

정준 상관 분석을 이용한 잡음 섞인 음성 신호의 분리

김선일*

*거제대학교

Segaration of Corrupted Speech Signals using Canonical Correlation Analysis

Seonil Kim*

*Koje College

E-mail : seonil@koje.ac.kr

요 약

음성 신호와 함께 섞인 자동차 배기 소음을 서로 분리해 내는 기술은 점점 음성을 중심으로 발전해가는 인터페이스를 현실화하는데 실질적으로 필요한 기술이다. 따라서 자동차 배기음이 섞인 음성 신호를 두 신호간의 독립성이 보장되지 않고 두 신호에 대한 사전 정보가 없는 상태에서 분리해 내기 위해 정준 상관 분석을 사용하여 두 신호를 분리해 내는 연구를 진행하였다. 정준 상관 분석을 이용하여 음성을 분리해 내기 위해서는 분석에 쓰이는 신호의 구성이 중요하다. 정준 상관 분석에 대해 알아보고 음성과 자동차 배기 소음이 섞인 두 개의 신호를 받아서 이를 재구성하여 정준 상관 분석을 이용하여 자동차 소음과 음성을 분리해 내었다. Blind Source Separation에 쓰이는 다른 방법과 비교했을 때 독립성이 보장되지 않는 신호에 대해서도 분리가 가능하므로 응용 대상이 상대적으로 넓어 실용적 응용이 가능하다고 할 수 있다.

ABSTRACT

The technology which is used for segregating voices signals from exhaust noise signals of a car is very practical one to realize the interfaces between men and machines using voices. The voice signals contaminated by exhaust noise signal of a car was separated by canonical correlation anagnosis(CCA) in an environment which does not guarantee the independence between signals and have prior informations. Rearrangement for the input signals is important in CCA. CCA was studied and segregation between source signals were performed by CCA through rearrangements of each of signals. It is possible to apply the technique to various signals since it is also possible to use CCA to the signals which are not independent

키워드

정준 상관 분석, CCA, 음성신호, 자동차 배기음, 분리

1. 서 론

텔레매틱스(Telematics)를 구현하고자 할 때 운 전의 안전성을 위해 궁극적으로 음성을 이용한 명령 전달이 이루어져야 한다. 하지만 자동차와 같은 환경에서는 주변 잡음에 노출될 수 밖에 없는 상황인데 그 중 우선 자동차 배기구에서 전달 되는 자동차 배기 소음이 가장 큰 부분을 차지한

다. 따라서 음성 신호와 자동차 배기 소음이 섞인 신호에서 음성 신호의 특성을 그대로 유지하면서 자동차 배기 소음을 어떻게 제거하느냐가 중요한 관건이 된다.

자동차 배기음은 음성 신호와 서로 연관성을 갖지 않아 상관관계가 거의 없다고 볼 수 있다. 따라서 신호 사이의 상관관계가 없다고 가정하고 정준 상관 분석(Canonical Correlation Analysis : CCA) 기법[1],[2]을 적용하여 두 개의 마이크로

섞여 들어온 각 음원들을 분리하고자 한다. 각 음원에 대한 정보는 전혀 없으므로 Blind Source Separation에 해당된다고 볼 수 있다. 따라서 두 개의 마이크에 두 개의 음원에서 나온 신호가 섞여서 입력되고 이들 신호에서 정준상관분석을 이용해 두 음원을 분리해 낸다.

II. 정준상관분석

평균값이 0 이고 n 개의 원소로 이루어진 랜덤 벡터 x 와 y 가

$$x = [x_1 \ x_2 \ \dots \ x_n]^T \quad (1)$$

$$y = [y_1 \ y_2 \ \dots \ y_n]^T \quad (2)$$

이고 x_i, y_i 가

$$x_i = [x_{i1} \ x_{i2} \ \dots \ x_{it}] \quad (3)$$

$$y_i = [y_{i1} \ y_{i2} \ \dots \ y_{it}] \quad (4)$$

인 신호일 때

$$z = [x \ y]^T \quad (5)$$

라고 하자. 이 때 랜덤 변수 x, y 사이의 공분산은

$$C = \begin{bmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{bmatrix} = E[z \ z^T] \quad (6)$$

여기서 x 와 y 사이의 정준 상관값은 다음 식을 풀어서 구할 수 있다.

$$C_{xx}^{-1} C_{xy} C_{yy}^{-1} C_{yx} w_x = \rho^2 w_x \quad (7)$$

$$C_{yy}^{-1} C_{yx} C_{xx}^{-1} C_{xy} w_y = \rho^2 w_y \quad (8)$$

벡터 $C_{xx}^{-1} C_{xy} C_{yy}^{-1} C_{yx}$ 의 고유벡터가 w_x 이고 고유값이 ρ^2 이며 $C_{yy}^{-1} C_{yx} C_{xx}^{-1} C_{xy}$ 의 고유벡터는 w_y 이고 고유값이 ρ^2 이다.

