \mathbf{w} = -(\partial E_p(\mathbf{w}) / \partial \mathbf{w}) [1 / \sqrt{(\mathbf{w} - \mathbf{w}^*)^T (\mathbf{w} - \mathbf{w}^*)}]^{\gamma} - \theta E_p^*(E_p^*(\mathbf{w})) \tag{11}$$

여기서, θ 는 터널링 페널티(tunneling penalty)이며, γ 의 값은 $(\partial E_p(\mathbf{w})/\partial \mathbf{w})$ 의 원점 \mathbf{w}^* 의 차수보다 크거나 같은 값을 가진다. 또한 $E_p^*(\mathbf{w}) = E_p(\mathbf{w}) - E_p(\mathbf{w}^*)$ 로 정의되며, $E_p^*(\cdot)$ 는 다음과 같이 두 개의 분리된 구분선형(piecewise linear) 함수로 정의된다.

$$E^*(z) = z \quad (z > 0)$$

$$0 \quad (z < 0)$$
(12)

식 (11)에서, 우편 첫 번째 항의 분모 $[(\mathbf{w}-\mathbf{w}^*)^T (\mathbf{w}-\mathbf{w}^*)]^2$ 항은 동적최적화로부터 구해진 국소최소점인 평형상태 \mathbf{w}^* 를 제거하기 위한 것이며, 두 번째 항은 하위유역에 존재하는 즉, $E_p(\mathbf{w}^{(0)}) \leq E_p(\mathbf{w}^*)$ 인 새로운 초기 상태 $\mathbf{w}^{(0)}$ 를 찾기 위한 것이다.

그러므로 동적터널링에 기초한 역전과 알고리즘은 식 (11)처럼 주어지고, 이때 출력층과 은닉층 각각의 $(\partial E_p(\mathbf{w})/\partial \mathbf{w})$ 는 다음 식과 같이 정의될 수 있다.

$$\partial E_p(\mathbf{w})/\partial \mathbf{w}_{kj} = - \eta \delta_{kp} i_{jp}$$

$$\partial E_p(\mathbf{w})/\partial \mathbf{w}_{ji} = - \eta \delta_{jp} x_{ip}$$
(13)

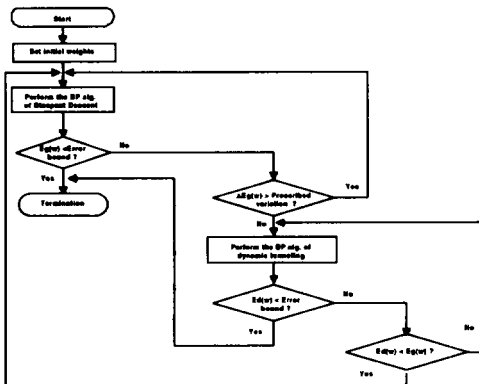
여기서 δ_{kp} 와 δ_{jp} 는 각각 식 (9)와 (10)으로 주어진다.

한편 추출된 특징들의 질을 평가하기 위해 다음의 SNR(signal to noise ratio)을 이용한다. 즉,

$$SNR = 10 \log_{10}[E(d_k^2)/E(d_k - y_k)^2]$$
(14)

이다. 여기서 $E[\cdot]$ 는 영상의 모든 픽셀에 대한 기대치이다.

따라서 기울기하강과 동적터널링을 조합한 역전과 알고리즘인 새로운 학습알고리즘의 흐름도를 그림으로 도시하면 (그림 4)와 같다. 그림에서, $E_g(\mathbf{w})$ 와 $E_d(\mathbf{w})$ 는



(그림 4) 제안된 학습 알고리즘의 흐름도

각각 기울기하강과 동적터널링에 기초한 역전과 알고리즘을 사용할 때 계산된 전체 오차함수값이다.

