

첨도에 의한 분석성분의 군집성을 고려한 독립성분분석

조 용 현*

요 약

본 논문에서는 첨도를 추가한 뉴우턴법과 할선법에 기초한 고정점 알고리즘의 독립성분분석을 각각 제안하였다. 여기서 첨도의 추가는 유사한 속성을 가지는 성분의 군집화된 분석순서를 얻기 위함이고, 고정점 알고리즘은 학습파라미터와 무관한 빠른 성분분석과 우수한 분석성능을 얻기 위함이다. 제안된 두 가지 독립성분분석 각각을 500개 샘플을 가지는 6개의 혼합신호와 512x512 픽셀을 가지는 10개의 혼합영상 분리에서 각각 적용한 결과, 제안된 두 가지 기법은 항상 일정한 분석순서를 유지하여 첨도가 추가되지 않은 기존의 기법들에서 알고리즘의 수행 때마다 랜덤하게 변하는 분석순서의 제약을 해결할 수 있었다. 특히 군집화의 속성을 가진 제안된 독립성분분석들은 신호나 영상의 분류나 식별에도 적용할 수 있다. 한편 할선법의 제안된 기법이 뉴우턴법의 제안된 기법보다 빠르면서도 우수한 분리성능이 있음을 확인하였다.

Independent Component Analysis for Clustering Analysis Components by Using Kurtosis

Yong-Hyun Cho[†]

ABSTRACT

This paper proposes an independent component analyses(ICAs) of the fixed-point (FP) algorithm based on Newton and secant method by adding the kurtosis, respectively. The kurtosis is applied to cluster the analyzed components, and the FP algorithm is applied to get the fast analysis and superior performance irrelevant to learning parameters. The proposed ICAs have been applied to the problems for separating the 6-mixed signals of 500 samples and 10-mixed images of 512x512 pixels, respectively. The experimental results show that the proposed ICAs have always a fixed analysis sequence. The results can be solved the limit of conventional ICA without a kurtosis which has a variable sequence depending on the running of algorithm. Especially, the proposed ICA can be used for classifying and identifying the signals or the images. The results also show that the secant method has better the separation speed and performance than Newton method. And, the secant method gives relatively larger improvement degree as the problem size increases.

키워드 : 독립성분분석(Independent Component Analysis), 고정점 알고리즘(Fixed-Point Algorithm), 첨도(Kurtosis), 군집화(Clustering), 뉴우턴법(Newton Method), 할선법(Secant Method)

1. 서 론

통계학이나 신호처리 등에서 적절한 데이터의 표현을 찾기 위해 일반적으로 선형변환을 이용하며, 그 중에서도 공분산(covariance)을 이용하는 2차원 방법인 주요성분분석(principal component analysis : PCA)[1-7]이 가장 널리 이용되고 있다. PCA는 데이터를 공분산에 의해 결정되는 orthonormal basis로 표현하며, 가우스(Gaussian) 분포의 입력 데이터에 대해서는 적절하다[1]. 일반적으로 통신이나 신호 및 영상 처리 등에서의 데이터에는 비가우스 분포도 있으며, 이러한 경우에도 효과적으로 적용될 수 있는 선형변환으로 PCA를 확장한 독립성분분석(independent component analysis : ICA)이 제안되었다[2-6]. 이는 데이터의 통계적 성질을 고려한

선형변환 기법으로 데이터의 분포정보를 이용하는 다차원 방법이다. 하지만 알려진 수치적 기법을 이용한 ICA에서는 데이터의 통계적 독립성을 판정하기 위해 확률밀도함수와 역행렬을 구한다든지 고차원의 모멘트를 구하는 등 복잡한 계산이 요구된다. 특히 ICA에서 신호모델을 추정하는 것은 혼합된 출력신호들 사이에 포함된 정보만을 이용하여 혼합행렬을 추정하는 것이며, 이를 위해 출력신호 상호간의 결합밀도함수를 이용함으로써 계산이 매우 복잡하다. 이러한 복잡한 계산에 따른 많은 비현실적인 계산시간이 요구되는 수치적 기법의 대안으로 간단한 신경망을 이용한 연구가 활발하게 진행되고 있다[4-9].

신경망을 이용한 ICA방법으로는 엔트로피를 최소화하는 것에 기초를 둔 고정점(fixed point : FP) 알고리즘, 비선형 PCA에 기초한 반복최소자승 알고리즘, 결합엔트로피를 최대화하기 위하여 확률적 기울기를 이용하는 Bell & Sejnowski

