

독립 성분 분석을 이용한 지각 필터의 성능 향상에 관한 연구

*구교식 **차형태

숭실대학교

*senia2@mms.ssu.ac.kr **hcha@ssu.ac.kr

A Study on Enhancement of Perceptual Filter's performance using Independent Component Analysis

*Koo, Kyo-Sik **Cha, Hyung-Tai

Soongsil University

요약

본 논문에서는 잡음이 첨가된 오디오 신호로부터 잡음을 추정하고 이에 따른 지각 필터 적용을 통한 음질 개선 알고리즘을 제안한다. 기존의 지각 필터는 고정된 잡음을 사용하여 잡음이 가변적일 경우 그 성능이 저하되었으며 독립 성분 분석만을 사용하여 잡음을 제거할 경우 잡음이 완전히 분리되지 못하고 잔류하게 된다. 그러나 제안된 잡음 추정 알고리즘은 독립성분 분석을 이용하여 잡음 에너지를 획득하고 이를 지각 필터에 적용함으로써 전 대역의 잡음 에너지를 효과적으로 제거할 수 있게 된다. 기존의 독립성분분석만을 이용한 방법과의 비교를 위하여 SSNR 비교를 수행하였고 그 결과를 통해 성능 개선을 확인 할 수 있었다.

1. 서론

오늘날 멀티미디어 기술의 발전은 영상과 음향에 대한 전반적인 기술의 발전을 가져왔다. 영상분야에 있어서는 화려한 색상과 넓은 화면, 고선명 화질, 3D 영상에 이르기 까지 눈부신 발전을 보여 왔다[1]. 그에 비례하여 음향 분야에서도 저장용량의 확대, 전송망의 개선, 처리 속도의 증가로 인하여 고음질에 대한 청자들의 요구가 급증하게 되었다. 그 중 무선 통신 기술과 휴대용 미디어 기기의 결합으로 영상통화와 같은 스피커폰 시스템의 사용이 증가함에 따라 잡음 제거와 관련된 기술이 더욱더 중요한 문제로 대두되게 되었다. 잡음은 음성 신호의 명료성을 떨어뜨리고 청자에게 불쾌감을 안겨주고 있다. 또한 멀티미디어와 관련된 시스템에서 배경 잡음은 시스템의 성능 저하와 관련된 주요 이유가 되고 있다. 이에 많은 알고리즘이 잡음 제거를 목적으로 개발되고 있으며 대표적인 알고리즘으로는 주파수 차감[3]이나 인간의 청각 특성을 이용한 방법[4][5] 등이 있다. 그러나 위의 방법 들은 잡음을 제거하기 위해 잡음에 대한 사전 정보가 필요하고 정보를 획득할 수 없을 시에는 묵음 구간을 통해 잡음을 추정한다[2]. 하지만 여러 환경에서 잡음의 특성은 시간에 따라 변화하므로 이를 효과적으로 제거하는 연구가 필요하다.

본 논문에서는 독립 성분 분석 기법 및 지각 필터를 이용하여 잡음을 효율적으로 제거하는 기법을 제안한다. 논문의 내용은 2장에서는 ICA 알고리즘과 지각 필터에 대해 설명하고 3장 및 4장에서 제안된 알고리즘에 대한 설명 및 시뮬레이션에 관한 결과를 제시한다. 마지막으로 5장에서 결론을 맺는다.

2. 독립성분분석 및 지각 필터

가. 독립성분분석(Independent Component Analysis)

독립 성분 분석 방법은 주성분 분석(principle component

analysis) 방법의 확장된 형태로써, 특정 신호를 구성하고 있는 독립된 성분들을 분리하는 방법으로 혼합된 신호로부터 그 신호의 기본 구성 요소를 이루는 성분을 분리해내는데 적용된다[6]. 주요 응용분야는 데이터 해석, Blind Source 분리, 특징 추출 기법 등이 있다. 주성분 분석 방법이 가우스 분포를 가지는 신호들 간의 상관성을 나타내는 공분산 행렬을 이용한 2차 선형변환 기법이라면, 독립 성분 분석 방법은 공분산행렬에 포함되지 않는 데이터의 비가우스(nongaussian) 분포 즉, 독립성을 가지는 변수에 대한 고차통계에 기초한 다차원 변환기법이다 [7].