4. 시뮬레이션 결과 및 분석

제안된 학습알고리즘을 이용한 신경망의 성능을 평가하기 위해서 3층 전향 신경망을 구성하였다. 다층신경망에서의 입력뉴런과 은닉뉴런 및 은닉뉴런과 출력뉴런 사이의 초기 연결가중치, 단층신경망의 입력뉴런과 출력뉴런 및 출력 뉴런사이의 측면초기 연결가중치는 각각 -0.5에서 +0.5 사이의 임의의 값으로 설정하였다. 학습은 전체 반복회수가 20,000이상이거나 전체 오차값이 허용치 이하일 때, 또는 전체 오차함수 값의 변화가 10^{-6} 이하일 때 종료되도록 하였다.

제안된 학습알고리즘의 다층신경망을 12×12 픽셀의 영상데이터들과 128×128 픽셀의 Lena 영상데이터를 대상으로 펜티엄 II-MMX200 (MHz) 컴퓨터로 시뮬레이션하여 그 타당성을 확인하였으며, 기존의 Sanger 방법과 Foldiak의 방법 및 기울기하강법에 기초를 둔 역전과 알고리즘을 이용한 신경망에 의한 결과와 비교 고찰하였다.

4.1 영상데이터 문제

주어진 입력데이터의 압축 및 주요특징을 추출하기 위하여 12×12 픽셀의 3가지 영상데이터를 이용하였다. 실험에 이용된 다층신경망의 입력뉴런과 출력뉴런 및 단층신경망의 입력뉴런수는 각각 144개로 하였으며, 은닉층 뉴런의 개수는 4, 8, 12, 16, 24, 32, 48, 64로 바꾸어 가면서 시뮬레이션 하였다. 여기서 허용 오차함수값은 10^{-4} 로하였다. 이때 은닉층 뉴런 수는 입력데이터로부터 추출되는 주요특징의 개수이다. (그림 5)는 시뮬레이션에 이용된 표본 영상데이터들을 나타낸 것이다.

12	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
12	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
12	3	4	4	4	4	4	4	4	4	3	2				
12	3	4	5	5	5	5	5	4	3	2					
12	3	4	5	6	6	6	5	4	3	2					
12	3	4	5	6	7	7	6	5	4	3	2				
12	3	4	5	6	7	7	6	5	4	3	2				
12	3	4	5	6	6	6	5	4	3	2					
12	3	4	5	5	5	5	5	4	3	2					
12	3	4	4	4	4	4	4	4	3	2					
12	3	3	3	3	3	3	3	3	3	2					
12	2	2	2	2	2	2	2	2	2	2					
15	2	4	6	3	1	4	3	7	4	5	4				
14	4	3	5	6	1	7	7	0	4	7	5				
13	7	1	5	7	1	3	4	0	6	0	3				
13	7	4	6	1	7	7	4	6	5	1	4				
17	2	1	6	7	3	5	3	1	2	4	1				
11	5	3	5	2	6	4	5	1	4	4	1				
16	6	5	7	0	2	6	6	5	2	4	5				
16	3	2	7	3	7	3	3	7	4	2	3				
13	5	6	2	2	3	5	3	4	1	4	5				
16	3	3	5	4	2	3	4	5	6	3					
11	0	3	5	2	6	3	7	6	3	5	7				
13	5	2	1	3	5	2	1	5	6	1	1				

(a) 영상데이터 1

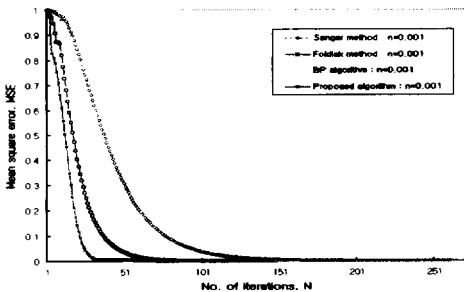
(b) 영상데이터 2

17	7	6	6	5	5	4	4	3	3	2	2
17	7	6	6	5	5	4	4	3	3	2	2
16	6	7	7	6	6	5	5	4	4	3	3
16	6	7	7	6	6	5	5	4	4	3	3
15	5	6	6	7	7	6	6	5	5	4	4
15	5	6	6	7	7	6	6	5	5	4	4
14	4	5	5	6	6	7	7	6	6	5	5
14	4	5	5	6	6	7	7	6	6	5	5
13	3	4	4	5	5	6	6	7	7	6	6
13	3	4	4	5	5	6	6	7	7	6	6
12	2	3	3	4	4	5	5	6	6	7	7
12	2	3	3	4	4	5	5	6	6	7	7

