

# Kurtosis를 이용한 독립성분의 군집화에 관한 연구

조용현, 김아람\*  
대구가톨릭대학교 공과대학 컴퓨터정보통신공학부  
e-mail:yhcho@cu.ac.kr

## A Study on Clustering of Independent Components by Using Kurtosis

Yong-Hyun Cho and A-Ram Kim\*  
School of Computer and Information Comm. Eng.,  
Catholic Univ. of Daegu

### 요 약

본 연구에서는 뉴우턴법에 기초한 고정점 알고리즘의 신경망 기반 독립성분분석에 kurtosis를 추가한 독립성분의 군집화를 제안하였다. 여기서 뉴우턴법의 고정점 알고리즘은 엔트로피에 기초한 목적함수의 근을 구하는 근사화 방법으로 빠른 성분분석을 위함이고, kurtosis는 독립성분의 추출순서를 고려하지 않는 속성을 개선하기 위함이다. 제안된 기법을 256x256 픽셀의 8개 혼합영상의 분리에 적용한 결과, 제안된 방법은 기존의 독립성분분석에서 분석순서를 고려하지 않는 제약을 효과적으로 해결할 수 있음을 확인하였다.

### 1. 서론

데이터의 차원을 감소시키는 선형변환 기법으로 주요 성분분석(principal component analysis : PCA)<sup>[1-3]</sup>이 널리 이용되고 있다. 이는 공분산(covariance)을 이용하는 2차원 방법으로 가우스(Gaussian) 분포의 입력 데이터에 대해서는 적당하나 비가우스 분포의 데이터 표현은 부적당하다.

최근 이러한 PCA의 제약을 해결하기 위해 독립성분분석(independent component analysis : ICA)이 제안되었으며, 이는 데이터의 통계적 성질을 고려한 선형변환 기법으로 데이터의 분포정보를 이용하는 다차원 방법이다<sup>[4-6]</sup>. 하지만 알려진 ICA 알고리즘들에서는 데이터의 통계적 독립성을 판정하기 위해 확률밀도함수와 역행렬을 구한다든지 고차원의 cumulant나 moment를 구하는 등 복잡한 계산이 요구된다. 특히 복잡한 계산에 따른 많은 비현실적인 계산시간이 요구되는 수치적 기법의 대안으로 간단한 신경망을 이용하여 해결하려는 연구가 활발하게 진행되고 있다<sup>[5,6]</sup>.

신경망을 이용한 ICA 방법으로는 고정점(fixed point : FP) 알고리즘, recursive least square 알고리즘, 확률적 기법을 이용하는 Bell & Sejnowski(BS)알고리즘,

자연기울기(natural gradient)를 이용한 natural gradient 알고리즘 등이 제안되었다<sup>[5,6]</sup>. 일반적으로 고정점 알고리즘은 빠른 수렴속성을 가지지만 분리성능이 떨어지고, 기법을 이용하는 방법은 학습 파라미터의 설정에 따라서는 분리속도가 늦은 제약이 있다. 특히 이들 ICA 방법들에서는 신호의 독립성과 비가우스성의 조건들은 선행되어야 한다. 또한 출력되는 신호는 ICA에 신호의 입력 순서와 부호를 고려하지 않는 고유의 속성이 있다.

본 연구에서는 뉴우턴법에 기초한 고정점 알고리즘의 신경망 기반 독립성분분석에 kurtosis를 추가한 새로운 ICA를 제안한다. 여기서 뉴우턴법에 기초한 독립성분분석은 엔트로피 최적화로부터 유도된 기법으로, 빠른 분석을 위함이고, kurtosis는 그 결과의 군집화를 위함이다. 제안된 기법을 256x256픽셀(pixel)의 8개 영상으로부터 임의의 혼합행렬에 따라 발생하는 영상들을 각각 대상으로 실험하여 군집화의 성능을 확인하였다.

### 2. Kurtosis를 이용한 독립성분분석

ICA는  $m$ 개의 입력신호  $s$ 로부터 선형적으로 혼합된  $n$ 개의 신호  $x$ 가 알려져 있을 때, 혼합된 신호로부터 역으로  $m$ 개의 독립인 입력신호를 찾는 기법이다. 하지만 입

력신호들을 혼합하는데 이용된 혼합행렬  $A$ 는 알려져 있지 않으며, 혼합과정에서 잡음  $n$ 이 추가 될 수도 있다. 이때 혼합신호와 입력신호와의 관계는

