



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0135757  
(43) 공개일자 2017년12월08일

- (51) 국제특허분류(Int. Cl.)  
G10L 21/02 (2006.01) G10L 25/03 (2013.01)  
H04R 1/10 (2006.01)
- (52) CPC특허분류  
G10L 21/02 (2013.01)  
G10L 25/03 (2013.01)
- (21) 출원번호 10-2017-0067119
- (22) 출원일자 2017년05월30일  
심사청구일자 없음
- (30) 우선권주장  
16171986.9 2016년05월30일  
유럽특허청(EPO)(EP)

- (71) 출원인  
오티콘 에이/에스  
덴마크 디케이-2765 스모럼 콩게바겐 9
- (72) 발명자  
젠센 제스퍼  
덴마크 디케이-2765 스모럼 콩게바겐 9 오티콘 에이/에스 씨/오  
크젠크 올리크  
덴마크 디케이-1850 프레데릭스베르그 씨 쥐아이 콘게베지 163 2번  
(뒷면에 계속)
- (74) 대리인  
박장원

전체 청구항 수 : 총 24 항

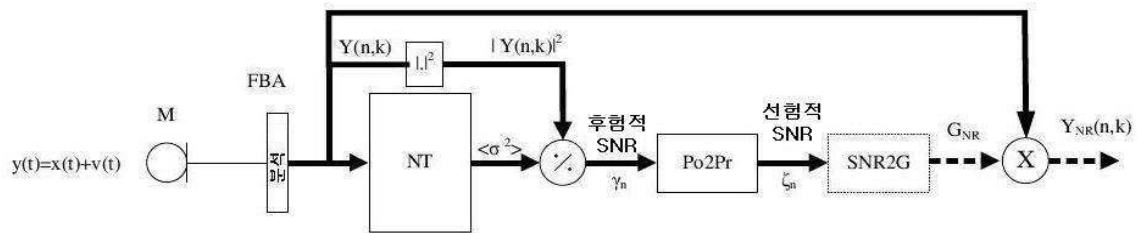
(54) 발명의 명칭 오디오 프로세싱 디바이스 및 사운드 신호의 신호-대-잡음비를 추정하는 방법

**(57) 요약**

본 출원은, a) 타겟 사운드 소스(TS)로부터의 타겟 스피치 신호 컴포넌트들(S(k,n)) 및 상기 타겟 사운드 소스 이외의 소스들로부터의 노이즈 신호 컴포넌트들(N(k,n))로 구성되는 시간 변화 사운드 신호를 나타내는 전기 입력 신호의 시간-주파수 표현(Y(k,n))을 제공하는 적어도 하나의 입력 유닛 - 여기서, k 및 n은 각각, 주파수 대역

(뒷면에 계속)

**대표도** - 도1a



및 시간 프레임 색인이다, b) 노이즈 검출 및/또는 감소 시스템을 포함하는 오디오 프로세싱 디바이스에 관한 것으로서, 상기 노이즈 검출 및/또는 감소 시스템은 b1) 상기 전기 입력 신호의 사후 신호 대 잡음비 추정 ( $\gamma(k,n)$ )을 결정하고, 그리고 b2) 재귀적 결정 지시 알고리즘에 기초하여 상기 사후 신호 대 잡음비 추정 ( $\gamma(k,n)$ )으로부터 상기 전기 입력 신호의 사전 타겟 신호 대 잡음 신호비 추정 ( $\zeta(k,n)$ )을 결정하기 위해 구성된다. n번째 시간 프레임에 대한 사전 타겟 신호 대 잡음 신호비 추정 ( $\zeta(k,n)$ )은, (n-1)번째 시간 프레임에 대한 사전 타겟 신호 대 잡음 신호비 추정 ( $\zeta(k,n-1)$ ) 및 n번째 시간 프레임에 대한 사후 신호 대 잡음비 추정 ( $\gamma(k,n)$ )으로부터 결정된다. 본 출원은 또한 사전 신호 대 잡음비를 추정하는 방법에 관한 것이다. 이로써, 개선된 잡음 감소는 제공될 수 있다. 본 발명은, 예를 들어, 보청기, 헤드셋, 이어폰, 능동적인 귀 보호 시스템, 핸드프리 전화 시스템들, 휴대 전화들 등에 대해 사용될 수 있다.

(52) CPC특허분류

*HO4R 1/1016* (2013.01)

(72) 발명자

**베텔센 안드레아스 더랜더**

덴마크 디케이-2765 스모럼 콩게바겐 9 오티콘 에이/에스 씨/오

**페데르센 미첼 시스킨드**

덴마크 디케이-2765 스모럼 콩게바겐 9 오티콘 에이/에스 씨/오

## 명세서

### 청구범위

#### 청구항 1

오디오 프로세싱 디바이스, 예를 들어, 보청기(hearing aid)에 있어서,

타겟 사운드 소스(target sound source)(TS)로부터의 타겟 스피치 신호 컴포넌트들  $S(k,n)$  및 상기 타겟 사운드 소스 이외의 소스들로부터의 노이즈 신호 컴포넌트들  $N(k,n)$ 로 구성된 시간 변화 사운드 신호(time variant sound signal)를 나타내는 전기 입력 신호의 시간-주파수 표현  $Y(k,n)$ 을 제공하기 위한 적어도 하나의 입력 유닛 -  $k$  및  $n$ 은, 각각, 주파수 대역 인덱스 및 시간 프레임 인덱스이며 - 과; 그리고

노이즈 감소 시스템을 포함하고,

상기 노이즈 감소 시스템은:

상기 전기 입력 신호의 후험적(posteriori) 신호 대 잡음비 추정치  $\gamma(k,n)$ 를 결정하고,

재귀적 알고리즘에 기초하여 상기 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 로부터 상기 전기적 입력 신호의 선형적(priori) 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하며, 그리고

$n-1$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n-1)$  및  $n$  번째 시간 프레임에 대한 상기 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 로부터 상기  $n$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하도록 되어 있는 것을 특징으로 하는

오디오 프로세싱 디바이스.

