

신경망 기반 독립성분분석을 위한 효율적인 학습알고리즘

박용수*, 조용현

대구가톨릭대학교 공과대학 컴퓨터정보통신공학부
e-mail:yhcho@cuth.cataegu.ac.kr

An Efficient Learning Algorithm for Independent Component Analysis Based on Neural Networks

Yong-Soo Park*, Yong-Hyun Cho
School of Computer and Information Comm. Eng.,
Catholic Univ. of Daegu

요약

본 연구에서는 효율적인 학습알고리즘을 가지는 신경망 기반 독립성분분석 기법을 제안하였다. 제안된 기법은 할선법에 기초를 둔 fixed point 알고리즘의 신경망 기반 독립성분분석 기법이다. 여기서 할선법은 독립성분 상호간의 정보를 최소화하기 위해 negentropy를 최대화하는 과정에서 요구되는 1차 미분에 따른 계산량을 줄이기 위함이다. 제안된 기법을 500개의 데이터를 가지는 4개 신호들로부터 임의의 혼합 행렬에 따라 발생하는 혼합신호들을 각각 대상으로 시뮬레이션 한 결과, 우수한 분리성능과 빠른 학습 속도가 있음을 확인하였다.

1. 서론

통계학, 신호처리, 그리고 신경망 등에서 적당한 변환에 의해 알맞은 데이터의 표현을 찾기 위하여 일반적으로 선형변환을 이용한다. 여러 가지 방법과 원리들이 적당한 선형변환을 찾기 위해서 개발되었으나, 그 중에서도 주요성분분석(principal component analysis : PCA)^[1-3]이 가장 널리 이용되고 있다. 이는 선형변환 기법으로 공분산(covariance)을 이용하는 2차원 방법이다. 특히 PCA는 입력 데이터의 공분산에 의해 결정되는 orthonormal basis로 표현되며, 가우스(Gaussian) 변수의 모든 정보는 공분산행렬 내에 포함되기 때문에 가우스 분포의 입력 데이터에 대해서는 적당한 것이다. 하지만 일반적으로 통신이나 신호 및 영상 처리 등에서 입력 데이터에는 비가우스 분포도 있으며, 이러한 경우에도 효과적으로 적용될 수 있는 선형변환 방법의 연구가 요구된다.

최근 PCA의 제약을 해결하기 위한 대안으로 독립 성분분석(independent component analysis : ICA)이 제안되었으며, 이는 데이터의 통계적 성질을 고려한 선형변환 기법으로 데이터의 분포정보를 이용하는 다차원 방법이다^[4-6]. 하지만 알려진 수치적 기법의 ICA 알고리즘들에서는 데이터의 통계적 독립성을 판정하기 위해 확률밀도함수와 역행렬을 구한다든지 고차원의 cumulant나 moment를 구하는 등 복잡한 계산이 요구된다. 특히 ICA에서 신호모델을 추정하는 것은 혼합된

출력신호들 사이에 포함된 정보만을 이용하여 혼합행렬을 추정하는 것으로 이를 위해 출력신호 상호간의 결합밀도함수(joint density function)를 이용하며, 그 함수의 계산은 매우 복잡하다. 문제의 규모가 커질수록 계산량은 기하급수적으로 증가되고 그에 따른 계산의 복잡성과 소요시간도 증가되어 비현실적이다.

최근 입력과 출력 사이에 선형성을 가지는 간단한 신경망을 이용하여 해결하려는 연구가 활발하게 진행되고 있다^[5,6]. 지금까지 발표된 신경망을 이용한 ICA 방법으로는 엔트로피를 최소화하는 것에 기초를 둔 fixed point(FP) 알고리즘, 비선형 PCA에 기초한 recursive least square 알고리즘, 결합엔트로피를 최대화하기 위하여 확률적 기울기를 이용하는 Bell & Sejnowski(BS) 알고리즘, natural gradient 알고리즘 등이 제안되었다^[5].