(c) 영상데이터 3

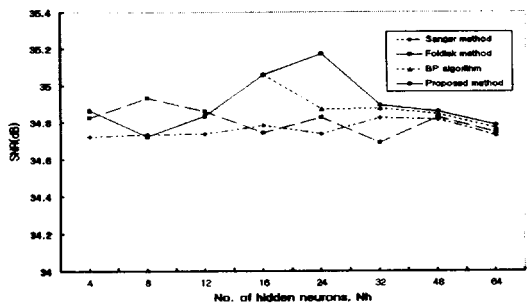
(그림 5) 12×12 픽셀을 가진 표본 영상데이터

(그림 6)은 영상데이터 1을 대상으로 Sanger 방법, Foldiak 방법, 역전파 알고리즘, 그리고 제안된 알고리즘을 각각 이용한 신경망의 평균자승오차(mean square error : MSE) 값의 변화를 나타낸 것이다. 이때 학습율 η 와 은닉층 뉴런 수는 각각 0.001과 12로 설정하였다. 그림에서 보면, 제안된 알고리즘과 역전파 알고리즘은 동일한 수렴속도를 나타낸다. 이는 기울기하강의 속성만이 이들 두 알고리즘에서 동시에 이용되었음을 보여준다. 특히 4가지 학습알고리즘 모두가 최적해로 수렴되었음도 확인할 수 있다. 이 경우는 단층신경망이나 다층신경망의 초기 연결가중치들이 우연히도 최적해로 수렴가능한 영역에 설정된 경우이다. 일반적으로 최적해로 수렴가능한 영역에 초기 연결가중치를 설정시키는 구체적인 방안이 없으며 대부분 경험적으로 설정한다. 한편 그림에서 보면 단층신경망을 이용하는 경우에 Foldiak 방법이 Sanger 방법보다는 빠른 수렴속성을 가지며, 또한 다층신경망을 이용한 경우가 단층신경망을 이용한 경우보다도 빠른 수렴속성이 있음을 알 수 있다. 이는 주요 특징추출을 위한 입력영상 데이터의 픽셀간 상관관계에 의존하는 것으로 추측된다.



(그림 6) 학습 알고리즘들에 대한 MSE의 변화

(그림 7)은 표본 영상데이터 1에 대해서 은닉층 뉴런의 개수와 SNR간의 관계를 보여준다. 여기서도, η 와 랜덤시드(random seed) 값을 각각 0.001과 10으로 설정하였다. 그림에서 입력층과 은닉층 신호가 동일한 정확도를 가질 때, 압축률은 은닉층 뉴런의 개수에 일치하는 1/36~4/9사이가 된다. 또한 그림에서는 은닉층 뉴런의 개수가 증가함에 따라 SNR이 조금씩 증가하고 있음을 보여준다. 이것은 은닉층 뉴런이 주요특징 추출을 위해 요구된 영상의 질을 보존하기 위해서 반드시 필요하다는 것을 나타낸다. 그리고 제안된 알고리즘에서의 SNR이 Sanger 방법, Foldiak 방법, 역전파 알고리즘의 그것보다 약간 우수하다. 이는 제안된 알고리즘이 보편화에 더욱 많은 영향을 준다는 것을 나타낸다. 또한 다층신경망을 이용하는 경우가 단층신경망을 이용하는 경우보다 전체적으로 큰 값을 가짐도 알 수 있다. 다층신경망을 이용하는 경우에 은닉층 뉴런의 수가 16개까지는 제안된 알고리즘과 역전파 알고리즘의 SNR은 동일한 값을 가진다. 이는 동일한 학습율과 초기 연결 가중치값에 대하여 오직 기울기하강의 속성만 동작한 경우이다. 한편, 표본 영상데이터 3 개 모두에 대하여 출력뉴런의 수를 바꾸어 가면서 실험해 본 결과, 영상데이터 3이다른 2 개의 나머지 영상데이터보다 더 큰 SNR 값을 가짐을 확인할 수 있었다. 이 결과에서 보면 영상데이터의 복잡도에 따라 SNR이 영향을 받는다는 것을 확인할 수 있다.