#### 청구항 2

제1항에 있어서,

상기 재귀적 알고리즘은, 적응 시간 상수(adaptive time constant)를 갖는 저역 통과 필터를 구현하도록 되어 있는 것을 특징으로 하는

오디오 프로세싱 디바이스.

#### 청구항 3

제1항 또는 제2항에 있어서,

상기 저역 통과 필터의 적응 시간 상수는, 제1 (후험적) 신호 대 잡음비 추정치 및/또는 제2 (선형적) 신호 대 잡음비 추정치에 따라 결정되는 것을 특징으로 하는

오디오 프로세싱 디바이스.

#### 청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

특정 주파수 인덱스  $k$ 에 대한 상기 저역 통과 필터의 적응 시간 상수는, 오직 상기 주파수 인덱스  $k$ 에 대응하는 제1 (후험적) 신호 대 잡음비 추정치 및/또는 제2 (선형적) 신호 대 잡음비 추정치에 따라 결정되는 것을 특징으로 하는

오디오 프로세싱 디바이스.

#### 청구항 5

제1항 내지 제3항 중 어느 한 항에 있어서,

특정 주파수 인덱스  $k$ 에 대한 상기 저역 통과 필터의 적응 시간 상수 또는 저역 통과 차단 주파수는, 미리 정의된 방식 또는 적응형 방식(adaptive scheme)에 따라 다수의 주파수 인덱스들  $k'$ 에 대응하는 제1 (후험적) 신호 대 잡음비 추정치 및/또는 제2 (선형적) 신호 대 잡음비 추정치에 따라 결정되는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 6**

제1항 내지 제5항 중 어느 한 항에 있어서,

특정 주파수 인덱스  $k$ 에 대한 상기 저역 통과 필터의 적응 시간 상수는, 하나 이상의 검출기들로부터의 입력들에 따라 결정되는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 7**

제1항 내지 제6항 중 어느 한 항에 있어서,

상기 노이즈 감소 시스템은,  $\gamma(k,n)$ 이 1보다 크거나 같다는 가정하에서,  $n$ 번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하도록 되어 있는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 8**

제1항 내지 제7항 중 어느 한 항에 있어서,

상기 노이즈 감소 시스템은, 상기  $n-1$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n-1)$ 로부터 그리고 상기  $n$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 의 최대 우도 SNR 추정자(maximum likelihood SNR estimator)  $\zeta^{ml}(k,n)$ 로부터 상기  $n$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하도록 되어 있는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 9**

제8항에 있어서,

상기 노이즈 감소 시스템은  $\text{MAX}\{\zeta_{\min}^{ml}(k,n); \gamma(k,m)-1\}$ 로서 상기 최대 우도 SNR 추정자  $\zeta^{ml}(k,n)$ 를 결정하도록 구성되고, MAX는 최대 연산자이고,  $\zeta_{\min}^{ml}(k,n)$ 는 최대 우도 SNR 추정자  $\zeta^{ml}(k,n)$ 의 최소값인 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 10**

제1항 내지 제9항 중 어느 한 항에 있어서,

상기 노이즈 감소 시스템은, 상기 후험적 신호 대 잡음비 추정치  $\gamma$ 의 비선형 평활화에 의해 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta$ 를 결정하거나, 또는 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta$ 로부터 도출된 파라미터를 결정하는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 11**

제10항에 있어서,

상기 노이즈 감소 시스템은, 높은 SNR 조건들에서 보다 낮은 SNR 조건들에서 평활화를 더 가능하게 하는 SNR-의존 평활화를 제공하는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 12**

제1항 내지 제11항 중 어느 한 항에 있어서,

상기 노이즈 감소 시스템은, 낮은 SNR 조건들에 대해  $\xi_n^{ML}$  에 비해 음(negative)의 바이어스를 제공하도록 되어 있는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 13**

제12항에 있어서,

상기 노이즈 감소 시스템은, 저-대-고(low-to-high) SNR 조건 및 고-대-저(high-to-low) SNR 조건으로부터 설정가능한 변경(configurable change)을 가능하게 하는, 재귀적 바이어스(recursive bias)를 제공하도록 되어 있는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 14**

제8항 내지 제13항 중 어느 한 항에 있어서,

상기 노이즈 감소 시스템은,

상기 n-1 번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k, n-1)$ 로부터 그리고 재귀적 알고리즘:  $s_n - s_{n-1} = (s_n^{ML} + \rho(s_{n-1}) - s_{n-1})\lambda(s_{n-1})$  에 따른 상기 n 번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k, n)$ 의 최대 우도 SNR 추정자  $\zeta^{ML}(k, n)$ 로부터 상기 n 번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k, n)$ 를 결정하도록 되어 있고,