독립 성분 분석을 정의하기 위하여 일반적으로 통계적인 모델(statistical model)을 사용한다. n 개의 관측된 랜덤 변수를 y_1, y_2, \dots, y_n 이라고 한다면 각각의 변수 y_i 는 n 개의 미지의 랜덤 변수 x_1, x_2, \dots, x_n 의 선형결합으로 이루어진다고 가정한다. 여기서 x_i 는 독립 성분으로써 직접 관측될 수 없는 latent 변수이다. 본 논문에서는 관측된 변수들인 \mathbf{y} 를 $\mathbf{y} = [y_1, y_2, \dots, y_n]^T$ 로 독립 성분들인 \mathbf{x} 를 $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ 로 표기한다. \mathbf{y} 와 \mathbf{x} 의 관계는 식 (1)과 같이 모델링될 수 있다.

$$\mathbf{Y} = \mathbf{A}\mathbf{X} = \sum_{i=1}^n a_i x_i \quad (1)$$

여기서 \mathbf{A} 는 $n \times n$ 의 역행렬이 존재하는 정방행렬로써 혼합 행렬(mixing matrix)이라고 하며, 혼합 행렬의 열벡터인 $\mathbf{a}_i, i=1, \dots, n$ 는 독립 성분 분석의 기저벡터(basis vector) 또는 기저 함수(basis function)라고 부른다. 만일 \mathbf{x} 의 확률 밀도함수가 주변밀도(marginal density) 함수의 곱으로 표현된다면, 즉 독립 성분으로 표현된다면, 혼합된 입력 데이터로부터 원 신호를 분리해 낼 수 있다. 이 통계적 모델을 독립 성분 분석이라고 한다. 결국 독립 성분 분석 방법의 목적은 독립이 아닌 기저를 사용하는 입력데이터 \mathbf{y} 를 독립된 새로운 기저를 사용하는 좌표계로 변환해 주는 선형 변환 \mathbf{W} 를 찾는 것이다. 식 (2)에서 \mathbf{u} 는 찾고자

하는 원 신호의 추정 값을 의미하며, 학습 벡터(training vectors)를 이용하여 \mathbf{u} 의 성분들을 최대한 독립적으로 만들어 감에 의하여 \mathbf{W} 를 추정한다. 결국 가장 이상적인 독립 성분은 \mathbf{W} 가 \mathbf{A} 의 역행렬이 될 때를 말한다[6].

$$\mathbf{U} = \mathbf{W}\mathbf{Y} = \mathbf{W}\mathbf{A}\mathbf{X} \quad (2)$$

2.2 지각필터

인간의 청각적인 특성을 이용하여 각 임계 대역(Critical Band)에서 주파수 변별력을 고려하여 각 대역에 해당하는 신호들에 대한 임계 대역 에너지와 마스킹 특성을 고려한 청각 자극 에너지를 계산할 수 있다. 이를 바탕으로 잡음의 열화도가 심하게 나타나는 고주파 영역에서는 신호의 에너지보다 잡음의 에너지가 더 크게 나타난다는 특성을 이용하여 잡음을 추정하는 필터를 설계할 수 있다[8].

원래의 오디오 신호와 잡음 신호를 $x(n)$, $d(n)$ 이라 할 때, 잡음에 열화 된 신호 $y(n)$ 의 표현은 식 (3)과 같다.

$$y(n) = x(n) + d(n) \quad (3)$$

이 때 어떤 임계 대역에 대한 $x(n)$ 과 $y(n)$ 의 전력 스펙트럼을 $X(\omega, i)$, $Y(\omega, i)$ 라고 정의할 때, 신호와 잡음이 서로 상관관계가 없다고 한다면 신호와 잡음의 자극 에너지는 식(4)와 같은 관계를 가지게 된다.

$$Y(\omega, i) = X(\omega, i) + D(\omega, i) \quad (4)$$

이 때 i 는 시간 프레임에서의 인덱스를 의미하며 $X(\omega, i)$ 로 인한 기저막에서의 임계 대역 에너지는 식 (5)로 표현 가능하다.

$$X_a(z, i) = a_0(z)X_i(z, i) = a_0(z)\sum_{\omega_i}^{\omega_{zh}} X(\omega, i) \quad (5)$$

where, $0 \leq z \leq Z-1$

여기서 ω_{zh} 와 ω_{zl} 은 전체 임계 대역 Z 에서의 임의의 임계대역 z 의 주파수 경계를 나타낸다. $a_0(z)$ 는 외이(外耳)에서부터 중이(中耳)까지의 다양한 전달 특성 요소(transmission factor)들에 의한 주파수 의존 감쇠 특성을 나타내는 감쇠 성분이다. 단, 이때 $a_0(z)$ 는 외이부터 중이까지의 감쇠 특성을 나타낸 것이므로, 신호를 직접 내이(內耳)로 전달하는 경우 또는 난청 환경과 같이 이러한 특성의 변화가 있을 경우를 제외하고는 그 영향을 무시할 수 있다.