(그림 7) 은닉층 뉴런의 개수에 대한 SNR의 변화

<표 1>은 초기 연결가중치를 100번 변화시켜가면서 학습시킨 결과, 최적해로 수렴된 경우들에 대한 반복회수 및 CPU 시간의 평균 m 과 표준편차 σ 를 나타낸 것이다. 이 결과는 영상데이터 1에 대해 전역최소점으로 수렴된 결과들만을 나타내었고, 은닉층 뉴런의

개수와 학습율은 모든 방법에서 12와 0.001로 설정하였다. 각 시도는 랜덤시드를 변화시켜 수행하였다. N_{gh} , N_{nh} , N_{tp} , N_{pa} 와 N_{ca} 는 각각 Sanger 방법, Foldiak 방법, 기존의 역전파 알고리즘, 제안된 알고리즘, 그리고 동적터널링의 역전파 알고리즘 반복회수를 나타내며, t_{gh} , t_{nh} , t_{tp} 와 t_{pa} 는 Sanger 방법, Foldiak의 방법, 역전파와 제안된 알고리즘의 실행 CPU 시간을 나타낸 것이다. 표에서 보는 것처럼, 제안된 알고리즘을 이용한 영상의 주요특징 추출이 단층신경망을 이용하는 Sanger 방법과 Foldiak의 방법, 그리고 다층신경망을 이용하는 기존의 역전파 알고리즘보다 최적해로의 수렴속도와 수렴확률면에서 더욱 우수한 성능이 있음을 알 수 있다. 표에서 보면 단층신경망을 이용하는 경우에도 출력 뉴런들간에 측면연결을 갖는 Foldiak 방법이 Sanger 방법보다는 더 빠른 수렴속도와 약간 높은 수렴율을 가짐을 확인할 수 있다. 또한 영상데이터 2와 3의 경우 최적해로 수렴된 경우만을 고려할 때, 단층신경망을 이용하는 것이 다층신경망을 이용하는 것보다 빠른 수렴속성을 보여 이는 해비안규칙이 데이터 상호간의 상관관계에 비례하여 학습됨을 알 수 있다. 특히, 실험의 결과를 통해 제안된 알고리즘이 역전파 알고리즘보다 초기 가중치에 대한 의존도도 적음을 추정할 수 있다. 해비안규칙에 기초한 알고리즘인 Sanger 방법과 Foldiak 방법 및 기울기하강법에 기초한 역전파 알고리즘의 경우 최적해로의 수렴률이 초기 연결가중치의 설정에 의존함을 알 수 있다. 그러나 제안된 학습알고리즘의 경우는 이상과 같은 규모의 문제에 대해서는 모두 최적해로 수렴되었으며, 이는 결국 기존의 Sanger 방법, Foldiak 방법, 그리고 역전파 알고리즘에 비해 초기 연결가중치에 대한 의존도가 낮음을 알 수 있다.

4.2 Lenna 영상

128×128 픽셀의 8 비트(256 회색단계) Lenna 영상을 대상으로 데이터의 주요 특징추출을 위해 단층 및 다층신경망을 각각 이용하였다. 이용된 영상은 크기 8×8(64)의 비중복 블록으로 분할하였다. 이는 전체 128개의 학습표본을 가지도록 행과 열배열을 이용하여 크기 64의 1 차원 입력벡터의 연속으로 정렬하였다. 블록의 크기에 따라 일반적으로 주요 특징추출의 정확도와 알고리즘의 계산적 및 구조적 복잡도는 상관관계를 가진다. 그러나 대부분의 실 영상에서 이웃하는 픽셀들 사이에는 매우 많은 상관성을 가지며 픽셀사이의 공간적 거리에 따라 상관도도 대체로 비례하여 감소한다. 경험적으로 8×8(64)의 블록크기가 좋은 결과를 가지는 것으로 알려져 있다^[3]. 따라서 실험에 이용된 다층신경망의 입력뉴런과 출력뉴런 및 단층신경망의 입력뉴런 수는 각각 64개로 하였으며 허용 오차함수값은 10^{-3} 으로 하였다.