[†] 정 회 원 : 대구가톨릭대학교 컴퓨터정보통신공학부 교수
 논문접수 : 2004년 3월 20일, 심사완료 : 2004년 4월 22일

(BS) 알고리즘, BS 알고리즘의 확률적 기울기 대신 자연기울기를 이용한 자연기울기 알고리즘 등이 제안되었다[4-7]. 여기에서 결합엔트로피를 최대화하는 알고리즘은 원 신호를 더 잘 근사화 하지만 기울기의 변화를 위한 학습율과 같은 학습파라미터의 설정이 경험적이다. 특히 학습과정에서 학습파라미터의 설정에 따라서는 그 수렴이 느려지거나 알고리즘이 수렴되지 않는 경우도 발생된다. 따라서 이런 제약들을 해결하면서도 신경망이 가지는 병렬성, 분산성, 그리고 더 작은 메모리 요구 등의 속성을 가지는 FP 알고리즘이 제안되었다[7, 8]. FP 알고리즘은 엔트로피 최적화 방법으로부터 유도된 기법으로 뉴턴(Newton)법[10]에 기초를 두고 있으며, 그 수렴속도는 적어도 2차 함수정도로 지금까지 알려진 기법 중 가장 빠른 분석속도를 가진다. 그러나 뉴턴법에 기초한 고정점 알고리즘은 역혼합행렬의 경신을 위하여 비선형 방정식의 해를 구할 때 함수의 성질에 따라 그 수렴시간이 증가하거나 근에 수렴하지 못하고 발산하는 경우가 있다. 또한 기울기인 도함수 값이 0에 수렴할 경우 비현실적인 수렴현상도 일어날 수 있다. 이러한 뉴턴법이 가지는 제약을 해결하기 위하여 할선법[10]에 기초한 고정점 알고리즘이 제안되었다[9]. 이는 뉴턴법에서의 도함수를 함수 값을 이용하여 계산함으로써 알고리즘의 수행시간을 더욱 더 줄일 수 있도록 한 것이다. 하지만 이들 ICA에서는 원 신호 사이에 독립성과 비가우성의 조건이 선행되어야 하며, ICA가 적용된 후 출력되는 분석성분의 순서와 부호가 알고리즘의 수행 때마다 달라지는 제약이 있다. 따라서 ICA의 수행과 무관하게 유사한 속성을 가지는 성분들이 항상 일정하게 군집화된 분석순서를 유지하도록 한다면, 기존의 단순한 고정점 알고리즘 ICA보다 우수한 속성을 가진 ICA가 될 것이다.

본 연구에서는 뉴턴법과 할선법에 기초한 FP 알고리즘 각각에 침도를 추가한 ICA들을 제안한다. 여기서 FP 알고리즘은 학습파라미터와 무관하게 빠르고 정확한 성분의 분석을 위함이고, 침도의 추가는 유사한 속성을 가지는 성분의 군집화를 얻기 위함이다. 제안된 ICA들을 임의로 생성된 500개의 샘플을 가진 6개의 혼합신호와 512×512 픽셀을 가지는 10개의 혼합영상의 분리를 대상으로 각각 실험하여 타당성을 확인하였으며, 기존 FP 알고리즘의 ICA의 결과들과도 비교·고찰하였다.

2. 뉴턴법 및 할선법의 고정점 알고리즘에 침도를 추가한 독립성분석

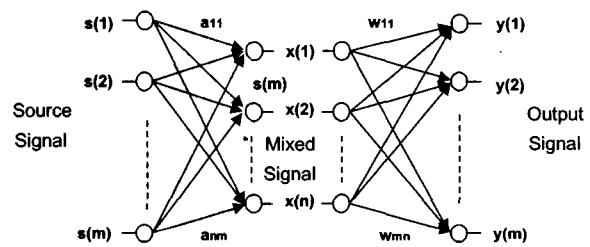
PCA가 가우스 분포를 가지는 신호들의 2차원 선형변환 기법이라면, ICA는 비가우스 분포인 군집성이나 독립성을 가지는 변수의 고차통계에 기초한 다차원 변환기법으로 PCA를 확장시킨 것이다. 따라서 ICA는 PCA보다 분포 데이터의 속성을 더욱 더 잘 나타내며, 은닉신호의 분리(blind source

separation : BSS)나 특징추출에 널리 이용되고 있다[4-6]. 여기에서는 BSS를 기준으로 ICA를 소개한다.

ICA는 m개의 입력신호 s로부터 선형적으로 혼합된 n개의 혼합신호 x가 알려져 있을 때, 혼합된 신호로부터 역으로 m개의 독립인 입력신호를 찾는 기법이다. 하지만 입력신호들을 혼합하는데 이용된 혼합행렬 A는 알려져 있지 않으며, 혼합과정에서 잡음 N이 추가 될 수도 있다. 이때 혼합신호와 입력신호와의 관계는

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

로 정의된다. 여기서 N은 보통 입력신호와 구별되지 않기 때문에 생략할 수도 있으며, $A=[a(1), a(2), \dots, a(m)]$ 으로 a(i)는 ICA의 기저벡터이다. 결국 ICA는 알려진 혼합신호로부터 혼합행렬의 역행렬 $A^{-1}(=W)$ 을 찾는 기법이다. 이때 혼합행렬 A와 역혼합행렬 W의 관계를 보면 (그림 1)과 같이 나타낼 수 있다. 여기서 ICA는 후단의 점선부분으로 $x = As$ 이고, $y = Wx$ 이다. 따라서 ICA는 알려진 혼합신호 x로부터 출력신호 y를 찾는 것이며, 궁극적으로는 역혼합행렬 W를 찾아서 원 신호 s의 근사값을 알아내는 것이다.



(그림 1) ICA의 구성도

일반적으로 신호의 통계성에 따른 상관성을 고려하여 좀 더 효율적인 ICA를 수행하기 위하여 영평균(zero-mean)과 whitening의 전처리 과정을 이용한다[4-6]. 신호의 영평균은 데이터의 1차적 통계성을 고려한 정규화로 학습알고리즘을 간략하게 하는 기법이며, 신호벡터 x에서 그 평균값 x^* 를 뺀 $x = x - x^*$ 로 구할 수 있다. 또한 whitening은 데이터의 2차적 통계성을 고려하여 상호간의 의존성을 줄이기 위한 기법이며, 신호 x의 공분산행렬 $E(xx^T)$ 을 단위행렬 I가 되도록 함으로써 구할 수 있다. 이를 위해 PCA가 널리 이용되며, PCA에 의한 whitening 행렬 $V = D^{(-1/2)}E^T$ 이다. 여기서 D는 고유값으로 구성되는 대각행렬이며, E는 고유벡터로 구성되는 직교행렬이다.

