

Gini 계수를 이용한 Blind Source Recovery 방법의 구현

Implementation of Blind Source Recovery Using the Gini Coefficient

정재웅*, 송은정**, 박영철***, 윤대희*

(Jae-Woong Jeong*, Eun-Jung Song**, Young-Cheol Park***, Dae-Hee Youn*)

*연세대학교 전기전자공학과 **삼성전자 선행연구팀 ***연세대학교 원주캠퍼스 정보기술학부
(접수일자: 2006년 9월 28일; 수정일자: 2007년 11월 27일 채택일자: 2007년 12월 11일)

UBSS (under-determined blind source separation)는 BMMR (blind mixing matrix recovery) 과정과 BSR (blind source recovery) 과정으로 구분된다. 일반적으로 이 두 과정은 취득된 데이터의 sparseness를 이용하여 수행되는데, 얼마나 sparseness를 정확히 측정하느냐에 따라 그 성능이 좌우된다. 본 논문에서는 Gini 계수를 이용한 sparseness의 측정 방법을 BSR 과정에 도입하여, l_1 -노름, l_q -노름과 쌍곡탄젠트 (hyperbolic tangent)를 이용하는 측정 방법들과 비교하였으며, 보다 정확한 sparseness 측정과 향상된 BSR 성능을 획득하였다. 이는 컴퓨터 모의 실험을 통하여 검증되었다.

핵심용어: blind source recovery, Gini 계수, l_1 -노름, l_q -노름, sparseness, under-determined blind source separation, 쌍곡탄젠트

투고분야: 음향 신호처리 분야 (1,2)

UBSS (under-determined blind source separation) is composed of the stages of BMMR (blind mixing matrix recovery) and BSR (blind source recovery). Generally, these two stages are executed using the sparseness of the observed data, and their performance is influenced by the accuracy of the measure of the sparseness. In this paper, as introducing the measure of the sparseness using the Gini coefficient to BSR stage, we obtained more accurate measure of the sparseness and better performance of BSR than methods using the l_1 -norm, l_q -norm, and hyperbolic tangent, which was confirmed via computer simulations.

Key words: Blind source recovery, Hyperbolic tangent, Gini coefficient, l_1 -norm, l_q -norm, Sparseness, Under-determined blind source separation

ASK subject classification: Acoustical Signal Processing (1,2)

I. 서론

신호 분리 기술은 여러 개의 신호가 혼합되었을 때, 혼합된 신호로부터 원신호를 추출하는 기술이며, 여러 개의 음성 신호가 공간을 통과하여 마이크로폰과 같은 센서에 의해 취득되었을 때, 이 신호들로부터 혼합되기 이전의 음성 신호를 분리해 내는 것을 한 예로 들 수 있다. 이러한 신호 분리 기술 중에서 특별히 원신호와 그 것이 혼합되는 과정에 대한 어떠한 정보도 이용하지 않는 것을 BSS (blind source separation)라고 지칭한다.

BSS를 원신호와 센서의 수를 기준으로 분류할 수 있으며, 원신호의 수에 비해 센서의 수가 적은 경우를 UBSS (under-determined BSS)라고 하는데, 일반적으로 UBSS과정은 BMMR (blind mixing matrix recovery)와 BSR (blind source recovery)의 두 과정으로 구분되어 수행된다 [1]. 특별히 UBSS 경우 원신호를 분리하기 위해서는 비선형적인 방법이 요구되며, 그 중 하나로 원신호의 sparseness를 이용하여 분리하는 방법이 널리 이용되고 있다.

Sparseness가 높다는 것은 어떠한 신호의 대부분의 에너지가 적은 지점에 집중되어 분포되어 있다는 것을 의미하며, 낮다는 것은 넓은 지점에 고르게 분포하고 있다는 것을 의미한다. 높은 sparseness를 갖는 한 가지

예로 평균값에 대부분의 에너지가 집중되어 있는 라플라시안 확률을 갖는 신호를 들 수 있다. Sparseness는 신호가 갖는 고유적인 특성이지만 특정한 변환을 통하여 그 정도를 변화시킬 수 있으며, 일반적으로 국소푸리에 변환 (STFT: short-time Fourier transform)이나 웨이블릿변환 등의 변환 방법들을 이용하여 sparseness를 증가시킬 수 있다.