$\rho(s_{n-1})$ 은 n-1 번째 시간 프레임의 바이어스 함수 또는 파라미터를 나타내고 그리고  $\lambda(s_{n-1})$ 은 n-1 번째 시간 프레임의 평활화 함수 또는 파라미터를 나타내는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 15**

제1항 내지 제14항 중 어느 한 항에 있어서,

상기 전기 입력 신호의 상기 시간-주파수 표현  $Y(k, n)$ 을 제공하는 분석 필터 बैं크를 포함하는 필터 बैं크를 포함하는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 16**

제15항에 있어서,

상기 분석 필터 बैं크는 오버샘플링(oversampling)되도록 제공되는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 17**

제16항에 있어서,

상기 n번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하는 알고리즘의 재귀적 루프는 고차 지연 요소, 예를 들어, 원형 버퍼(circular buffer)를 포함하는 것을 특징으로 하는 오디오 프로세싱 디바이스.

**청구항 18**

제16항 또는 제17항에 있어서,

상기 잡음 감소 시스템은, 상기 분석 필터 뱅크의 오버샘플링을 보상하는 것을 포함하는 상기 n번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하기 위한 알고리즘을 적응시키도록 되어 있는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 19**

제1항 내지 제18항 중 어느 한 항에 있어서,

보청기, 헤드셋, 이어폰, 귀 보호 디바이스 또는 이들의 조합과 같은 청각 디바이스(hearing device)를 포함하는 것을 특징으로 하는

오디오 프로세싱 디바이스.

**청구항 20**

타겟 스피치 컴포넌트들 및 노이즈 컴포넌트들로 구성되는 시간 변화 사운드 신호를 나타내는 전기 입력 신호의 시간-주파수 표현  $Y(k,n)$ 의 선형적 신호 대 잡음비  $\zeta(k,n)$ 를 추정하는 방법 - k 및 n은, 각각, 주파수 대역 인덱스 및 시간 프레임 인덱스이며 - 으로서,

상기 전기 입력 신호  $Y(k,n)$ 의 후행적 신호 대 잡음비 추정치  $\gamma(k,n)$ 를 결정하는 단계와,

재귀적 알고리즘에 기초하여 상기 후행적 신호 대 잡음비 추정치  $\gamma(k,n)$ 로부터 상기 전기적 입력 신호의 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하는 단계와, 그리고

n-1 번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n-1)$  및 n 번째 시간 프레임에 대한 상기 후행적 신호 대 잡음비 추정치  $\gamma(k,n)$ 로부터 상기 n 번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하는 단계를 포함하는 것을 특징으로 하는

선형적 신호 대 잡음비를 추정하는 방법.

**청구항 21**

제20항에 있어서,

상기 재귀적 알고리즘은, 적응 시간 상수를 갖는 저역 통과 필터를 구현하도록 되어 있는 것을 특징으로 하는 선형적 신호 대 잡음비를 추정하는 방법.

**청구항 22**

제20항 또는 제21항에 있어서,

상기 타겟 스피치 컴포넌트들의 크기들  $\hat{A}(k,n)$ 의 추정치들은 이득 함수 G가 곱해진 상기 전기 입력 신호  $Y(k,n)$ 로부터 결정되고, 상기 이득 함수 G는 상기 후행적 신호 대 잡음비 추정치  $\gamma(k,n)$  및 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 의 함수인 것을 특징으로 하는

선형적 신호 대 잡음비를 추정하는 방법.

**청구항 23**

프로세서 및 상기 프로세서로 하여금 제20항 내지 제22항 중 어느 한 항의 방법을 수행하게 하는 프로그램 코드 수단을 포함하는 데이터 프로세싱 시스템.

**청구항 24**

프로그램이 컴퓨터에 의해 실행될 때, 상기 컴퓨터로 하여금 제20항 내지 제22항 중 어느 한 항의 방법을 수행하도록 하는 명령어들을 포함하는 컴퓨터 프로그램.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 오디오 프로세싱 디바이스, 예를 들어, 보청기, 및 소리를 나타내는 전기 입력 신호의 신호 대 잡음비를 추정하는 방법에 관한 것이다. 본 발명은, 특히, 후험적 신호 대 잡음비 추정치의 (적응형 저역 차단 주파수를 갖는 저역 통과 필터링로 구현되는) 비-선형 평활화에 의한 선형적 신호 대 잡음비 추정치를 획득하기 위한 방식에 관한 것이다.

**발명의 내용**

[0002] 본 발명의 문맥에서, '후험적 신호 대 잡음비',  $SNR_{post}$ 는, 예를 들어, 노이즈 신호의 전력과 같은 하나 이상의 마이크로폰들에 의해 픽업되는 관측된 (이용가능한) 노이즈 신호(타겟 신호  $S$ 와 노이즈  $N$ 의 합,  $Y(t)=S(t)+N(t)$ )와, 시간  $t$ 에서의 특정 지점에서, 노이즈 신호의 전력과 같은 노이즈의 추정치 ( $\hat{N}(t)$ )인  $N(t)$  사이의 비율을 의미하는 것으로 얻어진다. 즉,  $SNR_{post}(t) = Y(t)/\hat{N}(t)$ , 또는  $SNR_{post}(t) = Y(t)^2/\hat{N}(t)^2$ 이다. '후험적 신호 대 잡음비',  $SNR_{post}$ 는, 예를 들어, 각 주파수 대역(인덱스  $k$ ) 및 시간 프레임(인덱스  $n$ )에 대한 값으로 시간-주파수 도메인으로 정의될 수 있다. 즉,  $SNR_{post} = SNR_{post}(k,n)$ 이고, 예를 들어,  $SNR_{post}(k,n) = |Y(k,n)|^2/|\hat{N}(k,n)|^2$ 이다. '후험적' 신호 대 잡음비의 발생의 예들은, 단일 마이크로폰 및 다중 마이크로폰, 각각에 대해 도 1a 및 도 1b에서 도시된다.