여기서 ρ 는

$$\rho = \frac{E[z_x' z_y]}{\sqrt{E[z_x' z_x] E[z_y' z_y]}} = \frac{w_x^T C_{xy} w_y}{\sqrt{w_x^T C_{xx} w_x w_y^T C_{yy} w_y}} \quad (9)$$

$$z_x = x^T w_x \quad (10)$$

$$z_y = y^T w_y \quad (11)$$

ρ 가 w_x 와 w_y 에 대해 최대값을 가질 때 이것을 최대 정준 상관이라고 하며 ρ 를 최대로 하는 w_x 와 w_y 는 앞 식에서 언급한 대로 고유값 문제를 풀어 구할 수 있다.

III. 신호의 재배열

자동차 배기음 신호와 음성 신호가 섞인 두 개의 신호 중 한 신호를 x_1 이라고 하고 또 다른 신호를 y_1 이라고 하자. 이 때 x_2, y_2 는

$$x_2 = [x_{22} \ x_{23} \ \dots \ x_{2i} \ \dots \ x_{2(t-1)}] \quad (12)$$

$$x_{2i} = x_{1(i-1)} + x_{1(i+1)} \quad (13)$$

$$y_2 = [y_{22} \ y_{23} \ \dots \ y_{2i} \ \dots \ y_{2(t-1)}] \quad (14)$$

$$y_{2i} = y_{1(i-1)} + y_{1(i+1)} \quad (15)$$

위 데이터를 이용하여 ρ 를 최대로 하는 w_x 나 w_y 를 구하면 분리된 음성 신호를 구할 수 있다.

IV. 실험

그림 1에 나타난 것과 같은 자동차 배기음 신호와 그림 2에 나타난 것과 같은 음성 신호를 하나는 0dB 다른 하나는 약 3dB 차이가 나도록 섞어 자동차에서 두 개의 마이크로폰을 통해 입력되는 것과 유사한 신호를 만들어 내었다(그림3, 그림 4). 이 신호를 각각 x_1, y_1 이라고 하고 식 (12), 식 (13), 식 (14), 식 (15)를 이용해 x_2, y_2 를 만들고 다음 식과 같은 x, y 벡터를 형성해 정준 상관 분석을 적용하여 그림 5, 그림 6과 같이 섞이기 전의 신호를 분리해 내었다.