더불어 임계 대역에 대한 청각 자극 에너지는 $X_a(z, i)$ 과 기저막(basilar membrane)에서의 신호 에너지의 확산 현상을 나타내는 지각적인 에너지 확산 함수(spreading function)와의 컨볼루션(convolution)을 통하여 기저막에서의 신호 에너지에 대한 응답을 나타낼 수 있다. 이는 식 (6)으로서 나타낼 수 있다[8].

$$X_e(z, i) = \sum_{v=0}^{Z-1} [SF(z, z)X_a(v, i)] \quad (6)$$

where, $0 \leq z \leq Z-1$

여기서, 확산 함수 $SF(v, z)$ 는 다음과 같이 표현되어 질 수 있다.

$$SF(v, z) = 15.81 + 7.5(\Delta z + 0.474) - 17.5\sqrt{1 + (\Delta z + 0.474)^2} \quad (7)$$

여기서, $\Delta z = z - v$ 는 바크 인덱스(Bark index)로 표현된 임계 대역 간 차이 값이며 청각 자극에너지는 임계 대역 에너지보다 변화가 급격하지 않다.

이런 특성을 이용하여 잡음에 열화 된 신호에서 잡음을 추정할 후 이를 통해 신호를 개선시키는 방법들이 많이 제안되고 있는데, 특히 신호의 자극 에너지를 이용하는 주파수 차감 방법이 그 대표적인 예이다. 주파수 차감 방식은 추정된 잡음의 주파수 스펙트럼을 잡음에 열화 된 주파수 스펙트럼에서 차감하는 방식으로 음질 개선 지각 필터 $H_e(z, i)$ 를 이용하여 식 (8)과 같이 나타낼 수 있다[5].

$$\begin{aligned} \widehat{X}_e(z, i) &= H_e(z, i)Y_i(z, i) \\ &= \left(1 - \frac{\widehat{D}_e(z, i)}{Y_e(z, i)}\right)Y_i(z, i) \\ &\text{for } 0 \leq z \leq Z-1 \end{aligned} \quad (8)$$

여기서 $\widehat{D}_e(z, i)$ 는 추정된 잡음의 청각 자극 에너지로서 추정된 잡음이 잡음에 의해 열화 된 신호보다 크지 않은 경우에만 고려하기로 한다. 식 (8)을 볼 때, 추정 잡음의 오차에 따라서 필터의 성능에 많은 차이가 생김을 알 수 있으며 만약에 인지 가능한 잡음의 에너지를 $H_e(z, i)$ 를 이용하여 $Y(z, i)$ 로부터 제거했다면 식 (9)와 같이 개선된 신호의 자극 에너지는 순수 신호의 에너지와 같게 될 것이다.

$$\widehat{X}_e(z, i) = X_e(z, i) \quad (9)$$

3. 제안된 알고리즘

주파수 차감 방식에서 잡음이 첨가된 신호의 목음 구간에서 세 프레임의 평균값을 잡음으로 추정하였다[2].

$$\widehat{D}(\omega, i) = \text{mean}\{Y(\omega, i)\}, \text{ for } 1 \leq i \leq 3 \quad (10)$$

그러나 식 (10)에서 구한 잡음은 모두 다른 SNR 값을 가지고 있는 각 프레임마다 동일한 값으로 적용되므로 최적화된 성능을 내기는 어렵다. 이에 따라 본 논문에서 사용된 잡음 추정 방식은 처리되는 각 단위마다 독립성분분석 기법을 이용하여 추정함으로써 좀 더 정확하고 가변적인 값을 획득할 수 있게 되고 이에 따라 필터 성능을 최대한으로 극대화 시킬 수 있다[2].