[0003] 본 발명의 문맥에서, '선형적 신호 대 잡음비',  $SNR_{prio}$ 는, 타겟 신호 크기  $S(t)$ (또는 타겟 신호 제곱  $S(t)^2$ )와 노이즈 신호 크기  $N(t)$ (또는 노이즈 신호 제곱  $N(t)^2$ )의 각각의 비율, 예를 들어, 시간  $t$ 에서 특정 지점에서 이러한 신호들의 추정치들 사이의 비율을 의미하는 것으로 얻어진다. 예를 들어,  $SNR_{prio} = SNR_{prio}(t) = \hat{S}(t)^2/\hat{N}(t)^2$ , 또는  $SNR_{prio} = SNR_{prio}(k,n)$ , 즉,  $SNR_{prio}(k,n) = |\hat{S}(k,n)|^2/|\hat{N}(k,n)|^2$ 이다.

[0005] 오디오 프로세싱 디바이스, 예를 들어, 보청기와 같은 청각 디바이스

[0006] 본 출원의 제1 양상에서, 오디오 프로세싱 디바이스가 제공된다. 오디오 프로세싱 디바이스, 예를 들어, 보청기 (hearing aid)는,

[0007] 타겟 사운드 소스(target sound source)(TS)로부터의 타겟 스피치 신호 컴포넌트들  $S(k,n)$  및 노이즈 신호 컴포넌트들  $N(k,n)$ 로 구성된 시간 변화 사운드 신호(time variant sound signal)를 나타내는 전기 입력 신호의 시간-주파수 표현  $Y(k,n)$ 을 제공하기 위한 적어도 하나의 입력 유닛 -  $k$  및  $n$ 은, 각각, 주파수 대역 인덱스 및 시간 프레임 인덱스이며 - 과; 그리고

[0008] 노이즈 감소 시스템을 포함하고,

[0009] 상기 노이즈 감소 시스템은:

- [0010] 상기 전기 입력 신호의 제1의 후험적(posteriori) 신호 대 잡음비 추정치  $\gamma(k,n)$ 를 결정하고,
- [0011] 재귀적 알고리즘에 기초하여 상기 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 로부터 상기 전기적 입력 신호의 제2의 선형적(priori) 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하며, 그리고
- [0012]  $n-1$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n-1)$  및  $n$  번째 시간 프레임에 대한 상기 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 로부터 상기  $n$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하도록 되어있다.
- [0013] 일 실시예에서, 재귀적 알고리즘은 단위 DC-이득 및 적응 시간 상수(또는 저역 차단 주파수)를 갖는 1차 IIR 저역 통과 필터를 구현한다.
- [0014] 본 출원의 제2 양상에서, 오디오 프로세싱 디바이스가 제공된다. 오디오 프로세싱 디바이스, 예를 들어, 보청기는,
- [0015] 타겟 사운드 소스(TS)로부터의 타겟 스피치 신호 컴포넌트들  $S(k,n)$  및 노이즈 신호 컴포넌트들  $N(k,n)$ 로 구성된 시간 변화 사운드 신호를 나타내는 전기 입력 신호의 시간-주파수 표현  $Y(k,n)$ 을 제공하기 위한 적어도 하나의 입력 유닛 -  $k$  및  $n$ 은, 각각, 주파수 대역 인덱스 및 시간 프레임 인덱스이며 - 과; 그리고
- [0016] 노이즈 감소 시스템을 포함하고,
- [0017] 상기 노이즈 감소 시스템은, 각각의 주파수 대역에 대해,
- [0018] 상기 전기 입력 신호의 제1의 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 를 결정하고,
- [0019] 재귀적 알고리즘에 기초하여 상기 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 로부터 상기 전기적 입력 신호의 제2의 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하도록 구성되며, 상기 재귀적 알고리즘은 적응 시간 상수 또는 저역 차단 주파수를 갖는 저역 통과 필터를 구현한다.
- [0020] 즉, 제2의 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 는 제1의 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 를 저역 통과 필터링함으로써 결정된다.
- [0021] 일 실시예에서, 저역 통과 필터의 적응 시간 상수 또는 저역 통과 차단 주파수는, 제1의 후험적 신호 대 잡음비 추정치 및/또는 제2의 선형적 신호 대 잡음비 추정치에 따라 결정된다.
- [0022] 일 실시예에서, 주어진 주파수 인덱스  $k$ (또한, 주파수 채널  $k$ 라고도 함)에 대한 저역 통과 필터의 적응 시간 상수 또는 저역 통과 차단 주파수는, 단지 상기 주파수 인덱스  $k$ 에만 대응하는 제1의 후험적 신호 대 잡음비 추정치 및/또는 제2의 선형적 신호 대 잡음비 추정치에 따라 결정된다.
- [0023] 일 실시예에서, 주어진 주파수 인덱스  $k$ (또한, 주파수 채널  $k$ 라고도 함)에 대한 저역 통과 필터의 적응 시간 상수 또는 저역 통과 차단 주파수는, 예를 들어, 미리정의된 (또는 적응성) 방식에 따라, 예를 들어, 적어도 이웃하는 주파수 인덱스들  $k-1$ ,  $k$ ,  $k+1$ 를 포함하는 다수의 주파수 인덱스들  $k'$ 에 대응하는 제1의 후험적 신호 대 잡음비 추정치 및/또는 제2의 선형적 신호 대 잡음비 추정치에 따라 결정된다.
- [0024] 일 실시예에서, 주어진 주파수 인덱스  $k$ (또한, 주파수 채널  $k$ 라고도 함)에 대한 저역 통과 필터의 적응 시간 상수 또는 저역 통과 차단 주파수는, 하나 이상의 검출기들(예를 들어, 온셋 표시기들, 바람 소리 또는 음성 검출기들 등)으로부터의 입력들에 따라 결정된다.
- [0025] 일 실시예에서, 저역 통과 필터는 1차 IIR 저역 통과 필터이다. 일 실시예에서, 1차 IIR 저역 통과 필터는 단위 DC-이득을 갖는다.
- [0026] 일 실시예에서, 주어진 시간 인스턴스  $n$ 에서의 저역 통과 필터의 적응 시간 상수 또는 저역 통과 차단 주파수는, 상기 시간 인스턴스  $n$ 에서의 제2의 선형적 타겟 신호 대 잡음비 추정치의 제1 최대 우도 추정치 및/또는 이전의 시간 인스턴스  $n-1$ 의 제2의 선형적 타겟 신호 대 잡음비 추정치의 추정치에 따라 결정된다.
- [0027] 따라서, 개선된 노이즈 감소가 제공될 수 있다.