본 연구에서는 효율적인 새로운 학습알고리즘의 신경망 기반 독립성분분석 기법을 제안한다. 제안된 기법에서는 엔트로피 최적화 방법에서 유도된 빠른 학습속도의 FP 알고리즘에서 그 계산을 간략화 하기 위해 할선법(Secant method)^[7]을 도입함으로써 더욱 빠른 학습성능을 가지도록 하였다. 제안된 기법을 500개의 데이터를 가지는 4개의 신호들로부터 임의의 혼합 행렬에 따라 발생하는 혼합신호들을 각각 대상으로 실험하여 성능을 확인하였다.

2. 효율적인 학습기법의 신경망 기반 ICA기법

ICA는 m 개의 입력신호 s 가 선형적으로 혼합된 n 개의 신호 x 가 알려져 있을 때, 혼합된 신호로부터 역으로 m 개의 독립인 입력신호를 찾는 기법이다. 여기서 입력신호들을 혼합하는데 이용된 혼합행렬 A 는 알려져 있지 않으며, 혼합과정에서 잡음 n 이 추가 될 수도 있다. 이때 혼합신호와 입력신호와의 관계는

$$x = As + n = \sum_{i=1}^m s(i)a(i) + n \quad (1)$$

로 표현된다. 여기서 잡음 n 은 보통 입력신호와 구별되지 않기 때문에 생략할 수도 있으며, $A=[a(1), a(2), \dots, a(m)]$ 으로 $a(i)$ 는 ICA의 basis vector이다. 결국 ICA는 알려진 혼합신호로부터 혼합행렬의 역행렬 A^{-1} 을 찾는 기법이다. 이상의 관계를 요약하면 그림 1과 같다. 그림에서 보면 혼합기(mixer)의 출력신호인 혼합신호는 알 수 있으나 혼합기의 특성인 A 와 입력신호인 원 신호 s 는 알려져 있지 않다. 결과적으로 입력신호와 출력신호가 일치하도록 하는 즉, $y = Wx = WA s$ 에서 $W=A^{-1}$ 의 관계를 만족토록 하는 것이다.

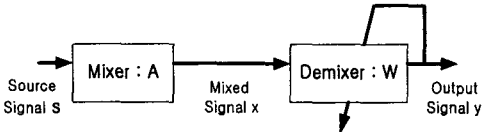


그림 1. ICA를 위한 구성도

한편, 그림 1로부터 혼합행렬 A 와 역혼합행렬 W 에 대하여 좀 더 상세히 살펴보면 다음 그림 2와 같은 구성도로 나타낼 수 있다. 그림에서 $x = As$ 이고, $y = Wx$ 이다. 이때 ICA는 출력신호 y 의 각 요소들이 통계적으로 독립성을 유지해야 하며 직교성의 제약조건은 포함하지 않는다. 따라서 ICA는 알려진 혼합신호 x 로부터 출력신호 y 를 찾는 기법으로 궁극적으로는 역혼합행렬 W 을 찾아서 원 신호 s 의 근사값을 알아내는 것이다. 여기서 영평균과 whitening된 혼합신호를 이용하면, 신호의 1차와 2차의 통계성에 따른 상관성이 이미 고려되어 좀 더 효율적인 ICA를 수행할 수 있다.

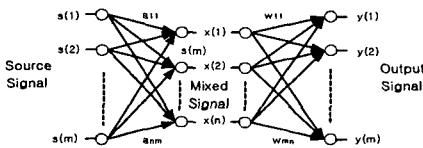


그림 2. 혼합행렬과 역혼합행렬의 상세 설명도

2.1 신호의 영평균과 Whitening

신호의 영평균은 데이터의 1차적 통계성을 고려한 정규화로 학습알고리즘을 간략하게 하는 기법이다^[5,6]. 이는 신호 데이터 벡터 x 에서 평균값 \bar{x} 를 뺀 차를 구함으로써 간단하게 구할 수 있다. 즉, 영평균된 새로운 데이터 벡터 x 는

$$x = x - \bar{x} \quad (2)$$

이다.