FP 알고리즘의 ICA는 엔트로피 최적화로부터 유도된 기법으로 지금까지 알려진 기법 중 가장 빠른 분석속도를 가진다[7]. 신호벡터 x가 whitening되어 있다고 가정할 때, 뉴턴법의 FP 알고리즘에 의한 근사화된 반복기법의 역혼합행렬 W 경신식은

$$\begin{aligned} \mathbf{W}^* &= \mathbf{W} - [\mathbf{E}\{\mathbf{xg}(\mathbf{W}^T\mathbf{x})\} - \beta \mathbf{W}] / [\mathbf{E}\{g'(\mathbf{W}^T\mathbf{x})\} - \beta] \\ \mathbf{W} &= \mathbf{W}^* / \|\mathbf{W}^*\| \end{aligned} \quad (2)$$

와 같다. 여기서 $\beta = \mathbf{E}\{\mathbf{W}^T\mathbf{xg}(\mathbf{W}^T\mathbf{x})\}$ 이다. 또한 식 (2)의 위 식 양쪽에 $\beta - \mathbf{E}\{g'(\mathbf{W}^T\mathbf{x})\}$ 를 곱해 구해지는 더욱 간단한 FP 알고리즘은

$$\begin{aligned} \mathbf{W}^* &= \mathbf{E}\{\mathbf{xg}(\mathbf{W}^T\mathbf{x})\} - \mathbf{E}\{g'(\mathbf{W}^T\mathbf{x})\}\mathbf{W} \\ \mathbf{W} &= \mathbf{W}^* / \|\mathbf{W}^*\| \end{aligned} \quad (3)$$

과 같다. 여기서 $g(\cdot)'$ 은 비선형 함수 $g(\cdot)$ 의 1차 미분이다. 한편 수치해석에서 할선법은 현재의 함수 값과 바로 이전의 함수 값을 이용하여 방정식의 근을 구하는 방법으로 뉴우턴법에서 별도의 1차 미분인 도함수 값의 계산을 요구하지 않는다[10, 11]. 따라서 뉴우턴법에 기초한 FP 알고리즘에서 목적함수의 근을 구하는데 할선법에 기초한 FP 알고리즘이 이용될 수 있다[9]. 이때 신호벡터 \mathbf{x} 가 whitening되어 있다고 가정할 때, 할선법의 FP 알고리즘에 의한 근사화된 반복기법의 역혼합행렬 \mathbf{W} 의 경신식은

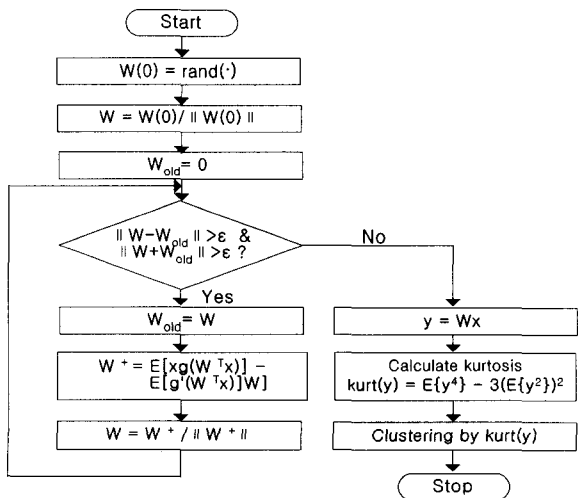
$$\begin{aligned} f(\mathbf{W}_{old}) &= [\mathbf{E}\{\mathbf{xg}(\mathbf{W}_{old}^T\mathbf{x})\} - \beta \mathbf{W}_{old}] \\ f(\mathbf{W}) &= [\mathbf{E}\{\mathbf{xg}(\mathbf{W}^T\mathbf{x})\} - \beta \mathbf{W}] \\ \mathbf{W}^* &= \mathbf{W} - f(\mathbf{W})\{[\mathbf{W} - \mathbf{W}_{old}] / \{f(\mathbf{W}) - f(\mathbf{W}_{old})\}\} \\ \mathbf{W} &= \mathbf{W}^* / \|\mathbf{W}^*\| \end{aligned} \quad (2)$$

와 같다. 여기서 \mathbf{W}_{old} 는 바로 전에 계산된 \mathbf{W} 의 값이다. 뉴우턴법이나 할선법의 FP 알고리즘에서 $g(\cdot)$ 함수로 $\tanh(\cdot)$ 함수가 일반적으로 이용된다[4-6]. 그러나 뉴우턴법이나 할선법에 기초한 FP 알고리즘의 ICA들에서도 성분을 분석할 때 순서를 고려하지 않는 제약은 여전히 남아있다. 한편 첨도는 비가우스성의 고전적 척도로 널리 이용되고 있다[4-6]. 이는 4차 cumulant로 다음의 식 (3)과 같이 계산된다.

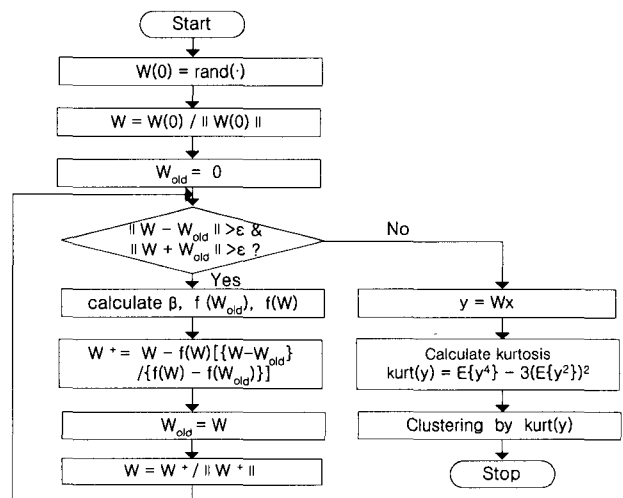
$$kurt(y) = E\{y^4\} - 3(E\{y^2\})^2 \quad (3)$$