- [0028] 노이즈 신호 컴포넌트들  $N(k,n)$ 은, 예를 들어, 타겟 사운드 소스 TS보다는 하나 이상의 다른 소스들  $NS_i$  ( $i=1, \dots, N_s$ )로부터 발생할 수 있다. 일 실시예에서, 노이즈 신호 컴포넌트들  $N(k,n)$ 은 타겟 신호(예를 들어, 해당 타겟 신호 컴포넌트의 우세한 피크보다 50ms 이상 늦게 사용자에게 도달하는 타겟 신호 컴포넌트들)로부터의 늦은 반향들을 포함한다.
- [0029] 즉,  $\zeta(k,n) = F(\zeta(k,n-1), \gamma(k,n))$ 이다. 선형적 SNR의 결정에서 후협적 SNR(SNR = 신호 대 잡음비)에 대한 가장 최근의 프레임 전력을 사용하는 것은, 예를 들어, SNR에 대한 큰 증가들이 일반적으로 짧은 시간에 발생하는 스피치 온셋들에서의 SNR 추정치에 대해 유용하다.
- [0030] 일 실시예에서,  $\gamma(k,n)$ 이 1보다 크거나 같다는 가정하에서,  $n$ 번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하도록 되어 있다. 일 실시예에서, 전기 입력 신호  $Y(k,n)$ 의 후협적 신호 대 잡음비 추정치  $\gamma(k,n)$ 는, 전기 입력 신호의 현재 값  $Y(k,n)$ 의 신호 전력 스펙트럼 밀도  $|Y(k,n)|^2$ 와 전기 입력 신호  $Y(k,n)$ 의 현재 노이즈 전력 스펙트럼 밀도의 추정치  $\langle \sigma^2 \rangle$  사이의 비율로서 정의된다. 즉,  $\gamma(k,n) = |Y(k,n)|^2 / \langle \sigma^2 \rangle$ 이다.
- [0031] 일 실시예에서, 상기 노이즈 감소 시스템은, 상기  $n-1$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n-1)$ 로부터 그리고 상기  $n$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 의 최대 우도 SNR 추정자  $\zeta^{ML}(k,n)$ 로부터 상기  $n$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하도록 되어 있다.
- [0032] 일 실시예에서, 상기 노이즈 감소 시스템은  $\text{MAX}\{\zeta_{\min}^{ML}(k,n); \gamma(k,n)-1\}$ 로서 상기 최대 우도 SNR 추정자  $\zeta^{ML}(k,n)$ 를 결정하도록 구성되고, MAX는 최대 연산자이고,  $\zeta_{\min}^{ML}(k,n)$ 는 최대 우도 SNR 추정자  $\zeta^{ML}(k,n)$ 의 최소값이다. 일 실시예에서, 최대 우도 SNR 추정자  $\zeta^{ML}(k,n)$ 의 최소값은, 예를 들어, 주파수 대역 인덱스에 의존할 수 있다. 일 실시예에서, 최소값  $\zeta_{\min}^{ML}(k,n)$ 은 독립적이다. 일 실시예에서,  $\zeta_{\min}^{ML}(k,n)$ 는 1과 동일하게 취해진다(즉, 로그 스케일 상에서 0dB이다). 이는, 예를 들어, 타겟 신호 컴포넌트들  $S(k,m)$ 이 무시할 수 있는 경우이다, 즉, 입력 신호  $Y(k,m)$ 에 노이즈 컴포넌트들  $N(k,m)$ 만이 존재하는 경우이다.
- [0033] 일 실시예에서, 상기 노이즈 감소 시스템은, 상기 후협적 신호 대 잡음비 추정치  $\gamma$ 의 비선형 평활화에 의해 선형적 타겟 신호 대 잡음비 추정치  $\zeta$ 를 결정하거나, 또는 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta$ 로부터 도출된 파라미터를 결정하도록 되어 있다. 상기 도출된 파라미터는, 예를 들어, 최대 우도 SNR 추정자  $\zeta^{ML}(k,n)$ 일 수 있다. 비선형 평활화는, 예를 들어, 적응형 저역 차단 주파수를 갖는 저역 통과 필터링에 의해, 예를 들어, 단위 DC-이득 및 적응형 시간 상수를 갖는 1차 IIR 저역 통과 필터에 의해 구현될 수 있다.
- [0034] 일 실시예에서, 상기 노이즈 감소 시스템은, 높은 SNR 조건들에서 보다 낮은 SNR 조건들에서 평활화를 더 가능하게 하는 SNR-의존 평활화를 제공하도록 되어 있다. 이는, 음악적 소음을 줄이는 이점을 가질 수 있다. 용어 '낮은 SNR 조건들' 및 '높은 SNR 조건들'은, 제1 및 제2 조건들을 나타내도록 의도되고, 참 SNR은 제2 조건들에서 보다 제1 조건들 하에서 더 낮다. 일 실시예에서, '낮은 SNR 조건들' 및 '높은 SNR 조건들'은, 각각, 0dB 아래 및 위인 것을 의미한다. 바람직하게, 평활화를 제어하는 시간 상수의 의존성은, SNR에 따른 점진적인 변화를 나타낸다. 일 실시예에서, 평활화에 관련된 시간 상수(들)은 SNR이 낮을수록 높다. '낮은 SNR 조건들'에서, SNR 추정치는, 일반적으로 '높은 SNR 조건들'에서 보다 상대적으로 열악하다(그리고 SNR이 낮을수록 신뢰도가 낮아져서, 구동기는 더 평활하게 한다).
- [0035] 일 실시예에서, 상기 노이즈 감소 시스템은, 낮은 SNR 조건들에 대한  $\xi_n^{ML}$ 에 비해 음(negative)의 바이어스를 제공하도록 되어 있다. 이러한 것은, 노이즈 전용 기간들에서 음악적인 노이즈의 가청성을 줄이는 이점을 가질 수 있다. 용어 "바이어스(bias)"는, 현재의 문맥에서, 최대 우도 SNR 추정자  $\zeta^{ML}(k,n)$ 의 기대값  $E(\xi_n^{ML})$ 과 선형