한편, 데이터의 2차적 통계성을 고려한 상호간의 의존성을 줄이기 위한 기법으로 whitening이 이용되

고 있다. whitening은 데이터 벡터 x 의 공분산행렬이 단위행렬값을 갖도록 함으로써 구할 수 있다. 즉, $E(xx^T)=I$ 가 되도록 한다. 이렇게 함으로써 whitening된 벡터의 성분들 상호간의 상관성이 줄어들어 강한 독립성분이 된다. PCA는 공분산행렬의 고유벡터와 고유치를 추정하는 수치적 기법으로 이루어진다. 일반적인 PCA whitening 행렬 V 는

$$V = D^{(-1/2)}E^T \quad (3)$$

이다. 여기서 D 는 고유값 λ 로 구성되는 대각행렬이며, E 는 고유벡터 c 로 구성되는 직교행렬이다. 본 연구에서는 혼합신호로부터 원 신호를 추출하기 전에 신호데이터를 영평균 정규화하고, 표준 PCA 기법으로 whitening하여 이용하였다.

2.2 할선법에 기초한 Fixed Point 알고리즘

독립성분분석을 위한 신경망 기반 기법들은 결합 엔트로피를 최대화하거나 엔트로피를 최소화하는 기법으로 제안되었다^[5,6]. 결합엔트로피를 최대화하기 위하여 기울기를 이용하는 알고리즘은 원 신호를 더 잘 근사하지만 학습율과 같은 파라미터의 설정이 경험적이며, 그에 따른 수렴속도도 느린 제약을 가진다. 특히 학습과정에서 학습율의 설정에 따라서는 그 수렴이 느려지거나 알고리즘이 수렴되지 않는 경우도 발생된다. 따라서 이런 제약들을 해결하면서도 신경망이 가지는 병렬성과 분산성, 그리고 더 작은 메모리 요구 등의 제약을 해결하기 위해 FP 알고리즘이 제안되었다^[3]. FP 알고리즘은 엔트로피 최적화 방법으로부터 유도된 기법으로 그 수렴속도가 적어도 2차 합수정도로 지금까지 알려진 기법 중 가장 빠른 학습속도를 가진다. 이는 신호 내에 포함된 상호정보를 고려한 것으로 그 척도는 독립성의 거리이다. 결국 상호정보를 최소화함으로써 ICA의 해를 구하는 기법이다.

상호정보는 랜덤변수 상호간의 독립성의 정보이론적 척도로 ICA 변환을 찾기 위한 판별로서 이용될 수 있다. 일반적으로 변수들 상호간의 정보를 표현하는 상호정보는 비가우시성의 척도로 이용되는 negentropy로 정의된다. n 개의 랜덤변수 벡터 $y = (y_1, \dots, y_n)^T$ 를 위한 negentropy $J(y)$ 는

$$J(y) = H(y_{gauss}) - H(y) \quad (4)$$

이다. 여기서 $H(y)$ 는 랜덤변수 y 의 엔트로피(entropy)이며, y_{gauss} 는 가우시 랜덤벡터이다. 이때 negentropy $J(y)$ 를 이용한 상호무관(decorrelation)한 n 개 랜덤변수 $y_i(i=1, \dots, n)$ 사이의 상호정보 $I(y_1, y_2, \dots, y_n)$ 는

$$I(y_1, y_2, \dots, y_n) = J(y) - \sum_{i=1}^n J(y_i) \quad (5)$$

로 정의된다. 결국 ICA에서 독립성을 최대화하기 위해 식 (5)로 정의되는 상호정보 $I(y_1, y_2, \dots, y_n)$ 를 최소화하는 것은 negentropy를 최대화하는 것이다. 여기서 negentropy의 최대화는 식 (5)의 우측 2번째 항의 최대화로 가능하다. 이를 위한 근사식으로 최대 엔트로피 원리를 이용한 식은