여기서 y 는 랜덤변수이며, 대부분 비가우스 랜덤변수의 첨도는 영이 아닌 양이나 음의 값을 가진다. 만약 랜덤변수들의 첨도가 음의 값을 가지면 subgaussian이고, 양의 값을 가지면 supergaussian이다. subgaussian 분포의 랜덤변수는 첨예하지 않고 일정한 확률밀도함수를 가지며, supergaussian 분포의 랜덤변수는 넓고 첨예한 확률밀도함수를 가진다. 균일 분포함수의 랜덤변수는 양의 첨도, 라플라스 분포함수의 랜덤변수는 음의 첨도, 그리고 가우스 분포함수의 랜덤변수는 영의 첨도 값을 가진다. 일반적으로 첨도는 계산적으로나 이론적으로 간단하기 때문에 ICA에서 비가우스성의 척도로 널리 이용된다. 본 연구에서는 기존 ICA들에서 성분의 분석순서를 고려하지 않는 제약을 해결하기 위해 FP 알고리즘에 계산이 간단한 첨도를 추가한 ICA들을 제안한다. 이는 유사한 속성을 가지는 랜덤변수는 유사한 첨도값을 가지는 특성을 이용한 것이다. 이렇게 하면 FP 알고리즘이 가지는 빠른 수렴속성을 그대로 살리면서도 유사한 속성을 가지는 성분끼리 군집화하여 좀 더 효율적으로 독립성분을 분석할 수 있다. 또한 첨도를 이용하면 크게는 subgaussian과 supergaussian 및 가우스 분포 신호들의 군집화가 가능하고, 각 군집내에서의 순서화도 가능하다. 결국 좀 더 빠르면서도 정확하게 성분을 분석하기 위하여 FP 알고리즘을 이용하고, 랜덤한 분석순서를 가지는 ICA 고유의 제약을 해결하기 위하여 첨도를 이용하면 매우 효과적인 ICA가 이루어 질 수 있다.

이상에서의 뉴우턴법이나 할선법에 기초한 FP 알고리즘 각각에 첨도가 추가된 제안된 ICA들에 의한 역혼합행렬 \mathbf{W} 의 경신과정을 도시하면 다음과 (그림 2)와 같다. 여기서 ϵ 은 고정점 알고리즘의 종료조건으로 1보다 적은 임의의 값이다.



(a) 뉴우턴법의 FP 알고리즘



(b) 할선법의 FP 알고리즘

(그림 2) FP 알고리즘에 첨도가 추가된 제안된 ICA

3. 실험 및 결과분석

제안된 군집속성을 가지는 ICA들의 타당성과 성능을 평가하기 위하여 1차원 혼합신호와 2차원 혼합영상을 각각 분리하는 문제에 적용하였다. 학습의 종료조건 $\epsilon = 0.0001$ 로 하였다. 신호의 분리성능을 평가하기 위해 원 신호 s_i 와 분리된 신호 y_i 사이의 평균절대오차 $mae(\text{mean absolute error})$

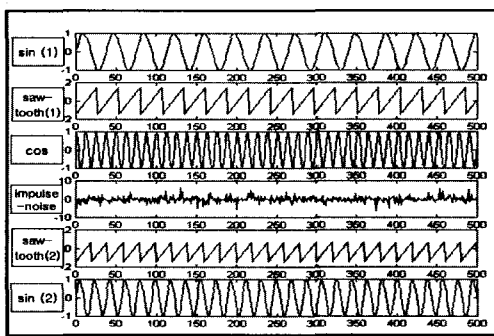
$$= \frac{1}{N^2} \sum_{i=1}^{N^2} |s_i - y_i|$$

를 정의하여 비교하였으며, 여기서 N^2 은 전체 픽셀수이다. 실험은 펜티엄 IV-2.2G 컴퓨터상에서 Matlab 5.1[11]로 실행하였고, 이때 $(\cdot)^3$ 의 비선형 함수를

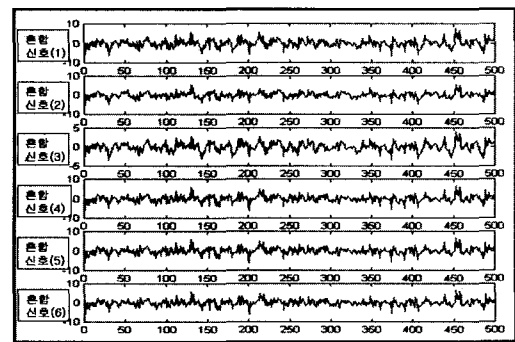
이용하였다.

3.1 1차원 신호

500개의 샘플을 가지는 1차원 배열의 6개 신호로부터 6×6 임의의 혼합행렬에 따라 생성된 6개의 혼합신호를 대상으로 실험하였다. (그림 3)은 실험에 이용된 6개의 원 신호와 생성된 6개의 혼합신호를 각각 나타낸 것이다. 특히 (그림 3)(b)의 혼합신호 각각에서는 원 신호들의 정보를 발견할 수 없으며, 실제에서 혼합신호들은 알려져 있으나 원 신호들은 미지의 신호들이다. 또한 실험에서는 영 평균과 whitening의 전처리된 혼합신호들을 이용하였다.