적 신호 대 잡음비  $\zeta(k,n)$ 의 기대값  $E(\zeta_n)$  사이의 차이를 반영하는데 사용된다. 즉, '낮은 SNR 조건들'에 대해 (예를 들어, 참 SNR < 0dB에 대해),  $E(\zeta_n^{ML}) - E(\zeta_n) < 0$ 이다(예를 들어, 도 3에 반영됨).

[0036] 일 실시예에서, 상기 노이즈 감소 시스템은, 낮은-대-높은(low-to-high) SNR 조건 및 높은-대-낮은(high-to-low) SNR 조건으로부터 설정가능한 변경(configurable change)을 가능하게 하는, 재귀적 바이어스(recursive bias)를 제공하도록 되어 있다.

[0037] n번째 시간 프레임에 대한 선형적 신호 대 잡음비의 대수 표현은,  $s_n = s(k,n) = 10 \log(\zeta(k,n))$ 로서 표현될 수 있고, 그리고 그에 대응하여, n번째 시간 프레임에 대한 최대 우도 SNR 추정자:  $s_n^{ML} = s^{ML}(k,n) = 10 \log(\zeta^{ML}(k,n))$ 이다.

[0038] 일 실시예에서, 상기 노이즈 감소 시스템은,

[0039] 상기 n-1 번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n-1)$ 로부터 그리고 재귀적 알고리즘:  $s_n - s_{n-1} = (s_n^{ML} + \rho(s_{n-1}) - s_{n-1})\lambda(s_{n-1})$  에 따른 상기 n 번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 의 최대 우도 SNR 추정자  $\zeta^{ML}(k,n)$ 로부터 상기 n 번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하도록 되어 있고,

[0040]  $\rho(s_{n-1})$ 은 바이어스 함수 또는 파라미터를 나타내고 그리고  $\lambda(s_{n-1})$ 은 n-1 번째 시간 프레임에 대한 n-1 번째 시간 프레임의 평활화 함수 또는 파라미터를 나타낸다.

[0041] 일 실시예에서,  $\rho(s_{n-1})$ 는

[0042]  $10 \log_{10} \frac{\xi}{\xi_{n-1}}$  의 값과 동일한 것으로 선택되고,

[0043]  $\xi$  는  $10 \log_{10} \Psi\left(\xi_{n-1}, \frac{\xi}{\xi_{n-1}}\right) = 0 \text{ dB}$  이고,

[0044]  $\Psi\left(\xi_{n-1}, \frac{\xi}{\xi_{n-1}}\right)$  는 식(8)에서 정의된 비선형 함수이다.