Sparseness는 혼합 행렬을 추정하는 BMMR 과정이나 원신호를 복원하는 BSR의 성능 향상에 큰 영향을 주는데, 이것은 추정하려고 하는 원신호의 sparseness가 높을수록 혼합 환경의 특성이 취득된 혼합 신호에 보다 잘 드러나게 되어 용이하게 혼합 행렬을 추정할 수 있기 때문이다. 현재까지 sparseness 특성을 이용하는 방법이 많이 제안되었는데, P. Bofill과 M. Zibulevsky [2]는 sparseness를 이용하여 혼합 행렬의 열벡터를 추정하고, 이를 이용해 원신호를 복원하는 일련의 과정에 대해 제안하였고, sparseness를 이용하여 혼합 행렬을 추정하는 방법으로서 O. Yilmaz와 S. Rickard [3]는 DUET (degenerate un-mixing estimation technique)을 F. Abrard와 Y. Deville [4]은 TIFROM (time-frequency ratio of mixture approach)을 제안하였다. 그리고, I. Takigawa [5] 등은 sparseness를 이용하여 원신호를 효율적으로 복원하는 방법을 제안하였다. 이처럼 sparseness가 높을수록 각 과정의 성능은 모두 향상될 수 있으나, 본 논문에서는 특별히 BSR 과정에 대해서 sparseness를 측정하는 방법을 달리 함으로써 나타날 수 있는 성능의 차이를 보이고 실험을 통해 이를 검증하였다.

본 논문은 다음과 같이 구성된다. 2장에서는 l_0 -노름, l_1 -노름과 쌍곡탄젠트, 그리고 Gini 계수를 정의하고, 이를 이용한 BSR 시스템의 구성에 대해 기술하며, 3장에서는 l_0 -노름, l_1 -노름과 쌍곡탄젠트, 그리고 Gini 계수를 이용한 sparseness의 측정과 이들 사이의 관계에 대해 기술한다. 4장에서는 컴퓨터 모의실험을 통하여 성능을 검증하며, 끝으로 5장에서 결론을 맺는다.

II. 노름, 쌍곡탄젠트와 Gini 계수

2.1. 노름과 쌍곡탄젠트

Sparseness의 측정 방법은 여러 가지가 있으나, 실제로 확정된 표준은 없다. 엄밀하게 어떤 신호의 sparseness가 높다는 것은 대부분의 신호가 0인 상태를

의미하지만, 실제적으로는 0에 근접한 매우 작은 값들이 존재하기 때문에, 몇몇의 신호만이 큰 값을 갖고 대부분의 신호는 상대적으로 매우 작은 값을 갖는 상태를 의미하게 된다. Sparseness를 측정하는 방법들로는 노름, 쌍곡탄젠트, 대수 (logarithm) 등을 이용하는 것들이 있으나, 일반적으로 노름 (l_0 -노름, l_1 -노름, l_q -노름 등)과 쌍곡탄젠트가 널리 이용되며, 이는 식 (1)과 같이 수식화된다 [6].

$$\begin{aligned} l_0 : \| \mathbf{x} \|_0 &= \# \{ j, x_j \neq 0 \} / N \\ l_1 : \| \mathbf{x} \|_1 &= \sum_j |x_j| \\ l_q : \| \mathbf{x} \|_q &= \left(\sum_j |x_j|^q \right)^{1/q} \\ \tanh_{a,b} &= \sum_j \tanh(ax_j^b) \end{aligned} \quad (1)$$

여기서 $\#(\cdot)$ 는 조건을 만족하는 데이터의 수를 의미하며 N 은 데이터의 총수를 나타낸다. 또한, a 와 b 는 양의 정수이며 특히 b 는 반드시 2이상의 수여야 한다 [6].