(a) 원 신호



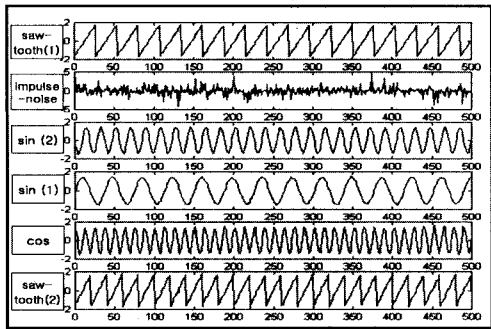
(b) 혼합신호

(그림 3) 500개의 샘플을 가진 6개 신호

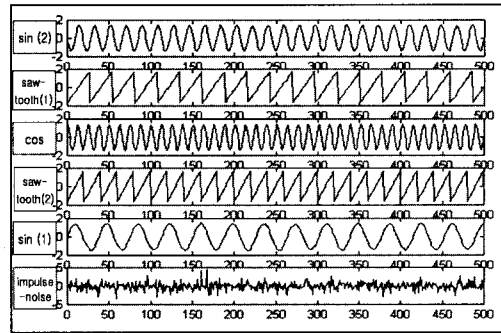
(그림 4)는 (그림 3)(b)의 혼합신호를 대상으로 뉴우턴법과 할선법에 기초한 FP 알고리즘에 첨도가 각각 추가된 제안된 ICA들에 의해 분리된 신호들을 나타낸 것이다. (그림 4)에서 보면, 두 가지 알고리즘 모두에서 원 신호와 분리된 신호들 사이의 분리순서는 서로 일치하지 않으며, 두 가지 알고리즘 사이에도 분리순서가 서로 일치하지 않는다. 또한 $\sin(2)$ 의 경우는 부호도 반대임을 알 수 있다. 이는 독립성분을 분석할 때 부호나 순서를 고려하지 않는 ICA의 고유한 속성이 그대로 반영되었기 때문이다. 한편 (그림 4)(a)와 (b)의 뉴우턴법과 할선법에 기초한 기존 FP 알고리즘의 ICA들 각각에 의해 분리된 신호들은 유사한 성분끼리 서로 군집화가 이루어지지 않음을 알 수 있다. 하지만 (그림 4)(c)와 (그림 4)(d)의 제안된 ICA들에서는 첨도에 따라 유사한 속성을 가지는 신호들이 서로 군집화되었음을 알 수 있다. 여기서는 먼저 saw-tooth 함수 (2)와 (1) 2개가 군집화되었고, 다음으로는 유사한 속성을 가지는 삼각함수인 cos 함수와 sin 함수 (1)과 (2) 2개, 그리고 다른 함수와 유사성이 거의 없는 impulse-noise순으로 군집화되어 분리됨을 알 수 있다. 결국 유사성이 있는 신호끼리의 군집화로 saw-tooth 함수는 saw-tooth 함수끼리 sin 함수와 cos 함수의 경우 한 무리로 군집화됨을 알 수 있다. 이는 제안된 ICA들은 신호나 영상의 분류나 식별에도 이용될 수 있음을 보여주는 것이다. 또한 (그림 4)(c)와 (그림 4)(d)를 보면 분리순서와 부

호에서 완전히 일치함을 알 수 있다. 이는 분리된 성분의 부호가 두 알고리즘에서 동일하기 때문이며, 순서는 첨도 값에 따라 결정되기 때문이다. 한편 기존의 ICA와 제안된 ICA를 각각 20회 실행하였을 때, 기존의 ICA들에서는 알고리즘의 수행 때마다 신호의 분석순서가 변하나 제안된 ICA들에서는 항상 일정한 분석순서가 유지됨을 확인하였다.

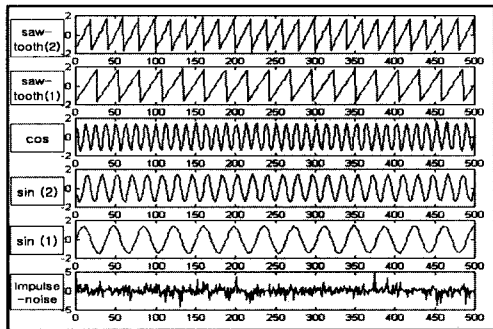
<표 1>은 혼합신호를 대상으로 기존 뉴우턴법과 할선법의 FP 알고리즘 ICA 각각과 첨도가 추가된 제안된 알고리즘의 ICA 각각을 20회 반복 시행한 후, 학습반복수 itr , CPU시간으로 정의된 분리시간 t , 그리고 평균절대오차 mae 의 평균을 나타낸 것이다. 표에서 6개의 신호를 분리하는데 소요된 학습반복수는 먼저 뉴우턴법의 기존 ICA와 제안된 ICA에서는 총 41.2회로 1개의 신호를 추출하는데 평균 약 6.87회 정도이며, 할선법의 기존 ICA와 제안된 ICA에서는 총 38.1회로 1개의 신호를 추출하는데 평균 약 6.35회 정도 수행되었다. 여기서 기존과 제안된 ICA들 각각에서 동일한 학습반복수를 가짐은 제안된 알고리즘의 ICA들도 기존 ICA들의 고정점 알고리즘을 그대로 이용하기 때문이다. 그러나 분리시간에서 뉴우턴법의 고정점 알고리즘 ICA는 총 1.16초로 평균 약 0.19초가 소요된 반면, 첨도가 추가된 제안된 ICA에서는 총 1.22초로 평균 약 0.2초 정도가 소요되었다. 또한 할선법의 고정점 알고리즘 ICA는 총 0.92초로 평균 약 0.15초가 소요된 반면, 첨도가 추가된 제안된 ICA에서는