$$J(y_i) = c[E\{G(y_i)\} - E\{G(\nu)\}]^2 \quad (6)$$

와 같이 정의된다. 여기서 G 와 c 는 각각 비2차(non-quadratic) 함수와 상수이고, ν 는 평균이 0이고 분산이 1인 표준화된 가우스 변수이며, 랜덤변수 y_i 도 평균이 0이고 분산이 1인 변수로 가정한다. negentropy의 근사화를 위한 식 (6)은 ICA 변환을 추정하기 위한 목적함수로 이용될 수 있다. 따라서 ICA에서 독립변수들의 상호간 무관성을 고려한 하나의 독립성분 $y_i = \mathbf{W}^T \mathbf{x}$ 를 찾기 위한 목적함수 $J_G(\mathbf{W})$ 는 식 (6)에 따라

$$J_G(\mathbf{W}) = [E\{G(\mathbf{W}^T \mathbf{x})\} - E\{G(\nu)\}]^2 \quad (7)$$

로 표현된다. 이를 전체 n 개의 독립성분 모두에 대해서 확장하면

$$\begin{aligned} & \text{maximize } \sum_{i=1}^n J_G(\mathbf{w}_i) \quad \text{w.r.t. } \mathbf{w}_i, i=1,2,\dots,n \\ & \text{subject to } E\{(\mathbf{w}_k^T \mathbf{x})(\mathbf{w}_j^T \mathbf{x})\} = \delta_{jk} \end{aligned} \quad (8)$$

이 된다. 여기서 \mathbf{w}_i 는 역혼합행렬 \mathbf{W} 의 행벡터 중 하나이며, ICA 변환은 $\mathbf{s} = \mathbf{W}\mathbf{x}$ 로 얻어진다. 결국 식 (7)이나 (8)로 정의된 ICA를 위한 목적함수를 최대화하기 위한 방법이 요구된다. 이러한 방법으로 가장 간단한 알고리즘은 기울기하강(gradient descent)에 기초를 둔 기법들이나, 이 방법에 기초한 알고리즘들의 제약들을 해결하기 위한 대안으로 FP 알고리즘이 제안되었다^[5].

FP 알고리즘은 알고리즘의 한 단계에 많은 수의 데이터들이 이용되는 batch mode로 계산을 수행하며, 신경망으로 수행될 수 있는 알고리즘이다. 데이터 벡터 \mathbf{x} 의 상관행렬 $E\{\mathbf{x}\mathbf{x}^T\}=\mathbf{I}$ 로 whitening되어 있다고 가정할 때, 하나의 독립성분을 위한 식 (7)로부터 FP 알고리즘을 유도하면 다음과 같다. 식 (7)에서 $J_G(\mathbf{W})$ 의 최대화는 반드시 $E\{G(\mathbf{W}^T \mathbf{x})\}$ 의 최적에서 얻어진다. 이때 Kuhn-Tucker 조건으로부터, $E\{(\mathbf{W}^T \mathbf{x})^2\} = \|\mathbf{W}\|^2 = 1$ 조건하에서 $E\{G(\mathbf{W}^T \mathbf{x})\}$ 의 최적은 다음의 값

$$E\{\mathbf{x}g(\mathbf{W}^T \mathbf{x})\} - \beta \mathbf{W} = 0 \quad (9)$$

으로 얻어진다. 여기서 $g(\cdot)$ 는 비2차 함수 G 의 미분이고, $\beta = E\{\mathbf{W}_0^T \mathbf{x}g(\mathbf{W}_0^T \mathbf{x})\}$ 인 상수이며 \mathbf{W}_0 는 \mathbf{W} 의 최적값이다. 따라서 방정식 (9)를 풀기 위해 빠른 수렴속도를 가지는 할선법^[7]을 도입한다.