Sparseness의 정확한 의미는 l_0 -노름으로 표현될 수 있기 때문에 이를 이용하면 최대의 sparseness를 갖는 완벽한 해답을 구할 수 있는 장점이 있으나, 그것을 푸는 방법이 완전히 최적화되어 있지 않을뿐더러 잡음에 대해서 강인하지 않는 한계가 있다. l_1 -노름을 이용하면 l_0 -노름과는 달리 유일하며 잡음에 대해서도 강인한 해답을 구할 수 있다. 그러나, l_1 -노름은 충분히 sparseness가 클 경우에만 좋은 성능을 내는 한계를 가지고 있다 [5]. l_q -노름은 q 를 0과 1사이에서 변화시키며 sparseness의 측정을 할 수 있으며, $0.1 < q < 0.5$ 일 때 좋은 성능을

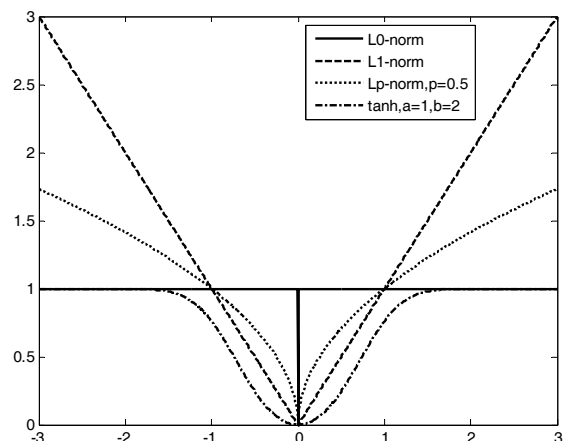


그림 1. 다양한 측정 방법에 따른 sparseness에 대한 영향
Figure 1. Sparseness contribution as various measures.

보이는 것으로 알려져 있다 [7]. 그리고 쌍곡탄젠트는 l_4 -노름을 근사화한 것인데, 이는 노름이 가지고 있는 미분이 불가능한 문제를 극복하기 위한 한 방법이며, 값이 커질수록 1에 접근하는 특성을 갖는다 [6]. 앞서 설명한 각각의 방법들이 sparseness 어느 정도의 영향을 주는 지에 대해 그림 1에 도시하였다.

2.2 Gini 계수

$\hat{s} = [\hat{s}_1, \dots, \hat{s}_N]$ 라는 데이터가 존재하며 $|\hat{s}_1|^p \leq |\hat{s}_2|^p \leq \dots \leq |\hat{s}_N|^p$ 와 같이 오름차순으로 정렬되어 있을 때, Lorenz 곡선은 식 (2)와 같이 정의 된다 [8].

$$L_p\left(\frac{i}{N}\right) = \frac{\sum_{j=1}^i |\hat{s}_j|^p}{\sum_{k=1}^N |\hat{s}_k|^p}, \text{ for } i = 0, \dots, N \quad (2)$$

여기서 $L_p(0) = 0$ 이며, $L_p(1) = 1$ 이 되며, p 는 Lorenz 곡선의 형태를 결정하는 인자이다. 그림 2에 Lorenz 곡선 ($p=2$)과 Gini 계수와와의 관계를 나타내었다.

그림 2에서 Lorenz 곡선을 기준으로 아래 부분의 영역을 식 (3)으로 계산할 수 있다.

$$A(\hat{s}) = \frac{1}{2N} \sum_{n=1}^N \left(L_2\left(\frac{n-1}{N}\right) + L_2\left(\frac{n}{N}\right) \right) \quad (3)$$

Gini 계수는 그림 2에서 Lorenz 곡선과 등가곡선 (line of equality) 사이 영역의 2배에 해당하는 값으로 식 (4)와 같이 정의된다.

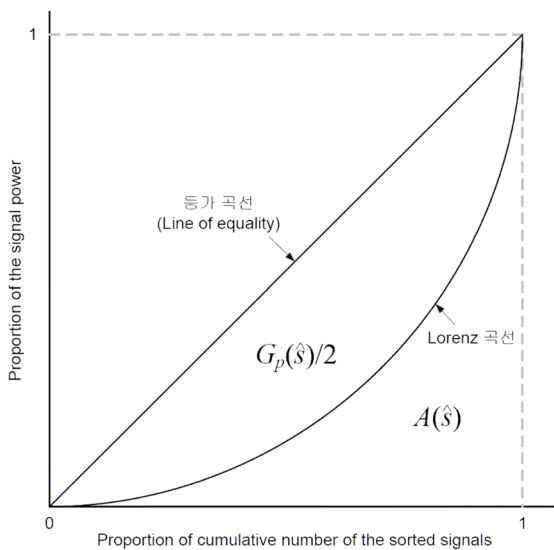


그림 2. Gini 계수와 Lorenz 곡선 ($p=2$)
Figure 2. Gini coefficient and Lorenz curve ($p=2$).

$$G_p(\hat{s}) = 1 - 2A(\hat{s}) \quad (4)$$

이와 같이 정의된 Gini 계수의 값은 모든 데이터가 분포할 수 있는 영역 내에 동일하게 분포되어 있을 때 최소가 되고, 적은 영역에 분포할수록 (sparseness가 높을수록) 값이 증가하게 되며 한 지점에만 분포해 있을 경우 최대값을 갖게 된다 [9].