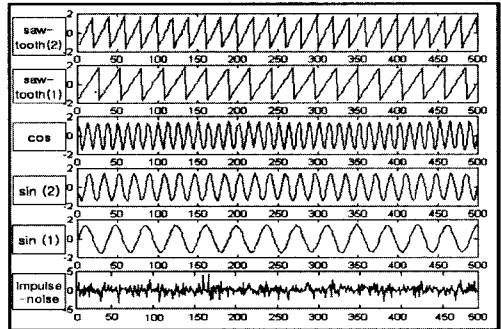
(a) 뉴우턴법의 고정점 알고리즘



(b) 할선법의 고정점 알고리즘



(c) 뉴우턴법의 제안된 알고리즘



(d) 할선법의 제안된 알고리즘

(그림 4) ICA에 의해 분리된 신호

<표 1> 6개의 혼합신호 분리에 따른 결과

원 신호	뉴우턴법의 고정점 알고리즘 ICA			뉴우턴법의 고정점 알고리즘과 점도의 ICA			할선법의 고정점 알고리즘 ICA			할선법의 고정점 알고리즘과 점도의 ICA		
	학습 반복수, n	분리 시간, t	평균절대 오차, mae	학습 반복수, n	분리 시간, t	평균절대 오차, mae	학습 반복수, n	분리 시간, t	평균절대 오차, mae	학습 반복수, n	분리 시간, t	평균절대 오차, mae
sin(1)	11.2	0.31	0.28	11.2	0.33	0.28	8.2	0.20	0.26	8.2	0.22	0.26
saw-tooth(1)	8.0	0.23	0.13	8.0	0.24	0.13	7.4	0.18	0.12	7.4	0.20	0.12
cos	8.1	0.23	0.28	8.1	0.24	0.28	8.6	0.21	0.26	8.6	0.23	0.26
impulse-noise	7.9	0.22	0.19	7.9	0.24	0.19	8.0	0.19	0.15	8.0	0.22	0.15
saw-tooth(2)	5.0	0.14	0.32	5.0	0.14	0.32	4.9	0.12	0.31	4.9	0.13	0.31
sin(2)	1.0	0.028	0.25	1.0	0.03	0.25	1.0	0.024	0.29	1.0	0.027	0.29
계	41.2	1.16	1.45	41.2	1.22	1.45	38.1	0.92	1.39	38.1	1.02	1.39
평균	6.87	0.19	0.24	6.87	0.20	0.24	6.35	0.15	0.23	6.35	0.17	0.23

총 1.02초로 평균 약 0.17초 정도가 소요되었다. 이는 제안된 ICA들에서는 기존 ICA들의 후처리 과정으로 점도를 계산하고 그에 따라 신호를 정렬하는데 각각 총 0.06초 정도와 총 약 0.1초 정도 더 소요되었기 때문이다. 하지만 원 신호와 분리된 신호의 평균절대오차에서는 기존의 ICA들과 제안된 ICA들 둘 다 각각 동일한 정확성을 가지며, 뉴우턴법에서는 총 1.45로 1개의 신호에 대해서는 평균 약 0.24 정도이고, 할선법에서는 총 1.39로 1개의 신호에 대해서는 평균 약 0.23 정도임을 알 수 있다. 따라서 할선법은 뉴우턴법에 비해 분리시간에서는 약 1.2배 정도 빠르며, 평균절대오차에서는 약 1.04배 정도 개선된 성능이 있음을 알 수 있다.

3.2 2차원 영상

512×512의 픽셀을 가지는 10개의 원 영상으로부터 10×10의 임의 혼합행렬에 의해 생성된 10개의 혼합영상들을 대상으로 실험하였다. (그림 5)는 10개의 원 영상과 생성된 10개의 혼합영상 각각을 나타낸 것이다. 여기에서도 혼합영상 각각에서 원 영상들의 정보들을 발견할 수 없으며, 혼합영상들은 알려져 있으나 원 영상들은 미지의 영상들이고, 전처리된 혼합영상들을 실험에 이용하였다.

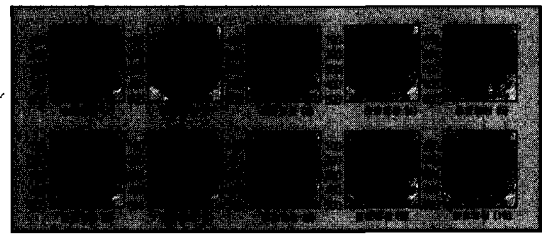
(그림 6) 역시 (그림 5)(b)의 혼합영상을 대상으로 할선법의 FP 알고리즘과 제안된 FP 알고리즘의 ICA를 각각 적용하여 분리한 10개의 영상들을 나타낸 것이다. 여기에서도 두 가지 알고리즘 모두와 상호사이에서 원 신호와 분리된 신

호들 사이의 분리순서는 서로 일치하지 않으며, (그림 6)(a)에서는 도장 (1)과 (2), 얼굴 (1), 지문 (2)가 (그림 6)(b)에서는 지문 (1), (2), (3), 도장 (1), 벽지 영상이 반대 부호임을 알 수 있다. 이 역시 부호나 순서를 고려하지 않는 ICA 고유의 속성 때문이다. 1차원 신호에서처럼 (그림 6)(a)와 (그림 6)(b)의 뉴우턴법과 할선법에 기초한 기존 FP 알고리즘의 ICA들 각각에 의해 분리된 신호들은 유사한 성분끼리 서로 군집화가 이루어지지 않음을 알 수 있다. 하지만 (그림 6)(c)와 (그림 6)(d)의 제안된 ICA들에서는 첨도에 따라 유사한 속성을 가지는 신호들이 서로 군집화되었음을 알 수 있다. 여기서는 먼저 벽지 영상, 도장 (1)과 (2) 2개, 레나 및

얼굴 (2)와 (1) 2개, 그리고 지문 (2), (4), (3), (1) 4개가 군집화되어 분리됨을 알 수 있다. 결국 유사성이 있는 신호끼리의 군집화로 도장은 도장끼리 얼굴은 얼굴끼리 지문은 지문끼리 한 무리로 군집화됨을 알 수 있다. 특히 유사한 속성을 가지는 벽지는 도장과 레나는 얼굴 군집화됨도 알 수 있다. 이는 제안된 ICA들이 신호나 영상의 분류나 식별에도 이용될 수 있음을 다시 한 번 더 보여 주는 것이다. 여기에서도 신호에서처럼 기존의 ICA들과 제안된 ICA들을 각각 20회 실행하였을 때, 기존의 ICA에서는 알고리즘의 수행 때마다 분석순서가 달라지나 제안된 ICA들에서는 항상 일정한 분석순서가 유지되었다.