할선법은 현재의 함수 값과 바로 이전의 함수 값만을 이용함으로써 해를 구할 수 있어 기존의 Newton 기법에 기반을 둔 알고리즘에서처럼 별도의 1차 미분 계산이 요구되지 않는다. 이는 Newton 기법에 비해 할선법이 반복에 소요되는 시간이 적게 걸리는 근거이다. 따라서 β 를 \mathbf{W}_0 대신에 \mathbf{W} 의 현재값으로 나타낸 근사화된 반복방법의 할선법에 기초한 역혼합행렬 \mathbf{W} 를 구하는 식은

$$\begin{aligned} f(\mathbf{W}^*) &= [E\{\mathbf{x}g(\mathbf{W}^{*T} \mathbf{x})\} - \beta \mathbf{W}^*] \\ f(\mathbf{W}) &= [E\{\mathbf{x}g(\mathbf{W}^T \mathbf{x})\} - \beta \mathbf{W}] \\ \mathbf{W}^* &= \mathbf{W} - f(\mathbf{W})/[f(\mathbf{W}) - f(\mathbf{W}^*)] \\ \mathbf{W}^* &= \mathbf{W}^*/\|\mathbf{W}^*\| \end{aligned} \quad (10)$$

와 같다. 여기서 \mathbf{W}^* 는 \mathbf{W} 의 새롭게 경신된 값이며, \mathbf{W}^* 은 바로 전에 계산된 \mathbf{W} 의 값이고, $\beta = E\{\mathbf{W}^T \mathbf{x}g(\mathbf{W}^T \mathbf{x})\}$ 이다. 결국 식 (10)은 할선법에 기초를 둔 ICA를 위한 FP 알고리즘이다. 여기서는 역혼합행렬 \mathbf{W} 를 구하는데 별도의 1차 미분인 Jacobian matrix의 계산이 요구되지 않아 보다 빠른 계산시간을 가진다.

제안된 할선법에 기초를 둔 FP 알고리즘의 역혼합행렬 \mathbf{W} 를 경신하는 과정을 요약하면 다음과 같다.

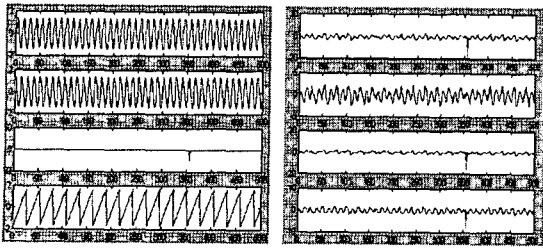
- 단계 1 : $\mathbf{W}(0) = \text{rand}()$ 로 설정한다.
- 단계 2 : $\mathbf{W} = \mathbf{W}(0)/\|\mathbf{W}(0)\|$ 의 값을 계산한다.
- 단계 3 : $\mathbf{W}^* = \mathbf{0}$ 로 설정한다.
- 단계 4 : $\|\mathbf{W} - \mathbf{W}^*\| > \epsilon$ 이고 $\|\mathbf{W} + \mathbf{W}^*\| > \epsilon$ 이면 다음의 식에 따라 \mathbf{W} 를 계산하고, 그렇지 않으면 과정을 종료한다.
 - (1) β , $f(\mathbf{W}^*)$, 그리고 $f(\mathbf{W})$ 를 계산한다.
 - (2) $\mathbf{W}^* = \mathbf{W} - f(\mathbf{W})/[f(\mathbf{W}) - f(\mathbf{W}^*)]$ 를 계산한다.
 - (3) $\mathbf{W}^* = \mathbf{W}$ 로 둔다.
 - (4) $\mathbf{W} = \mathbf{W}^*/\|\mathbf{W}^*\|$ 를 계산한다.