III. Sparseness의 측정과 BSR 방법

3.1. 노름과 쌍곡탄젠트를 이용한 BSR 방법

l_4 -노름을 이용한 sparseness의 측정은 다음과 같이 수식화 될 수 있다. 취득된 신호 x 와 추정된 혼합 행렬 \hat{A} 이 주어졌을 때, \hat{s} 을 구하기 위해서 일반적으로 최대우도원리 (maximum likelihood)를 이용하며, 이것은 곧 후험적 (a posteriori) 확률을 최대화하는 \hat{s} 을 찾는 것을 의미한다 [10]. \hat{s} 의 확률 $P(\hat{s})$ 를 라플라시안 확률로 가정하였을 때, 식 (5)와 같이 \hat{s} 은 l_4 -노름을 최소화하는 것으로 유도된다.

$$\begin{aligned} \hat{s} &= \arg \max_{\hat{A} \hat{s} = x} P(\hat{s} | x, \hat{A}) = \arg \max_{\hat{A} \hat{s} = x} P(x | \hat{s}, \hat{A}) P(\hat{s}) = \arg \max_{\hat{A} \hat{s} = x} P(\hat{s}) \\ &= \arg \max_{\hat{A} \hat{s} = x} e^{-\lambda(|\hat{s}_1| + \dots + |\hat{s}_N|)} = \arg \min_{\hat{A} \hat{s} = x} |\hat{s}_1| + \dots + |\hat{s}_N| \quad (5) \\ &= \arg \min_{\hat{A} \hat{s} = x} \|\hat{s}\|_1 \end{aligned}$$

결국 l_4 -노름을 이용한 최적화 문제는 식 (6)과 같이 표현된다.

$$\min \sum_{j=1}^N |\hat{s}_j|, \text{ subject to } \hat{A} \hat{s} = x \quad (6)$$

또한 l_4 -노름과 쌍곡탄젠트를 이용한 최적화 문제도 그림 1에서 볼 수 있듯이 sparseness에 주는 영향이 상이할 뿐 같은 형태로, 즉 각각을 최소화 하는 것으로 표현될 수 있다.

BSR 과정은 BMMR 과정으로부터 추정된 혼합 행렬을 이용하여 원신호를 복원하는 과정이다. UBSS에 있어서 (특히 순간적인 혼합 환경에서 [11]) 추정된 혼합 행렬은 정방행렬이 아니기 때문에 단순한 역행렬을 통해 신호를 복원할 수 없다. 따라서, I. Takigawa [5] 등은 combinatorial method를 제안하였는데 그 과정은 다음

과 같다. 먼저 혼합 행렬을 구성하는 여러 개의 열벡터 중에 일부를 선택하여 다양한 조합의 정방행렬을 구성하며, 이를 이용하여 식 (7)과 같이 여러 개의 원신호의 추정치들을 각각 형성한다.

$$\hat{\mathbf{s}}_i = [\mathbf{a}_{j_1} \cdots \mathbf{a}_{j_M}]^{-1} \mathbf{x}, \quad i = 1, \dots, N, \mathbf{C}_M, \quad j_1, \dots, j_M \in \{1, \dots, N\} \quad (7)$$

여기서 M 은 마이크로폰의 수를, N 은 원신호의 수를 의미하며, 총 $N C_M$ 개의 추정치가 형성된다. 이렇게 형성된 여러 개의 추정치 $\hat{\mathbf{s}}_i$ 중에서 가장 sparseness가 큰 것을 최종적인 추정치 $\hat{\mathbf{s}}$ 으로 결정하게 되는데, sparseness를 측정하는 과정에서 앞서 설명한 여러 가지 측정방법이 적용될 수 있다.