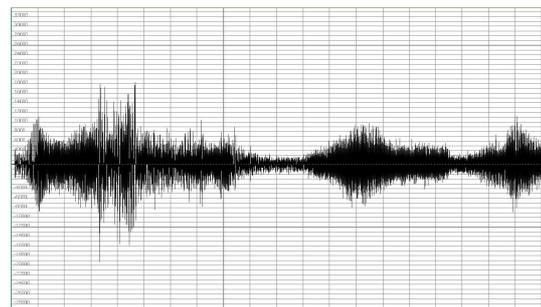


그림 1. 자동차 배기음

$$x = [x_1 \ x_2]^T \quad (16)$$

$$y = [y_1 \ y_2]^T \quad (17)$$

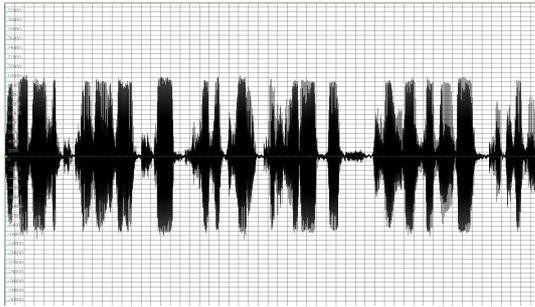


그림 2. 음성 신호

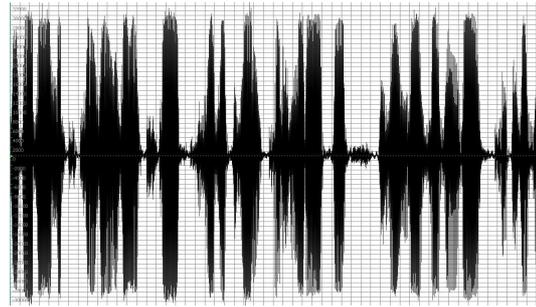


그림 6. 분리된 음성 신호

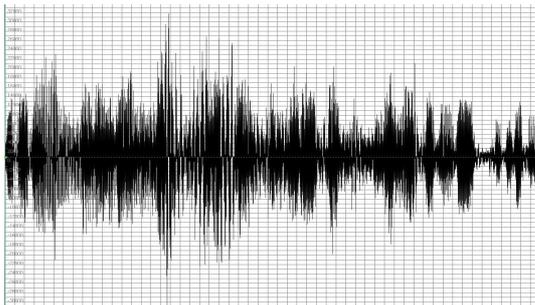


그림 3. 섞인 신호 A

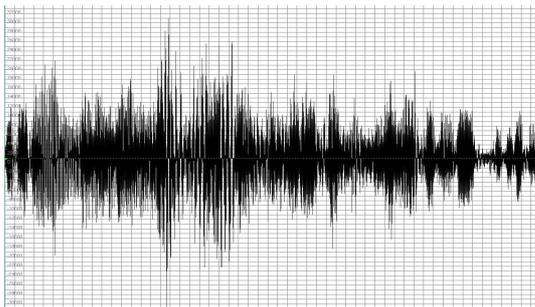


그림 4. 섞인 신호 B

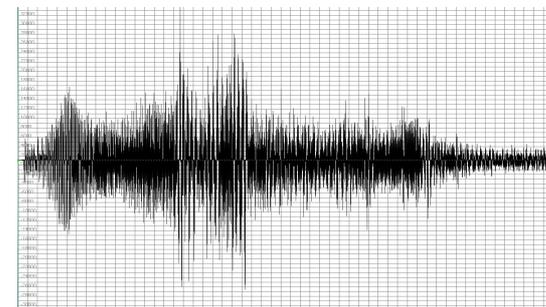


그림 5. 분리된 잡음

V. 결 론

Blind Source Separation에는 기존에 잘 알려진 Independent Component Analysis[3],[4],[5],[6] 방법이 있으나 정준 상관 분석보다 더 엄격한 전제 조건, 즉 두 신호가 서로 독립적이어야 한다는 조건을 충족시켜야 한다. 하지만 정준 상관 분석은 독립성보다 약한 상관 관계를 이용함으로써 많은 신호들이 이러한 전제조건을 충족시킬 수 있는 잇점이 있다. 그리고 ICA의 일부 알고리즘들은 합성 행렬을 구할 때 최소값 또는 최대값을 찾아가는 Gradient Method를 사용하게 되는데 이 때 사용되는 초기값에 따라 Local Minimum에 빠져서 실제 분리된 신호의 모양이 그 때 그 때 달라지는 경향을 실험을 통해서 확인할 수 있으나 정준 상관 분석은 그런 현상을 발견할 수 없었고 항상 일정한 파형이 관찰되었다.

정준 상관 분석이 적용되어야 할 대상과 ICA를 적용하여야 할 대상을 어떻게 구분할지와 어느 방법이 어느면에서 유리한지는 추가 연구를 이용해 알아봐야 할 과제이다.

참고문헌

- [1] W. Liu, D. Mandic, and A. Cichocki, "Analysis and Online Realization of CCA Approach for Blind Source Separation," IEEE Transaction on Neural Networks, Vol. 18, No. 3, September 2007.
- [2] D. Weenink, "Canonical Correlation Analysis," Institute of Phonetic Sciences, University of Amsterdam, Proceedings 25, pp. 81-99, 2003
- [3] A. Hyvarinen, "Survey on Independent Component Analysis," Neural Computing Surveys 2, pp. 94-128, 1999
- [4] 김선일, "잡음 섞인 한국어 인식을 위한 ICA 비교 연구," 대한전자공학회지, 제45권, IE편, 제3호, 2008. 9.
- [5] 김선일, "ICA로 분리한 신호의 분류," 대한전자공학회논문지, 제 47권 IE편, 제 4호, 2010. 12

[6] 김선일, "주파수 영역 자기 공분산 기울기를 이용한 음성과 자동차 소음 신호의 구분," 한국해양정보통신학회논문지, 제15권 제10호, pp. 2093-2099, 2011. 10.