[0045] 일 실시예에서, 평활화 함수  $\lambda(s_{n-1})$ 는

[0046] 0dB 크로싱의 위치에서(즉,  $s_n^{ML} - s_n = \rho(s_{n-1})$ 일 때),

[0047] 함수  $10 \log_{10} \Psi\left(\xi_{n-1}, \frac{\xi}{\xi_{n-1}}\right) = 0 \text{ dB}$  의 기울기( $s_n^{ML}$ 에 관한)와 동일한 것으로 선택된다(도 3 참조).

[0048] 일 실시예에서, 오디오 프로세싱 디바이스는, 상기 전기 입력 신호의 상기 시간-주파수 표현  $Y(k,n)$ 을 제공하는 분석 필터 बैं크를 포함하는 필터 बैं크를 포함한다. 일 실시예에서, 다수의 주파수 서브-대역 신호들  $Y(k,n)$ ,  $k=1, 2, \dots, K$ 로서 이용가능하다. 일 실시예에서, 선형적 신호 대 잡음비 추정치  $\zeta(k,n)$ 는, 이웃하는 주파수 서브-대역 신호에서 후행적 신호 대 잡음비 추정치(예를 들어,  $\gamma(k-1,n)$  및/또는  $\gamma(k+1,n)$ )에 의존한다.

[0049] 일 실시예에서, 상기 오디오 프로세싱 디바이스는, 상기 분석 필터가 오버샘플링(oversampling)되도록 제공된다. 일 실시예에서, 상기 오디오 프로세싱 디바이스는, 분석 필터 बैं크가 DFT-변조된 분석 필터 बैं크이도록 제공된다.

[0050] 일 실시예에서, 상기 n번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하는 알고리즘의 재귀적 루프는 고차 지연 요소, 예를 들어, 원형 버퍼(circular buffer)를 포함한다. 일 실시예에서, 상기 고차 지연 요소는 상기 분석 필터 बैं크의 오버샘플링을 보상하도록 되어 있다.

- [0051] 일 실시예에서, 상기 잡음 감소 시스템은, 상기 분석 필터 뱅크의 오버샘플링을 보상하도록 상기 n번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\hat{\zeta}(k,n)$ 를 결정하기 위한 알고리즘을 적용시키도록 되어 있다. 일 실시예에서, 상기 알고리즘은 평활화 파라미터( $\hat{\lambda}$ ) 및/또는 바이어스 파라미터( $\hat{\rho}$ )를 포함한다.
- [0052] 일 실시예에서, 2개의 함수  $\hat{\lambda}$  및  $\hat{\rho}$ 는, 추정된 SNR의 재귀 함수로서 평활화의 양 및 SNR 바이어스의 양을 제어한다.
- [0053] 일 실시예에서, 평활화 파라미터( $\hat{\lambda}$ ) 및/또는 바이어스 파라미터( $\hat{\rho}$ )는 샘플링 속도를 보상하도록 적용된다. 예를 들어, 도 5 참조. 일 실시예에서, 상이한 오버샘플링 속도들은 파라미터  $\alpha$ 를 적용시킴으로써 보상된다. 예를 들어, 도 8 참조.
- [0054] 일 실시예에서, 오디오 프로세싱 디바이스는, 보청기, 헤드셋, 이어폰, 귀 보호 디바이스 또는 이들의 조합과 같은 청각 디바이스(hearing device)를 포함한다.
- [0055] 일 실시예에서, 오디오 프로세싱 디바이스는, 예를 들어, 사용자의 청력 손상을 보완하고 그리고/또는 도전적인 음향 환경을 보완하기 위해, 주파수 종속 이득 및/또는 레벨 종속 압축 및/또는 (주파수 압축의 유무에 관계없이) 하나 이상의 주파수 범위들에서 하나 이상의 다른 주파수 범위들로 전이를 제공하도록 구성된다. 일 실시예에서, 상기 오디오 프로세싱 디바이스는, 입력 신호들을 개선하고 그리고 프로세싱된 출력 신호를 제공하기 위한 신호 프로세싱 유닛을 포함한다.
- [0056] 일 실시예에서, 오디오 프로세싱 디바이스는, 프로세싱된 전기 신호에 기초한 음향 신호로서 사용자에게 의해 인식된 자극을 제공하기 위한 출력 유닛을 포함한다. 일 실시예에서, 출력 유닛은 인공 와우의 다수의 전극들 또는 골전도 청각 디바이스의 진동기를 포함한다. 일 실시예에서, 출력 유닛은 출력 트랜스듀서를 포함한다. 일 실시예에서, 출력 트랜스듀서는, 음향 신호로서 자극을 사용자에게 제공하기 위한 수신기(라우드스피커)를 포함한다. 일 실시예에서, 출력 트랜스듀서는 사용자에게 두개골 뼈의 기계적 진동으로서 자극을 제공하기 위한 진동기(예를 들어, 뼈 부착형 또는 뼈 고정식 청력 디바이스)를 포함한다.
- [0057] 일 실시예에서, 오디오 프로세싱 디바이스는 사운드를 나타내는 전기 입력 신호를 제공하기 위한 입력 유닛을 포함한다. 일 실시예에서, 입력 유닛은, 입력 사운드를 전기적 입력 신호로 변환하기 위한 입력 트랜스듀서, 예를 들어, 마이크론을 포함한다. 일 실시예에서, 입력 유닛은, 사운드를 포함하는 무선 신호를 수신하고 그리고 상기 사운드를 나타내는 전기 입력 신호를 제공하기 위한 무선 수신기를 포함한다.
- [0058] 일 실시예에서, 상기 오디오 프로세싱 디바이스는, 휴대용 디바이스, 예를 들어, 로컬 에너지 소스, 예를 들어, 배터리(예를 들어, 재충전 배터리)를 포함하는 디바이스이다.
- [0059] 일 실시예에서, 바이노럴 보청기 시스템의 일부를 형성하는 주어진 보청기의 선형적 SNR 추정치는, 상기 바이노럴 보청기 시스템의 양 보청기들로부터의 후협적 SNR 추정치들에 기초한다. 일 실시예에서, 바이노럴 보청기 시스템의 일부를 형성하는 주어진 보청기의 선형적 SNR 추정치는, 상기 주어진 보청기의 후협적 SNR 추정치 및 상기 바이노럴 보청기 시스템의 다른 보청기의 선형적 SNR 추정치에 기초한다.
- [0060] 일 실시예에서, 오디오 프로세싱 디바이스는, 입력 트랜스듀서(마이크로폰 시스템 및/또는 직접 전기 입력(예를 들어, 무선 수신기))와 출력 트랜스듀서 사이의 순방향 (또는 신호) 경로를 포함한다. 일 실시예에서, 신호 프로세싱 유닛은 순방향 경로에 위치된다. 일 실시예에서, 신호 프로세싱 유닛은 사용자의 특정 요구에 따라 주파수 종속 이득을 제공하도록 구성된다. 일 실시예에서, 오디오 프로세싱 디바이스는, 입력 신호를 분석(예를 들어, 레벨, 변조, 신호의 타입, 음향 피드백 추정들을 결정)하고, 그리고 순방향 경로의 프로세싱을 가능하게 제어하는 기능 컴포넌트들을 포함하는 분석 (또는 제어) 경로를 포함한다. 일 실시예에서, 분석 경로 및/또는 신호 경로의 일부 또는 모든 신호 프로세싱은, 주파수 도메인에서 수행된다. 일 실시예에서, 분석 경로 및/또는 신호 경로의 일부 또는 모든 신호 프로세싱은 시간 도메인에서 수행된다.
- [0061] 일 실시예에서, 분석 (또는 제어) 경로는 순방향 경로보다 더 적은 채널들(또는 주파수 서브-대역들)에서 동작한다. 이것은, 예를 들어, 휴대용 오디오 프로세싱 디바이스와 같은 오디오 프로세싱 디바이스, 예를 들어, 보청기에서 전력을 절약하기 위해 행해질 수 있고, 전력 소비는 중요한 파라미터이다.
- [0062] 일 실시예에서, 음향 신호를 나타내는 아날로그 전기 신호는, 아날로그-디지털(AD) 변환 프로세스에서 디지털 오디오 신호로 변환되고, 아날로그 신호는 미리정의된 샘플링 주파수 또는 속도  $f_s$ 로 샘플링되며,  $f_s$ 는 시간  $t_n$