3.2. Gini 계수를 이용한 BSR 방법과 노름 또는 쌍곡탄젠트를 이용한 BSR 방법의 비교

앞장에서 살펴본 바와 같이 Gini 계수가 커질수록 신호의 sparseness는 증가하게 되므로 식 (8)과 같이 이것을 최대화하는 값을 추정치 $\tilde{\mathbf{s}}$ 로 결정한다.

$$\max G_p(\tilde{\mathbf{s}}), \quad \text{subject to } \hat{\mathbf{A}}\tilde{\mathbf{s}} = \mathbf{x} \quad (8)$$

l_1 -노름을 이용하는 sparseness 측정 방법에 의한 BSR은 식 (6)과 같이 데이터의 절대값의 합이 최소화되는 것을 추정치 $\hat{\mathbf{s}}$ 으로 결정하며, l_1 -노름이나 쌍곡탄젠트를 이용하는 방법 역시 그림 1에서 볼 수 있듯이 sparseness에 대한 영향에 있어서 약간의 차이가 있을 뿐 크게 다르지는 않다. 반면, 식 (8)과 같이 Gini 계수를 최대화 한다는 것은 식 (4)의 $A(\tilde{\mathbf{s}})$ 을 최소화 하는 것이므로, 이는 식 (3)에서 확인할 수 있듯이 절대값의 p -제곱 형태의 누적치의 합이 최소화되는 것을 추정치 $\tilde{\mathbf{s}}$ 로 결정하게 된다. 특별히 $p=1$ 일 경우에는 오름차순으로 나열되어 있는 값들의 절대값의 누적치의 합이 최소화되는 것을 의미하게 된다. 이것은 Gini 계수를 이용하는 방법이 sparseness가 낮은 신호에 대해서는 l_1 -노름을 이용하는 방법과 비슷한 값을 나타내지만, sparseness가 높은 신호에 대해서는 Gini 계수를 이용할 경우 작은 값일수록 더 많이 누적되기 때문에 sparseness가 극명해지며, 이것은 보다 정확하게 sparseness를 측정할 수 있다는 것을 의미한다. 이와 같은 사실을 다음 장에서 실험을 통해 검증하였다.

IV. 모의 실험

4.1. 노름, 쌍곡탄젠트와 Gini 계수의 sparseness 측정 비교

l_1 -노름, l_q -노름, 쌍곡탄젠트, 그리고 Gini 계수를 이용하는 네 가지 방법으로 sparseness를 측정하기 위해 다음과 같은 실험을 하였다. 실험 대상 신호는 1000 샘플 길이의 가우시안 확률을 갖는 4개의 신호와 라플라시안 확률을 갖는 1개의 신호이며, -1과 1사이 분포하고 동일한 평균값과 에너지를 갖게 하였다. 이 신호들에 대해 위의 네 가지 방법으로 각각 sparseness를 측정하고, 라플라시안 확률을 갖는 신호가 전체에 대해 차지하는 비율을 계산하였으며, 1000번 반복하여 평균값을 취하였다. 노름과 쌍곡탄젠트를 이용한 방법은 그것을 최소화 하는 것이고 Gini 계수는 그것을 최대화하는 것이므로, Gini 계수를 그대로 이용하지 않고 식 (4)의 A 값을 이용하였으며, $p=2$ 를 사용하였다. l_q -노름에 대해서는 $q=0.4$ 를 사용하였으며, 쌍곡탄젠트에 대해서는 $a=1$, $b=2$ 를 사용하였다.

5개의 신호 중 라플라시안 확률을 갖는 신호, 즉 가장 sparseness가 높은 신호가 전체 신호에 대해 차지하는 비율이 낮다는 것은 실제 BSR 과정에서 여러 개의 추정치 중에서 최종 추정치를 선택할 때 선택되는 확률이 높아지는 것을 의미하며, 이는 BSR 성능의 향상으로 이어질 수 있다. 실험에 대한 결과는 표 1에 도시하였다.

표 1. sparseness가 가장 높은 신호가 전체에 대해 차지하는 비율
Table 1. Ratio of the sparsest signal to the whole signal.

l_1 노름 (%)	l_q 노름 (%)	\tanh (%)	Gini 계수 (%)
19.3	18.7	19.2	15.4

표 1에서 볼 수 있듯이 라플라시안 확률을 갖는 신호가 차지하는 비율이 Gini 계수를 사용한 방법에서 약 3~4%정도 낮게 나타나는 것을 확인 할 수 있으며, 이는 Gini 계수를 사용할 때 노름이나 쌍곡탄젠트를 사용할 경우와 비교하여 BSR 과정에 있어서 sparseness가 높은 신호가 선택될 확률이 높은 것을 의미하며, 또한 BSR 성능이 향상될 수 있음을 의미한다.