위의 과정에서 벡터 \mathbf{x} 는 whitening된 신호벡터이며, ϵ 은 1보다 작은 임의의 값이다. 특히 경신의 식에서 $g(\cdot)$ 는 비선형함수로 일반적으로 $(\cdot)^3$ 과 $\tanh(\cdot)$ 의 함수값을 가진다. 비선형함수의 선택은 $\tanh(\cdot)$ 의 대칭적인 알고리즘의 경우에 supergaussian의 분포특성을 가지는 신호들은 분리할 수 없다는 것만 제외하면 어느 함수나 무방하다. 그러나 일반적으로 $(\cdot)^3$ 함수는 학습속도가 빠르며, $\tanh(\cdot)$ 함수는 좀 더 안정한 특성을 가진다. 본 연구에서는 상대적으로 이들 함수에 따른 분리성능은 거의 유사한 특성을 가짐을 확인하였으며, 여기서는 $\epsilon = 0.0001$ 과 $(\cdot)^3$ 함수를 적용한 결과들을 나타낸다.

3. 시뮬레이션 결과 및 분석

제안된 학습알고리즘의 신경망 기반 ICA의 타당성과 성능을 평가하기 위해서 인공적으로 생성된 4개의 신호를 각각 대상으로 시뮬레이션 하였다. 시뮬레이션에서는 500개의 데이터를 가지는 4개의 독립된 신호를 대상으로 임의의 값을 가진 혼합행렬로부터 생성된 신호들로 제안된 알고리즘을 수행하였다. 복원 성능을 평가하기 위해 각 신호의 절대합오차(absolute sum error : ASE)를 정의하여 원 신호와 복원된 신호의 정합성을 각각 비교하였다. 이때 절대합오차 $ASE = \frac{1}{500} \sum_{i=1}^{500} |x_i - y_i|$ 이며, 여기서 x_i 와 y_i 는 각각 원 신호와 복원된 신호의 i 번째 값이다. 시뮬레이션은 펜티엄 IV-1.5G 컴퓨터 상에서 Matlab 5.1로 실행하였다.

그림 3은 원 신호와 임의의 혼합행렬로부터 생성된 혼합신호 4개를 각각 나타낸 것이다. 그림 3(a)의 원 신호는 sin, cos, tan, 그리고 ramp 함수이다. 또한 그림 3(b)의 혼합신호는 4x4의 임의의 랜덤값을 가지는 혼합행렬 \mathbf{A} 로부터 생성되며, 그 각각에서는 원 신호들을 발견할 수 없다. 실제에서 혼합신호들은 알려져 있으나 원 신호들은 미지의 신호들이다. 시뮬레이션에서는 혼합된 신호들을 입력신호로 하여 ICA를 수행한다.



(a) 원 신호 (b) 혼합된 신호

그림 3. 500개의 4개 독립된 신호

그림 4는 그림 3(b)의 혼합신호들을 대상으로 할선법을 이용한 제안된 FP 알고리즘의 ICA를 수행한 후 복원된 4개 신호들이다. 그림에서 원 신호와 복원된 신호들 사이의 순서는 일치되지 않음을 알 수 있다. 이는 수행과정에서 신호의 순차를 고려하지 않는 ICA의 속성이 반영되었기 때문이다. 여기서 원 신호와 복원된 신호를 순서에 맞추어 서로 비교하면 육안으로 상당히 유사함을 알 수 있다. 복원된 신호는 각각 혼합된 신호의 독립성분이며, 여기서 4개의 sin, cos, tan, 그리고 ramp 함수 각각을 찾아내는데 소요되는 알고리즘의 반복수는 각각 37회, 3회, 9회, 그리고 8회 수행되었다. 이는 500개 데이터를 가지는 4개의 분리신호를 찾는 데 총 57회 알고리즘이 반복수행 되었음을 의미한다. 특히 그림에서 cos 함수의 경우 복원된 신호는 원 신호와 반대의 부호를 가진다. 이 또한 ICA에서 독립성분을 추출할 때 순서뿐만 아니라 부호도 고려되지 않는 속성이 반영되었기 때문이다.