즉 신호와 잡음은 혼합행렬 \mathbf{A} 에 의해 선형으로 섞이는 것으로 가정할 때 두 개의 마이크로폰에 수신된 두 신호에서 잡음만을 추정하도록 시도하게 된다. 이를 위해서 그림 5와 같은 독립성분추출모델을 이용하여 혼합행렬의 역행렬을 추정하게 된다.

분리 행렬 \mathbf{W} 를 결정하는 독립성분분석 알고리즘으로는 여러 가지 알고리즘들이 있으나 본 연구에서는 $u[n]$ 의 비가우시안성

(nongaussianity)이 최대로 되도록 하는 고속독립성분분석(Fast ICA) 알고리즘의 Complex-valued version을 사용하였다[9].

Hyvarinen과 Oja에 의해 제안된 Fast ICA는 kurtosis의 극점을 찾는 것이 비가우시안(non-Gaussian) 독립 요소를 찾아내는 것과 동일하다는 사실에 근거하여 출력신호 $u=wy$ 의 kurtosis를 극대 혹은 극소화 시키는 방법이다[11][12]. 여기서 w 는 분리행렬 W 의 한 열 벡터(row vector)를 나타낸다. 이 방법에서 학습의 기준으로 사용한 오차 함수는 식 (11)과 같으며 F 는 제약조건 $\|w\|^2 = 1$ 을 위한 항이다.

$$J(w) = \langle (wy)^4 \rangle - 3 \|w\|^4 + F(\|w\|^2) \quad (11)$$

$$w^+ = \langle y(w(p)y)^3 \rangle - 3w(p) \quad (12)$$

오차함수로부터 유도된 간단한 학습 방법은 p 가 학습 횟수(iteration)를 나타낸다고 할 때, 식 (12)를 계산한 후 $w(p+1) = w^+ / \|w^+\|$ 에 따라 가중치 벡터를 변경하는 것이다. 위와 같이 w 의 학습에 의해 신호원에 해당하는 신호가 찾아지면, 이 신호를 y 에서 제거한 후 또 다른 신호원에 해당하는 신호를 같은 방식으로 찾아낸다.

먼저 두 개의 마이크를 통해 입력된 신호들을 일정 크기의 프레임으로 분리한 후 푸리에 변환을 적용하여 주파수 영역으로 변환한다. 그 다음으로 각 프레임마다 독립성분분석을 적용하여 잡음신호를 추정한다.

$$U(f, \tau) = \mathbb{W}(f)Y(f, \tau) \quad (13)$$

이어서 독립성분분석의 문제인 amplitude ambiguity를 개선하기 위하여 projection back 방법을 적용한다. 이는 식 (14)와 같이 나타낼 수 있다[10].

$$\hat{U}(f, \tau) = \mathbb{W}^+(f)U(f, \tau) \quad (14)$$

\mathbb{W}^+ 는 W 의 Moore-penrose pseudo inverse matrix를 의미하며 이 작업을 통하여 추정된 노이즈가 마이크에 녹음된 실제 크기와 유사하게 된다.

마지막으로 추정된 인지 가능한 잡음은 자극 에너지로 변환되어 잡음에 열화 된 신호의 자극 에너지, $Y(f, \tau)$ 와 같이 지각 필터에 적용되게 된다. 그 결과로 잡음이 제거되게 된다.

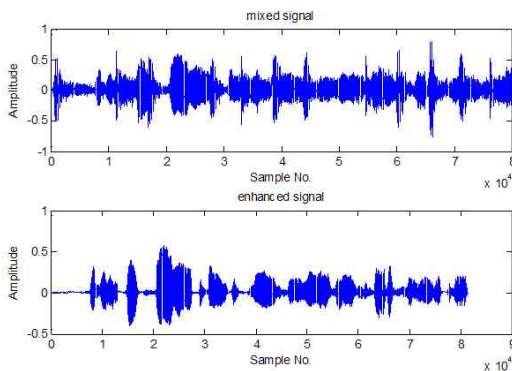


그림 1. 잡음에 열화된 신호(상) & 개선된 신호(하)

제안된 알고리즘의 Block diagram은 다음과 같다.