4.2. BSR 성능 평가

BSR의 성능을 평가하기 위해 다음과 같은 실험을 하였다. 3개의 음원 ($N=3$)이 혼합되어 2개의 마이크로폰 ($M=2$)으로 취득되는 2×3 의 혼합 행렬에 대해 실험하였

고, 완벽한 BMMR 과정을 통해 오차가 없는 혼합 행렬이 추정된 것으로 가정하였다. 각각의 결과는 무작위로 발생된 20개의 혼합 행렬에 대한 결과의 평균값으로 나타내었으며, 3개의 음원의 순서를 무작위로 선정한 3개의 세트에 대하여 실험하였다. 6명의 남성 화자와 5명의 여성 화자가 읽은 11개의 한국어 문장을 원신호로 사용하였으며, 각각의 원신호는 16kHz로 샘플링 되었고, 5초 길이로 사용되었다. Gini 계수의 값을 구하기 위해

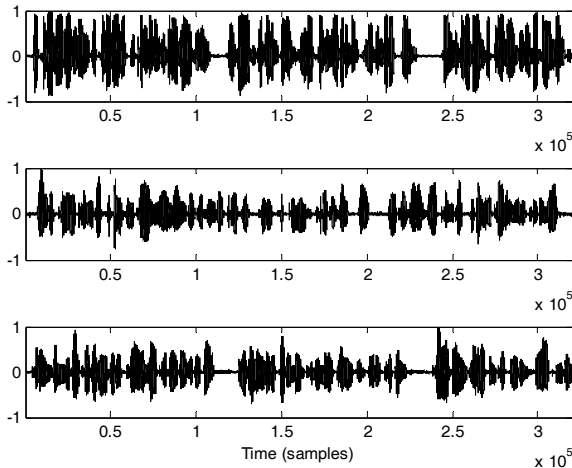


그림 3. 혼합되기 이전의 음성신호
Figure 3. Original speech signals before mixing.

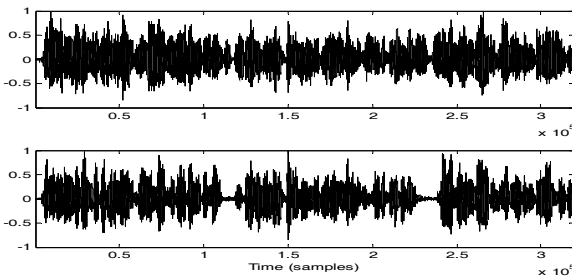


그림 4. 혼합된 음성신호
Figure 4. Mixed speech signals.

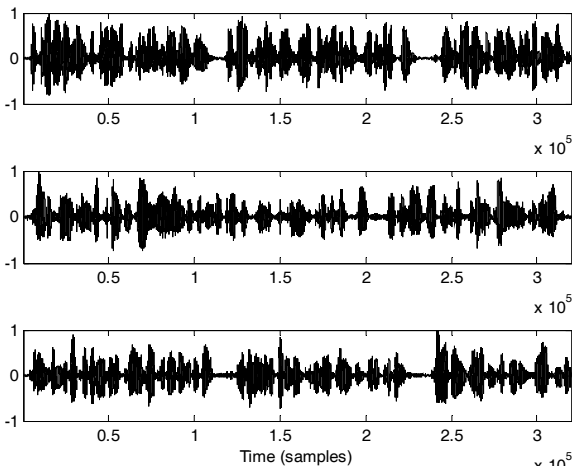


그림 5. 분리된 음성신호
Figure 5. Separated speech signals.

$p=2$ 를 사용하였으며, l_1 -노름에 대해서는 $q=0.4$ 를, 쌍곡 탄젠트에 대해서는 $a=1, b=2$ 를 사용하였다.