<표 2>는 은닉층 뉴런의 개수를 12로 고정하고 초기 연결가중치를 100번 변화시켜 학습시킨 결과, 최적해로 수렴된 경우들에 대한 반복 회수 및 CPU 시간의 평균 m 과 표준편차 σ 를 나타낸 것이다. 여기서 학습율은 모든 학습방법에 대해서 0.001로 고정하였다. 표 1의 결과에서 처럼, 표 2에서도 제안된 알고리즘을 이용한 영상의 주요특징 추출이 단층신경망을 이용하는 Sanger 방법과 Foldiak의 방법, 그리고 다층신경망을 이용하는 기존의 역전파 알고리즘보다 최적해로의 수렴속도와 수렴확률면에서 더욱 우수한 성능이 있음을 알 수 있다. 한편, SNR의 값들을 비교하면 다층신경망을 이용하는 경우가 단층신경망을 이용하는 경우보다 더 큰 값을 가짐을 알 수 있다. 특히 <표 1>과 <표

<표 1> 12×12 픽셀을 가진 표본 영상데이터에 대한 100번 시도에 따른 실험 결과

Image Input	m	Sanger Method					Foldiak Method					BP Algorithm					Proposed Algorithm				
		N_{gh}	t_{gh}	MSE	SNR	P_c	N_{nh}	t_{nh}	MSE	SNR	P_c	N_{tp}	t_{tp}	MSE	SNR	P_c	N_{pa} (N_{gh} , N_{nh})	t_{pa}	MSE	SNR	P_c
Data 1	m	324.98	0.5	9.85×10^{-5}	34.76	87	214.43	0.34	9.48×10^{-5}	34.94	88	100.55	0.31	9.86×10^{-5}	34.75	92	93.49(91.2, 49)	0.31	9.84×10^{-5}	34.76	100
	σ	38.99	0.51	1.03×10^{-6}	0.05		83.01	0.48	7.89×10^{-6}	0.42		70.21	0.49	1.15×10^{-5}	0.05		32.11(31.17, 4.29)	0.49	1.79×10^{-5}	0.08	
Data 2	m	230.09	0.36	9.76×10^{-5}	35.99	85	153.57	0.25	9.21×10^{-5}	36.33	87	287.81	0.89	9.95×10^{-5}	35.91	94	221.82(212.29, 9.53)	0.78	9.93×10^{-5}	35.92	100
	σ	44.60	0.48	1.45×10^{-6}	0.06		57.93	0.44	1.33×10^{-5}	1.06		85.23	0.80	6.89×10^{-7}	0.03		69.24(57.65, 19.51)	0.54	7.44×10^{-7}	0.03	
Data 3	m	150.13	0.24	9.61×10^{-5}	37.68	86	95.46	0.15	8.93×10^{-5}	38.15	88	338.22	1.04	9.95×10^{-5}	37.53	93	252.42(248.54, 3.88)	0.92	9.94×10^{-5}	37.54	100
	σ	36.66	0.43	2.22×10^{-6}	0.10		44.60	0.36	1.65×10^{-5}	1.37		124.52	1.10	7.03×10^{-7}	0.03		112.35(101.48, 3.72)	0.73	7.27×10^{-7}	0.03	

m : mean, σ : standard deviation, t_c : CPU time(sec), MSE : mean square error, SNR : signal to noise ratio, P_c : convergence ratio(%)