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

로 정의된다. 여기서  $\mathbf{n}$ 은 보통 입력신호와 구별되지 않기 때문에 생략할 수도 있으며,  $\mathbf{A}=[\mathbf{a}(1), \mathbf{a}(2), \dots, \mathbf{a}(m)]$ 으로  $\mathbf{a}(i)$ 는 ICA의 basis vector이다. 결국 ICA는 알려진 혼합신호로부터 혼합행렬의 역행렬  $\mathbf{A}^{-1}$ 을 찾는 기법이다. 혼합행렬  $\mathbf{A}$ 와 역혼합행렬  $\mathbf{W}$ 에 대하여 상세히 살펴보면 다음 그림 1과 같은 구성도로 나타낼 수 있다. 여기서  $\mathbf{x} = \mathbf{A}\mathbf{s}$ 이고,  $\mathbf{y} = \mathbf{W}\mathbf{x}$ 이다. 이때 ICA는 출력신호  $\mathbf{y}$ 의 각 요소들이 통계적으로 독립성을 유지해야 하며 직교성의 제약조건은 포함하지 않는다. 그림에서 보면 ICA는 혼합행렬과 일치하는 역혼합행렬을 찾는 과정에서 출력신호가 독립성을 가지도록 하는 기법이다. 따라서 ICA는 알려진 혼합신호  $\mathbf{x}$ 로부터 출력신호  $\mathbf{y}$ 를 찾는 기법으로 궁극적으로는 역혼합행렬  $\mathbf{W}$ 을 찾아서 원신호  $\mathbf{s}$ 의 근사값을 알아내는 것이다. 이때 영평균과 whitening된 혼합신호를 이용하면, 신호의 1차와 2차의 통계성에 따른 상관성이 이미 고려되어 좀 더 효과적으로 ICA 기법을 수행할 수 있다.

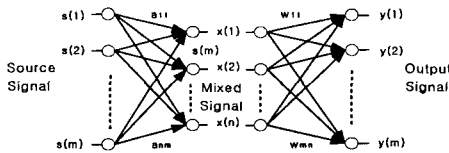


그림 1. 혼합행렬과 역혼합행렬의 상세 설명도

### 2.1 신호의 영평균과 Whitening

신호의 영평균은 데이터의 1차적 통계성을 고려한 정규화로 학습알고리즘을 간략하게 하는 기법이다<sup>[4-6]</sup>. 이는 신호 데이터 벡터  $\mathbf{x}$ 에서 평균값  $\mathbf{x}^*$ 를 뺀 차를 구함으로써 간단하게 구할 수 있다. 즉, 영평균된 새로운 데이터 벡터  $\mathbf{x}$ 는

$$\mathbf{x} = \mathbf{x} - \mathbf{x}^* \quad (2)$$

이다.

한편, 데이터의 2차적 통계성을 고려한 상호간의 의존성을 줄이기 위한 기법으로 whitening이 이용되고 있다<sup>[3]</sup>. whitening은 데이터 벡터  $\mathbf{x}$ 의 공분산행렬이

단위행렬값을 갖도록 함으로써 구할 수 있다. 즉,  $E(\mathbf{x}\mathbf{x}^T)=\mathbf{I}$ 가 되도록 한다. whitening된 벡터의 성분들은 상호간의 상관성이 줄어들어 비상관성이 더 강한 독립조건을 위해 요구되는 전처리 기법이다. 이러한 whitening의 기법으로 PCA가 널리 사용된다. 일반적인 PCA whitening 행렬  $\mathbf{V}$ 는

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

이다. 여기서  $\mathbf{D}$ 는 고유값  $\lambda$ 로 구성되는 대각행렬이며,  $\mathbf{E}$ 는 고유벡터  $\mathbf{c}$ 로 구성되는 직교행렬이다. PCA whitening이 가지는 장점은 이미 잘 알려진 표준 소프트웨어를 이용할 수 있으며, 또 다른 장점으로는 원신호, 즉 독립성분의 개수를 추정하는 편리한 수단이 되는 것이다.

### 2.2 고정점 알고리즘의 ICA와 Kurtosis

독립성분분석을 위한 고정점 알고리즘은 신경망이 가지는 병렬성과 분산성, 그리고 더 작은 메모리 요구 등의 제약을 해결하기 위해 제안된 알고리즘이다<sup>[5]</sup>. 고정점 알고리즘은 엔트로피 최적화 방법으로부터 유도된 기법으로 그 수렴속도가 적어도 2차 함수정도로 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.