위와 같은 조건으로 신호를 분리하기 위해서 먼저 혼합된 신호를 국소푸리에 변환하여 시간-주파수 영역으로 변환 하였는데, 이는 시간-주파수 영역에서는 시간 영역에 비해 일반적으로 신호의 sparseness가 크게 나타나며, 특별히 Gini 계수에 대해서도 그러하기 때문이다 [9]. 이렇게 시간-주파수 영역으로 변환된 신호에 대해 BSR과정을 수행하였다. 그리고 역국소푸리에 변환 (inverse STFT)을 통해 시간 영역으로 재변환 하였다. 그림 3~5는 각각 혼합되기 이전의 원신호, 혼합된 신호, 분리된 신호를 나타내고 있다. 그림 4에서 볼 수 있듯이 혼합된 신호는 그림 3의 원신호에 비해 sparseness가 떨어지는 것을 확인할 수 있으며, 그림 5의 분리된 신호는 원신호의 수준으로 복원되는 것을 확인할 수 있다.

보다 객관적인 BSR의 성능을 평가를 위해 source-to-distortion ratio (SDR), source-to-interference ratio (SIR), source-to-artifacts ratio (SAR)의 평가 기준을 사용하였다. 각각을 정의하기 위해 추정된 신호 $\hat{\mathbf{s}}$ (또는 $\tilde{\mathbf{s}}$) 를 식(9)와 같이 분해할 수 있으며 [12],

$$\hat{\mathbf{s}} = \hat{\mathbf{s}}_{target} + \mathbf{e}_{interf} + \mathbf{e}_{artif} \quad (9)$$

이를 이용하여 각각의 평가 기준을 정의하면 식 (10)~(12)와 같다.

$$SDR = 10 \log_{10} \frac{\sum \hat{s}_{target}^2}{\sum (e_{interf} + e_{artif})^2}, \quad (10)$$

$$SIR = 10 \log_{10} \frac{\sum \hat{s}_{target}^2}{\sum e_{interf}^2} \quad (11)$$

$$SAR = 10 \log_{10} \frac{\sum (\hat{s}_{target} + e_{interf})^2}{\sum e_{interf}^2}. \quad (12)$$

SDR, SIR, SAR이 모두 성능 평가의 기준으로 사용될 수 있지만, 신호 분리의 성능 평가에 있어서는 상호 혼합되어 있는 신호가 얼마나 제거 되었는지가 중요하기 때문에, SIR이 보다 중요한 의미를 갖는다고 할 수 있다.

표 2에 각각의 방법을 사용하였을 때의 SDR, SIR, SAR의 결과값을 나타내었다. 표 2에서 볼 수 있듯이,

표 2. l_1 노름, l_q 노름, 쌍곡탄젠트, Gini 계수를 이용한 BSR의 성능
 Table 2. Performances of BSR using l_1 norm, l_q norm, hyperbolic tangent and Gini coefficient.

l_1 노름			l_q 노름		
SDR(dB)	SIR(dB)	SAR(dB)	SDR(dB)	SIR(dB)	SAR(dB)
11.68	18.78	11.80	10.46	19.30	11.14
tanh			Gini 계수		
SDR(dB)	SIR(dB)	SAR(dB)	SDR(dB)	SIR(dB)	SAR(dB)
10.67	17.73	11.43	11.36	22.24	11.94

Gini 계수를 이용했을 때 여타의 다른 방법을 이용했을 때 보다 나은 BSR의 성능을 나타냈다. 물론 SDR과 SAR에 대해서는 큰 성능 향상은 발생하지 않았지만, 신호 분리라는 관점에서 성능을 평가할 수 있는 SIR의 성능은 다른 방법들이 20dB를 밑도는 것에 반해 약 3~4dB의 성능 향상이 있는 것으로 확인되었다.

V. 결론 및 토의

본 논문에서는 sparseness를 측정하는 방법으로서 Gini 계수를 소개하고 UBSS의 한 과정인 BSR에 이를 적용하는 방법에 대해 제안하였다. 또한 노름과 쌍곡탄젠트를 이용한 방법과 비교 설명하였으며 이 방법들 사이의 관계를 파악하였다. 컴퓨터 모의실험을 통해 sparseness 측정 능력과 BSR 성능을 평가하였으며, Gini 계수를 사용한 방법이 노름이나 쌍곡탄젠트를 이용한 방법 보다 좋은 sparseness 측정 능력과 BSR 성능을 나타내는 것을 확인하였다.

본 논문에서는 순간적인 혼합신호에 대한 이론과 실험에 대해 연구하였으나, 일반적으로 현실적으로 발생하는 신호들은 잔향이 있는 콘볼루션 혼합 신호가 대부분을 차지한다. 본 연구에서 진행한 순간 혼합 신호에 대한 결과는 그 자체로도 의미가 있지만, 주파수 영역에서 접근하는 콘볼루션 혼합 신호 분리의 연구 [13]에도 상당 부분 기여할 것으로 생각된다.