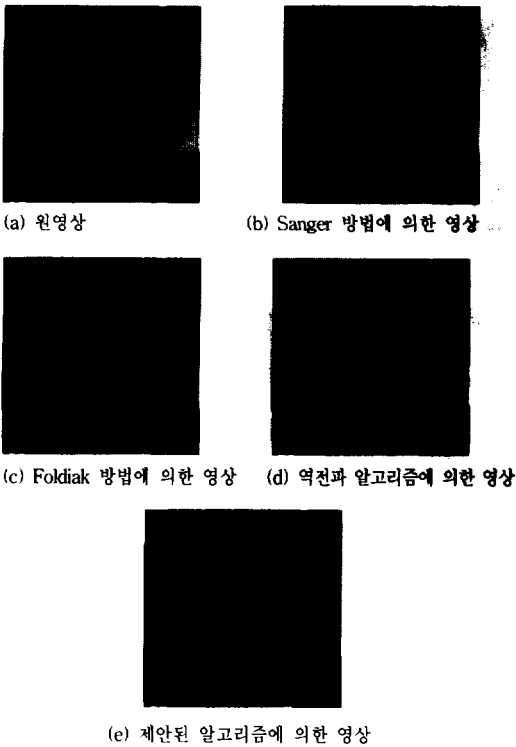
〈표 2〉 128×128 Lenna 영상데이터에 대한 100번 시도에 따른 실험 결과

	Sanger Method		Foldiak Method		BP Algorithm		Proposed Algorithm	
	m	σ	m	σ	m	σ	m	σ
N	1795.2	134.11	1543	374.56	1140.6	96.55	1103.4	113.83
t	415	31.13	402.6	107.67	595.2	44.46	528.0	60.18
MSE	0.004813	8.85×10^{-5}	0.004942	7.34×10^{-5}	0.004757	4.23×10^{-5}	0.004798	3.71×10^{-6}
SNR	14.53	0.06	14.54	0.05	14.57	0.03	14.57	0.04
P_c	42		54		75		89	

m : mean, σ : standard deviation, t : CPU time(sec), MSE : mean square error, SNR : signal to noise ratio, p_c : convergence ratio(%)

2>를 비교하면, 문제의 규모가 크질수록 제안된 방법의 성능개선 정도가 더 우수함을 확인할 수 있다. 이는 문제의 규모가 크질수록 초기 연결가중치의 설정은 더욱 어려우며 이에 따른 학습성능의 의존도도 더욱 크짐을 알 수 있다. 또한 표에서는 제안된 알고리즘을 이용할 경우에 89번의 시도에 대해서만 최적해로 수렴되었음을 알 수 있다. 그러나 나머지 11번의 시도는 학습알고리즘의 반복수를 20,000번 이상 더욱 오래 수행시키면 가능함을 확인할 수 있었다.

(그림 8)은 128×128 픽셀의 Lenna 영상에 대해 각각의 방법으로 복원된 영상을 나타낸 것이다. (그림 8)의 (a)는 실험에 이용된 128×128 픽셀의 Lenna 원영상이고, (b), (c), (d), 그리고 (e)는 각각 12 개의 주요 특징추출 후, 이들을 이용하여 재구성된 영상을 보여준다. 그림에서 보면, Sanger 방법에 의해 추출된 12 개의 주요 특징들로부터 재구성된 결과 영상 (b)가 가장 심한 왜곡이 있으며 잡음도 다소 존재하고 있음을 알 수 있다. 또한 제안된 방법에 의해 얻어진 특징으로 재구성된 영상 (e)는 다른 방법에 의해 재구성된 영상들보다 시각적인 왜곡이 적음을 알 수 있으며, 역전파 알고리즘을 이용하는 경우가 다른 다층신경망을 이용한 경우보다 더 작은 왜곡을 가짐도 확인할 수 있다. 이는 은닉층을 가지는 다층신경망이 단층신경망에 비해 비선형 사상(nonlinear mapping)의 속성을 가지므로 인식이나 분류 등에서 좀 더 우수한 특성이 있음을 나타내는 것으로 추측된다. 따라서 제안된 학습알고리즘의 다층신경망을 이용하면 효과적으로 입력영상의 주요 특징들을 추출할 수 있으며, 또한 영상의 압축효과도 아주 우수함을 알 수 있다.