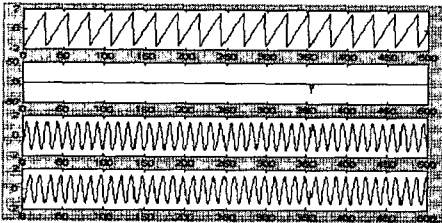


그림 4. 500개의 4개 복원된 신호

표 1은 20번의 실험을 통하여 각각의 대응되는 원 신호와 복원된 신호에 대해 데이터단위로 그 절대합 오차와 학습반복수의 평균을 각각 나타낸 것이다. 표에서 보면 sin과 cos 함수의 경우 절대합오차로 표현되는 정도는 0.28로 동일하며, 학습반복수에서도 각각 31과 35로 비슷한 값을 가진다. 또한 tan과 ramp 함수의 정도는 0.04로 동일하며, 학습반복수에 있어서는 12와 26회로 많은 차이를 보인다. 여기서 보면 sin과 cos 함수가 tan과 ramp 함수보다 상대적으로 큰 값의 ASE를 가지며, 학습반복수에 있어서는 sin, cos, 그리고 ramp 함수가 tan 함수보다 상대적으로 큰 값을 가진다. 이는 변화가 상대적으로 심한 신호일수록 복원에 따른 오차와 학습반복수는 증가됨을 확인할 수 있다. 결국 학습반복수가 증가되면 절대합오차는 감소하여 이들의 관계는 상호 반비례함도 알 수 있다. 따라서 제안된 기법의 ICA를 이용하면 평균적으로 절대합오차는 약 0.16

정도, 학습반복수는 약 26회 정도에서 500개의 데이터를 가지는 4개의 신호를 복원할 수 있다.

표 1. 원 신호와 복원된 신호의 절대합오차와 학습반복수

신호	절대합오차	학습반복수
sin 함수	0.28	31
cos 함수	0.28	35
tan 함수	0.04	12
ramp 함수	0.04	26
평균	0.16	26

4. 결 론

본 연구에서는 효율적인 학습알고리즘을 가지는 신경망 기반 독립성분분석 기법을 위해 할선법에 기초를 둔 효과적인 fixed point 알고리즘을 제안하였다. 여기서 할선법은 상호정보를 최소화하기 위해 negentropy를 최대화하는 과정에서 1차 미분의 계산 대신 현재와 바로 이전 함수의 값만을 이용함으로써 계산량을 줄이기 위함이다.

제안된 신경망 기반 독립성분분석 기법으로 500개의 데이터를 가지는 4개 신호들로부터 임의의 혼합행렬에 따라 발생하는 혼합신호들을 각각 대상으로 실험한 결과, 우수한 분리성능과 빠른 학습속도가 있음을 확인하였다.

향후 제안된 알고리즘의 독립성분분석 기법을 다양한 분야에 좀 더 큰 규모의 문제와 기존의 fixed point 알고리즘과의 비교연구가 뒤따라야 할 것이다.

[참 고 문 헌]

- [1] 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
- [2] S. Haykin, Neural Networks : A Comprehensive Foundation, Prentice-Hall, 2ed, London, 1999
- [3] J. Karhunen and J. Joutsensalo, "Generation of Principal Component Analysis, Optimization Problems, and Neural Networks," Neural Networks, Vol. 8, No. 4, pp. 549-562, 1995
- [4] P. Comon, "Independent Component Analysis -A New Concept?," Signal Processing, vol.36, No.3, pp. 287-314, Apr. 1994
- [5] T.W. Lee, Independent Component Analysis : Theory and Applications, Kluwer Academic Pub., Boston, 1998
- [6] J. Karhunen, "Neural Approaches to Independent Component Analysis and Source Separation", "4th European Symp., Artificial Neural Network, ESANN96, Burges, Belgium, pp. 249-266, Apr. 1996
- [7] K. Atkinson, Elementary Numerical Analysis, John Wiley & Sons, Inc., New York, 1993