감사의 글

본 논문은 연세대학교 학술연구비 지원 (2002-5-0024)에 의한 것임.

참고 문헌

1. F. Theis and E. Lang, "Formalization of the two step approach to overcomplete BSS," Proc. of SIP 2002, 207-212, 2002.
2. P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," Signal Processing, **81**, 2353-2362, 2001
3. O. Yilmaz and S. Rickard, "Blind source separation of speech mixtures via time frequency masking," IEEE Trans. on Signal Processing, **52** (7), 1830~1847, July 2004.
4. F. Abrard and Y. Deville, "Blind separation of dependent signals using the time-frequency ratio of mixtures approach," Proc. 7th International Symposium on Signal Processing and its Applications (ISSPA 2003), 2003.
5. I. Takigawa, M. Kudo, and J. Toyama, "Performance analysis of minimum l_1 norm solution for underdetermined source separation," IEEE Trans. on Signal Processing, **52** (3), 582-591, 2004.
6. J. Karvanen and A. Cichocki, "Measuring sparseness of noisy signals," ICA03, 2003.
7. R. Saab, Ö. Yilmaz, M. J. McKeown, and R. Abugharbieh, "Underdetermined Sparse Blind Source Separation With Delays," Proceedings of Signal Processing with Adaptive Sparse Structured Representations Workshop (SPARS '05), Rennes, France, November 2005.
8. C. Gini, "Measurement of inequality of incomes," Economic Journal, **31**, 124-126, 1921.
9. S. Rickard and M. Fallon, "The Gini Index of Speech," Conference on Information Sciences and Systems, Princeton, NJ, USA, March 2004.
10. S. Winter, W. Kellermann, H. Sawada and S. Makino, "MAP based underdetermined blind source separation of convolutive mixtures by hierarchical clustering and l_1 norm minimization," EURASIP Journal on Advances in Signal Processing, vol. 2007, 2007
11. P. O'Grady, B. Pearlmutter, and S. Rickard, "Survey of sparse and non sparse methods in source separation," International Journal of Imaging Systems and Technology, **15** (1), 2005.
12. E. Vincent, R. Gribonval and C. Fevotte, "Performance measurement in blind audio source separation," IEEE Trans. on Audio, Speech and Language Processing **14**, (4), 1462-1469, 2006.
13. P. Smaragdis, "Blind Separation of Convolved Mixtures in the Frequency Domain," Neurocomputing, 22:21 **34**, 1998

저자 약력

• 정재웅 (Jae-Woong Jeong)



2003년 2월 : 연세대학교 공과대학 전기전자공학과 (공학사)
 2005년 8월 : 연세대학교 공과대학 전기전자공학과 (공학석사)
 2005년 9월~현재 : 연세대학교 공과대학 전기전자공학과 박사과정
 ※ 주요연구분야 : 디지털 신호처리, 오디오 신호처리, 3D 오디오, crosstalk cancellation, blind source separation, dereverberation

• 송 은 정 (Eun-Jung Song)



2004년 2월 : 연세대학교 공과대학 전기전자공학과
(공학사)
2007년 2월 : 연세대학교 공과대학 전기전자공학과
(공학석사)
2007년 3월~현재 : 삼성전자 선행연구팀
※ 주요연구분야 : 디지털 신호처리, 오디오 신호처리, 3-D
오디오

• 박 영 철 (Young-Cheol Park)



1986년 2월 : 연세대학교 전자공학과 졸업
1988년 2월 : 연세대학교 전자공학과 석사
1993년 2월 : 연세대학교 전자공학과 박사
현재 : 연세대학교 정보기술학부 교수
※ 주요연구분야 : 디지털 신호처리, 오디오 신호처리, 음성
신호처리, 적응 신호처리

• 윤 대 희 (Dae-Hee Youn)



1977년 2월 : 연세대학교 전자공학과 졸업
1979년 8월 : Kansas State Univ. 석사
1982년 2월 : Kansas State Univ. 박사
현재 : 연세대학교 전기전자공학과 교수
※ 주요연구분야 : 디지털 신호처리, 적응 신호처리, 음성
신호처리, 오디오 신호처리