(그림 8) 4 가지 학습 알고리즘을 이용한 12 개의 추출된 특징들로부터 복원된 128×128 픽셀의 Lenna 영상들

5. 결 론

본 논문에서는 새로운 학습 알고리즘의 3층 전향 신경망을 이용한 입력데이터의 주요 특징추출에 대해서 제안하였다. 제안된 학습알고리즘은 기울기하강과 동적터널링의 속성을 조합한 역전파 알고리즘이다. 여기서 기울기하강의 속성은 빠른 수렴을 위한 것이고, 동적터널링의 속성은 국소최적해로부터 탈출하여 전역 최적해로의 수렴을 위한 것이다.

제안된 학습알고리즘을 이용한 다층신경망을 12×12 픽셀(pixel)의 영상데이터들과 128×128 픽셀의 Lenna 영상데이터를 대상으로 시뮬레이션한 결과, 기존의 단

층신경망을 이용하는 Sanger 방법이나 측면연결을 가지는 단층신경망을 이용하는 Foldiak 방법 및 기술기하강에 기초를 둔 역전과 알고리즘을 이용한 다층신경망에 의한 결과와 각각 비교할 때 더욱 우수한 수렴성능과 특징 추출성능이 있음을 확인할 수 있었다. 특히, 초기 연결가중치의 설정에 대한 융통성도 더욱 우수함을 알 수 있었다.

향후 제안된 학습 알고리즘의 다층신경망의 병렬구현에 대한 연구가 계속 진행되어야 하며, 통신이나 영상 인식 등과 같은 좀 더 많은 분야에의 응용에 대한 연구가 남아 있다.

참 고 문 헌

[1] A. Cichock and R. Unbehauen, 'Neural Networks for Optimization and Signal Processing,' John Wiley & Sons., New York, 1993.

[2] P. Foldiak, "Adaptive Network for Optimal Linear Feature Extraction," International Joint Conference on Neural Networks, Washington D. C., Vol.1, pp.401-406, June 1989.

[3] S. Bannour, A. Mahmood, and A. Sadjadi, "Principal Component Extraction Using Recursive Least Squares Learning," IEEE Trans. on Neural Networks, Vol.6, No.2, pp.457-469, March 1995.

[4] K. I. Diamantaras and S. Y. Kung, 'Principal Component Neural Networks : Theory and Applications, Adaptive and Learning Systems for Signal Processing, Communications, and Control,' John Wiley & Sons, Inc., 1996.

[5] N. Sonehara, M. Kawato, S. Miyake, and K. Nakane, "Image Data Compression Using Neural Network Model," International Joint Conference

on Neural Networks, Washington D. C., Vol.2, pp.35-41, June 1989.

[6] M. Arozullah and A. Namphol, "A Data Compression System Using Neural Network Based Architecture," International Joint Conference on Neural Networks, San Diego, Vol.1, pp.531-536, June 1990.

[7] D. E. Rumelhart and J. L. McClelland, 'Parallel Distributed Processing,' MIT Press, Cambridge, MA., 1986.

[8] J. A. Freeman and D. M. Skapura, 'Neural Networks : Algorithms, Applications, and Programming Techniques,' Addison Wesley, London, 1991.

[9] Y. Yao, "Dynamic Tunneling Algorithm for Global Optimization," IEEE Trans. On Systems, Man, and Cybernetics, Vol.19, No.5, pp.1222-1230, Sept/Oct. 1989.



조 용 현

e-mail : yhcho@cuth.cataegu.ac.kr

1979년 2월 경북대학교 공과대학 전자공학과(공학사)

1981년 2월 경북대학교 대학원 전자공학과(공학석사)

1993년 2월 경북대학교 대학원 전자공학과(공학박사)

1983년 9월~1984년 2월 삼성전자(주) 사원

1984년 3월~1987년 2월 한국전자통신연구소 연구원

1987년 3월~1997년 2월 영남전문대학 전자과 부교수

1997년 3월~현재 대구효성가톨릭대학교 전자·정보공학부 조교수

관심분야 : 신경망, 병렬분산처리, 신호처리, 영상인식, 교환기 등