고정점 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 목적함수  $J_G(\mathbf{W}) = [E(G(\mathbf{W}^T\mathbf{x})) - E(G(\nu))]^2$ 의 최소화는 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

$$E(\mathbf{x}g(\mathbf{W}^T\mathbf{x})) - \beta \mathbf{W} = 0 \quad (4)$$

로 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 (4)를 풀기 위해 빠른 수렴속도를 가지는 뉴턴의 기법을 도입한다. 이를 위해 식의 좌변항을  $F(\mathbf{W})$ 로 나타내고, 그것의 Jacobian matrix  $\mathbf{J}F(\mathbf{W})$ 를 구하면 다음과 같다. 즉,

$$JF(W) = E\{xx^T g'(W^T x)\} - \beta I \quad (5)$$

이다. 한편 데이터 벡터  $x$ 가 이미 whitening되어 있어 뉴우턴의 기법에 따른 해를 구하는데 이용되는  $JF(W)$ 의 역행렬을 간단하게 계산하기 위해 식 (5)의 우변 첫째 항을 다음과 같이 근사화할 수 있다. 즉,

$$E\{xx^T g'(W^T x)\} = E\{xx^T\}E\{g'(W^T x)\} = E\{g'(W^T x)\}I \quad (6)$$

이다. 식 (6)을 식 (5)에 대입하면 Jacobian matrix는 대각선 행렬이 되어 쉽게 역행렬을 계산할 수 있다. 따라서  $\beta$ 를  $W_0$  대신에  $W$ 의 현재값으로 나타내면 근사화된 반복기법의 역혼합행렬  $W$ 를 구하는 뉴우턴법은

$$W^* = W - [E\{xg(W^T x)\} - \beta W] / [E\{g'(W^T x)\} - \beta], \\ W^* = W^* / \|W^*\| \quad (7)$$

와 같이 정의된다. 여기서  $W^*$ 는  $W$ 의 새롭게 경신된 값이고,  $\beta = E\{W^T xg(W^T x)\}$ 이다. 결국 식 (7)은 뉴우턴법에 기초한 ICA를 위한 FP 알고리즘이다. 또한 식 (7)의 뒷식 양쪽에  $\beta - E\{g'(W^T x)\}$ 를 곱해 구해지는 더욱 간단한 뉴우턴법의 FP 알고리즘은

$$W^* = E\{xg(W^T x)\} - E\{g'(W^T x)\}W, W^* = W^* / \|W^*\| \quad (8)$$

이 된다.

한편 기존의 고정점 알고리즘은 독립성분을 분석할 때 순서를 고려하지 않는 제약을 가지고 있다. 이러한 제약을 해결하기 위해 본 논문에서는 독립성분의 kurtosis를 이용하였다. 이는 유사한 성분은 유사한 kurtosis를 가진다는 속성을 이용한 것이다. 이렇게 하면 뉴우턴법이 가지는 빠른 수렴속성을 그대로 살리면서도 은닉신호 분리시 유사한 속성을 가지는 성분끼리 군집화하여 좀 더 효율적으로 독립성분분석을 얻을 수 있기 때문이다. 일반적으로 이용되는 신호  $y$ 의 kurtosis 계산식은

$$kurt(y) = E\{y^4\} - 3(E\{y\}^2)^2 \quad (9)$$

와 같다.

고정점 알고리즘과 kurtosis를 이용한 제안된 ICA의 역혼합행렬  $W$ 의 경신과정을 요약하면 다음과 같다.

단계 1 : 초기 역혼합행렬  $W(0)$ 를 임의의 값 rand()로 설정한다.

단계 2 :  $W = W(0) / \|W(0)\|$ 의 값을 계산한다.

단계 3 :  $W_{old} = 0$ 로 설정한다.

단계 4 :  $\|W - W_{old}\| > \epsilon$ 이고  $\|W + W_{old}\| > \epsilon$ 이면 다음 식에 따라  $W$ 를 계산하고, 그렇지 않으면 과정을 종료한다.

(1)  $W_{old} = W$ 를 계산한다.

(2)  $W^* = E\{xg(W^T x)\} - [E\{g'(W^T x)\}W + \alpha(W - W^*)]$ 를 계산한다.

(3)  $W^* = W^* / \|W^*\|$ 를 계산한다.

단계 5 :  $y = W^* x$ 를 계산한다.

단계 6 : 식 (9)를 이용하여 kurtosis를 계산하고 순서화한다.

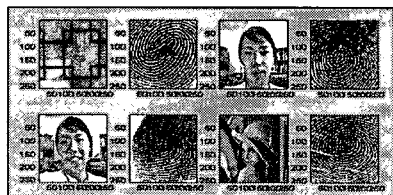
단계 7 : 순서화된 kurtosis에 따라 분석성분을 정렬한다.

위의 과정에서 벡터  $x$ 는 whitening된 신호벡터이며,  $\epsilon$ 은 1보다 작은 임의의 값이다. 본 실험에서는  $\epsilon = 0.0001$ 과 경신 식에서  $g(\cdot)$ 는  $(\cdot)^3$ 함수를 적용하였다.