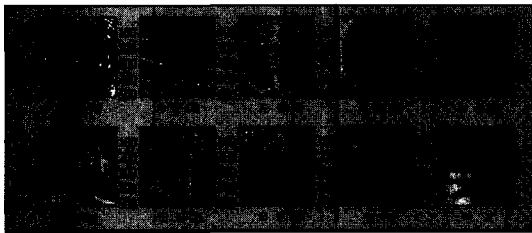


(a) 원 영상

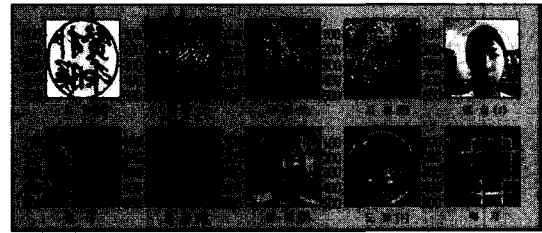


(b) 혼합영상

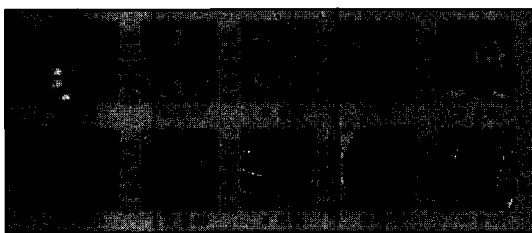
(그림 5) 512×512 픽셀의 10개 영상



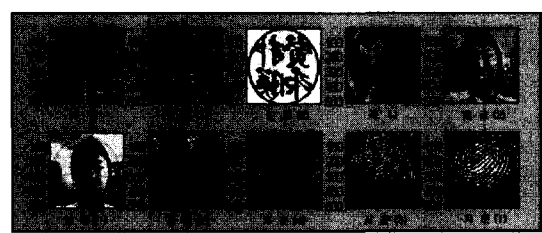
(a) 뉴우턴법에 기초한 고정점 알고리즘



(b) 할선법에 기초한 고정점 알고리즘



(c) 뉴우턴법의 제안된 알고리즘



(d) 할선법의 제안된 알고리즘

(그림 6) 10개의 분리된 영상

<표 2>도 혼합영상을 대상으로 기존 뉴우턴법과 할선법의 FP 알고리즘 ICA 각각과 첨도가 추가된 제안된 알고리즘의 ICA 각각을 20회 반복 시행한 후, 학습반복수 itr , CPU 시간으로 정의된 분리시간 t , 그리고 평균절대오차 mae 의 평균을 나타낸 것이다. 표에서도 10개의 신호를 분리하는데 소요된 학습반복수는 먼저 뉴우턴법의 기존 ICA와 제안된 ICA에서는 총 98.3회로 1개의 신호를 추출하는데 평균 약 9.83회 정도이며, 할선법의 기존 ICA와 제안된 ICA에서는 총 78.8회로 1개의 신호를 추출하는데 평균 약 7.88회 정도

수행되었다. 여기서도 기존과 제안된 ICA들 각각에서 동일한 학습반복수를 가짐은 제안된 ICA들에서도 기존 ICA들의 고정점 알고리즘을 그대로 이용하기 때문이다. 그러나 분리시간에서 뉴우턴법의 고정점 알고리즘 ICA는 총 128.77초로 평균 약 12.88초가 소요된 반면, 첨도가 추가된 제안된 ICA에서는 총 133.68초로 평균 약 13.37초 정도가 소요되었다. 또한 할선법의 고정점 알고리즘 ICA에서는 총 101.62초로 평균 약 10.16초가 소요된 반면, 첨도가 추가된 제안된 ICA에서는 총 105.33초로 평균 약 10.53초 정도가 소요되었

다. 이는 제안된 ICA들에서는 기존 ICA들의 후처리 과정으로 첨도를 계산하고 그에 따라 신호를 정렬하는데 각각 총 4.91초 정도와 총 약 3.71초 정도 더 소요되었기 때문이다. 하지만 원 신호와 분리된 신호의 평균절대오차에서는 기존의 ICA들과 제안된 ICA들 둘 다 각각 동일한 정확성을 가지며, 뉴우턴법에서는 총 11.52로 1개의 신호에 대해서는 평균 약 1.15 정도이고, 할선법에서는 총 10.96으로 1개의 신호에 대해서는 평균 약 1.1 정도임을 알 수 있다. 따라서 할선법은 뉴우턴법에 비해 분리시간에서는 약 1.27배 정도 빠르며, 평균절대오차에서는 약 1.05배 정도 개선된 성능이 있음을 알 수 있다. 또한 신호분리와 비교하면 영상분리는 보다 큰 규모의 문제로 분리시간과 평균절대오차에서 성능의

개선 정도가 더욱 더 큼을 확인할 수 있다.

이상의 혼합신호와 혼합영상의 분리실험을 통하여, 제안된 ICA들은 첨도가 가지는 계산의 용이성을 살리면서도 성분의 분석순서를 고려치 않는 기존 ICA들의 제약을 효과적으로 해결할 수 있다. 또한 제안된 ICA들은 유사한 첨도 값을 가지는 성분끼리의 군집화 특성을 가져 여러 가지 신호나 영상을 식별하거나 분류하는데 효과적인 수단으로도 적용될 수 있다. 한편 할선법에 기초한 기존과 제안된 ICA가 뉴우턴법에 기초한 기존과 제안된 ICA보다 분석시간과 분석율에서 좀 더 우수한 성능을 가지며, 문제의 규모가 클수록 그 성능의 개선정도도 증가함을 알 수 있었다.