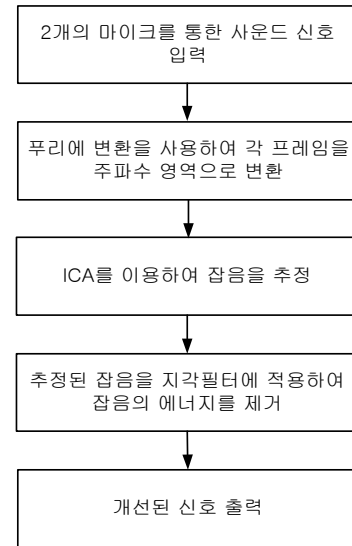


그림 2. 제안된 알고리즘의 Block diagram

4. 실험 및 결과

본 논문은 잡음에 의해 열화 된 신호에서 잡음을 제거하고 음질을 개선하기 위한 알고리즘을 제안하였다.

실험에서 사용된 잡음은 CD에서 추출한 44.1kHz, 16bit 오디오 신호이며 원 신호로는 음성 신호를 사용하였다. 잡음에 의해 열화 된 신호에서 독립성분분석 기법만을 이용하여 잡음을 분리한 신호와 추정된 잡음을 지각필터에 통과시켜 개선시킨 신호를 가지고 SSNR(Segmental Signal-to-noise Ratio)를 비교하였다.

표 1. 세그멘탈 신호 대 잡음 비의 개선량 비교(dB)

Input noise	Method (SSNR)	
	Only ICA	PF
Classic	2.12	4.5901
Jazz	5.341	8.011
Soprano	2.9032	5.2208

표 1에서 볼 수 있듯이 독립성분분석만을 적용시킨 경우 보다 제안된 방법의 필터의 SSNR 결과가 평균 2dB 이상 성능이 개선되었음을 확인 할 수 있다.

5. 결론

본 논문은 일반 생활 환경에서 두 개의 마이크로폰을 통해 수신된 두 개의 신호를 독립성분분석을 이용하여 잡음을 추정한 후, 지각 필터에 적용하여 잡음을 제거함으로써 잡음에 열화 된 신호의 음질을 개선시키는 알고리즘을 제안하였다. 이를 통해 원 신호와 유사한 개선 결과를 얻을 수가 있었다. 또한 기존 방식과의 SSNR 비교를 통해서도 객관적인 성능 개선을 확인할 수 있었다.

다음 연구될 분야는 음성인식이나 오디오 신호 처리 시 입력으로

잡음이 많이 첨가된 신호가 들어올 경우에 잡음을 효과적으로 추정하여 오차를 줄임으로서 지각필터로 최적의 신호를 입력해 주기 위한 연구가 진행될 예정이다.

참 고 문 헌

1. 구교식, 차형태, “스테레오 시스템을 위한 머리전달함수의 개선,” 한국 음향학회 논문지 27권, 4호, 207-214쪽, 2008
2. 서정국, 차형태, “잡음 패턴의 지능적 추정을 통한 음질 개선 알고리즘,” 한국 지능시스템학회 논문지 15권, 2호, 230-235쪽, 2005.
3. S. F. Boll, “Suppression of acoustic noise speech using spectral subtraction”, IEEE Trans. Acoustic, Speech, and Signal processing, Vol. ASSP-27, 1997
4. D. Tsoukalas, J. Mourjopoulos, and G. Kokkinakis, “Perceptual filters for audio signal enhancement,” Journal of Audio Eng. Soc, vol. 45, no. 1/2, pp. 22-35, Jan/Feb. 1997
5. 김현중, 차형태, “잡음 신호의 지각 패턴 제어를 통한 음질 개선 알고리즘 개발에 관한 연구,” 한국음향학회 하계학술대회 논문집 21권, 1(s)호, 199-202쪽, 2002.
6. 독립성분분석 방법을 이용한 홍채 특징 추출
7. Kwanghyuk Bae, Seung-In Noh, and Jaihie Kim, “Iris Feature Extraction Using Independent Component Analysis”, Lecture Notes on Computer Science 2688, pp 838- 844, Jun. 2003.
8. 엄혜영, 한현수, 홍민철, 차형태, “적응 지각 필터를 이용한 오디오 음질 개선 알고리즘,” 한국음향학회지 22권, 8호, 687-693쪽, 2003.
9. A. Hyvarinen and E. Oja, “A fast fixed-point algorithm for Independent Component Analysis,” Neural Computation, vol.9, pp.1483-1492, 1997
10. S. Ikeda and N. Murata, “A method of ICA in the frequency domain,” in Proc. Int. Workshop Ind. Compon. Anal. Blind Signal Separation, 1999, pp.365-371