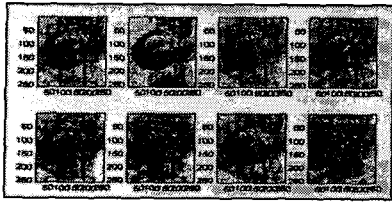
### 3. 시뮬레이션 및 결과분석

제안된 학습알고리즘의 신경망 기반 ICA의 타당성과 성능을 평가하기 위해  $256 \times 256$  픽셀을 가지는 8개의 영상을 대상으로 하였다. 분석성능을 평가하기 위해 각 신호의 절대합오차 ASE를 정의하여 원 신호와 복원된 신호의 정합성을 비교하였다. 이때 절대합오차  $ASE = 100 - \frac{1}{256 \times 256} \sum_{i=1}^{256} \sum_{j=1}^{256} |x_{ij} - y_{ij}|$ 이며, 여기서  $x_{ij}$ 와  $y_{ij}$ 는 각각 원 신호와 복원된 신호의  $ij$ 번째 픽셀값이다. 시뮬레이션은 팬티엄IV-2.2G 컴퓨터 상에서 Matlab 5.1로 실행하였다.

그림 2는 원 영상과 임의의  $8 \times 8$  혼합행렬로부터 생성된 혼합영상 8개를 각각 나타낸 것이다. 특히 그림 2(b)의 혼합영상 각각에서는 원 영상들을 발견할 수 없다. 실제에서 혼합영상들만 알려져 있으며 원 영상들은 미지의 영상들이다. 시뮬레이션에서는 혼합된 영상들을 전처리한 후 입력으로 이용하였다.



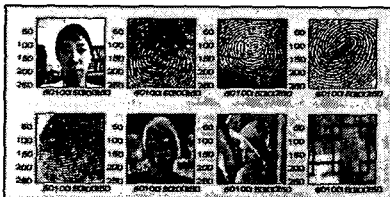
(a) 원 영상



(b) 혼합 영상

그림2. 256×256 픽셀의 8개 영상

그림 3은 그림 2(b)의 혼합신호들을 대상으로 고정점 알고리즘과 제안된 알고리즘의 ICA 각각을 수행한 후 분리된 영상들이다. 그림에서 각각의 알고리즘에서 원 영상과 복원된 영상들 사이의 순서와 부호가 일치되지 않음을 알 수 있다. 이는 수행과정에서 신호의 순차를 고려하지 않는 속성과 부호를 고려치 않는 ICA의 속성을 보여 주는 것이다. 한편 그림 3(a)의 고정점 알고리즘의 경우는 추출된 결과에서 유사한 영상의 군집화가 이루어지지 않음을 알 수 있다. 하지만 그림 3(b)의 제안된 알고리즘에서는 영상은 영상대로 지문은 지문대로 군집화가 이루어짐을 확인할 수 있다. 이는 군집화를 ICA의 결과에 반영한 것으로 분류문제나 식별 문제에 ICA가 적용될 수 있는 가능성을 제시한다. 특히 고정점 알고리즘의 경우 알고리즘의 수행마다 분리순서는 달라지나 제안된 알고리즘에서는 일정한 순서가 유지됨을 알 수 있다.



(a) 고정점 알고리즘



(b) 제안된 알고리즘

그림3. 256×256 픽셀의 8개 복원 영상

한편 본 연구에서는 8개의 혼합영상 분리에 제안된 알고리즘의 ICA를 20회 시행하여 8개의 영상을 분리하는데 소요되는 학습반복수, 분리시간, 그리고 픽셀

단위의 정합도 각각의 평균을 고찰하였다. 학습반복수는 총 59.4회로 평균적으로 1개의 영상을 추출하는데 약 7.4회 소요되었다. 또한 분리시간은 약 27.9초로 1개의 영상분리에는 약 3.5초가 소요되었다. 한편 정합도에서는 평균적으로 약 97.9정도의 높은 정합성이 있음을 알 수 있다.

#### 4. 결론

본 연구에서는 뉴우턴법에 기초한 고정점 알고리즘의 신경망 기반 독립성분분석에 kurtosis를 추가한 독립성분의 군집화를 제안하였다. 여기서 뉴우턴법의 고정점 알고리즘은 엔트로피에 기초한 목적함수의 근을 구하는 근사화 방법으로 빠른 성분분석을 위함이고, kurtosis는 독립성분의 추출순서를 고려하지 않는 속성을 개선하기 위함이다.

제안된 알고리즘의 신경망 기반 독립성분분석을 256×256 픽셀의 8개 혼합영상 분리에 적용한 결과, 기존의 고정점 알고리즘에서 추출순서가 고려되지 않는 속성을 개선한 군집화 특성을 얻을 수 있었다.

향후 제안된 알고리즘의 독립성분분석 기법을 다양한 분야에 응용하기 위한 연구가 뒤따라야 할 것이다.

#### [참고문헌]

- [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