〈표 2〉 10개의 혼합영상 분리에 따른 결과

원 영상	뉴우턴법의 고정점 알고리즘 ICA			뉴우턴법의 고정점 알고리즘과 첨도의 ICA			할선법의 고정점 알고리즘 ICA			할선법의 고정점 알고리즘과 첨도의 ICA		
	학습 반복수, n	분리 시간, t	평균절대 오차, mae	학습 반복수, n	분리 시간, t	평균절대 오차, mae	학습 반복수, n	분리 시간, t	평균절대 오차, mae	학습 반복수, n	분리 시간, t	평균절대 오차, mae
벽지	9.4	12.31	0.96	9.4	12.78	0.96	9.1	11.74	0.92	9.1	12.19	0.92
지문(1)	10.0	13.10	0.06	10.0	13.60	0.06	8.5	10.97	0.03	8.5	11.39	0.03
얼굴(1)	12.8	16.77	0.79	12.8	17.41	0.79	9.6	12.38	0.73	9.6	12.86	0.73
도장(1)	11.2	14.67	0.35	11.2	15.23	0.35	7.8	10.06	0.36	7.8	10.45	0.36
지문(2)	9.8	12.84	0.91	9.8	13.33	0.91	7.5	9.66	0.84	7.5	10.05	0.84
얼굴(2)	10.3	13.49	2.20	10.3	14.00	2.20	8.6	11.09	2.18	8.6	11.27	2.18
지문(3)	16.0	20.96	0.08	16.0	21.76	0.08	12.2	15.74	0.04	12.2	16.35	0.04
레나	8.3	10.87	4.55	8.3	11.29	4.55	7.4	9.55	4.26	7.4	9.92	4.26
도장(2)	9.5	12.45	1.12	9.5	12.92	1.12	7.1	9.16	1.10	7.1	9.51	1.10
지문(4)	1.0	1.31	0.50	1.0	1.36	0.50	1.0	1.29	0.50	1.0	1.34	0.50
계	98.3	128.77	11.52	98.3	133.68	11.52	78.8	101.62	10.96	78.8	105.33	10.96
평균	9.83	12.88	1.15	9.83	13.37	1.15	7.88	10.16	1.10	7.88	10.53	1.10

4. 결 론

본 논문에서는 뉴우턴법과 할선법의 고정점 알고리즘 각각에 첨도를 추가하여 군집성을 고려한 독립성분분석을 제안하였다. 여기에서 고정점 알고리즘은 학습파라미터의 설정과 무관한 빠른 성분분석과 좀 더 우수한 분석성능을 얻기 위함이고, 첨도는 유사한 속성을 가지는 성분의 군집화된 분석순서를 얻기 위함이다.

제안된 독립성분분석을 500개 샘플을 가지는 6개의 혼합신호와 512×512 픽셀을 가지는 10개의 혼합영상의 분리에 각각 적용하여 실험한 결과, 제안된 두 가지 ICA들은 항상 일정한 분석순서를 유지하여 기존의 ICA들에서 알고리즘의 수행 때마다 랜덤하게 변하는 분석순서의 제약을 해결할 수 있었다. 특히 군집화의 속성을 가진 제안된 독립성분분석들은 신호나 영상의 분류나 식별에도 적용할 수 있음을 확인하였다. 또한 신호와 영상을 대상으로 한 실험의 결과들을 비교할 때, 할선법의 ICA들이 뉴우턴법의 ICA들보다

빠르면서도 우수한 분리성능이 있으며, 문제의 규모가 클수록 그 성능 개선의 정도도 증가함을 알 수 있다.

향후 제안된 ICA를 좀 더 다양한 큰 규모의 문제와 분류나 식별을 위한 문제에 적용하는 연구가 지속적으로 이루어져야 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] P. Comon, "Independent Component Analysis - A New Concept?," *Signal Processing*, Vol.36, No.3, pp.287-314, Apr., 1994.
- [3] A. Hyvarinen, "Survey on Independent Component Analysis," *Neural Computing Surveys*, Vol.2, pp.94-128, Apr.,

1999.

[4] A. Hyvarinen, J. Karhunen and E. Oja, '*Independent Component Analysis*,' John Wiley & Sons, Inc., New York, May, 2001.

[5] T. W. Lee, '*Independent Component Analysis : Theory and Applications*,' Kluwer Academic Pub., Boston, Dec., 1998.

[6] S. Roberts and R. Everson, '*Independent Component Analysis : Principles and Practice*,' Cambridge Univ., 2001.

[7] A. Hyvarinen and E. Oja, "A Fast Fixed Point Algorithms for Independent Component Analysis," *Neural Computation*, Vol.9, No.7, pp.1483-1492, Oct., 1997.

[8] 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.

[9] Y. H. Cho and Y. S. Park, "Independent Component Analysis of Fixed Point Algorithm Based on Secant Method," *Proc. of the International Conference on MLMTA'03*, Las Vegas, U.S.A., pp.302-306, June, 2003.

[10] S. C. Chapra and R. P. Canale, '*Numerical Methods for Engineering*,' McGraw-Hill, Inc., 1988.

[11] J. W. Hines, '*MATLAB Supplement to Fuzzy and Neural Approaches in Engineering*,' John Wiley & Sons, Inc., June, 1997.



조 용 현

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

1979년 경북대학교 공과대학 전자공학과
(공학사)

1981년 경북대학교 대학원 전자공학과
(공학석사)

1993년 경북대학교 대학원 전자공학과
(공학박사)

1983년~1984년 삼성전자(주) 사원

1984년~1987년 한국전자통신연구소 연구원

1987년~1997년 영남이공대학 전자과 부교수

1997년~현재 대구가톨릭대학교 컴퓨터정보통신공학부 부교수

관심분야 : 신경망, 병렬분산처리, 신호처리, 영상인식, 교환기 등