(또는  $n$ )의 이산 포인트들에서 디지털 샘플들  $x_n$ (또는  $x[n]$ )을 제공하기 위한 (애플리케이션의 특정 요구들에 맞춰) 8 kHz 내지 48 kHz의 범위에 존재하고, 각 오디오 샘플은,  $t_n$ 에서 미리정의된 수  $N_s$ 의 비트들에 의해 음향 신호의 값을 나타내고,  $N_s$ 는, 예를 들어, 1 내지 16 비트들의 범위 내에 존재한다. 디지털 샘플  $x$ 는  $1/f_s$ 의 시간 길이를 갖는다. 예를 들어,  $f_s = 20$  kHz에 대해  $50\mu s$ 이다. 일 실시예에서, 다수의 오디오 샘플들이 시간 프레임 내에 배열된다. 일 실시예에서, 시간 프레임은 64 또는 128개의 오디오 데이터 샘플들을 포함한다. 다른 프레임 길이들은 실제 애플리케이션에 따라 사용될 수 있다. 일 실시예에서, 프레임은, 오버샘플링의 경우(예를 들어, 3.2 ms의 프레임 길이에 대응하는 임계 샘플링(프레임이 전혀 오버랩되지 않음)의 경우(예를 들어,  $f_s = 20$  kHz 이고, 그리고 프레임당 64 샘플들))에서 매 ms 또는 매 2ms 마다 시프트된다. 즉, 프레임들이 오버랩되어, 샘플들의 특정 부분, 예를 들어, 샘플들의 25% 또는 50% 또는 75% 만이 주어진 프레임에서 다음 프레임까지 새로워진다.

[0063] 일 실시예에서, 오디오 프로세싱 디바이스들은, 미리정해진 샘플링 속도, 예를 들어, 20kHz로 아날로그 입력을 디지털화하기 위해 아날로그-디지털(AD) 변환기를 포함한다. 일 실시예에서, 오디오 프로세싱 디바이스들은, 예를 들어, 출력 트랜스듀서를 통해 사용자에게 표시되도록, 디지털 신호를 아날로그 출력 신호로 변환하는 디지털-아날로그(DA) 변환기를 포함한다.

[0064] 일 실시예에서, 오디오 프로세싱 디바이스, 예를 들어, 마이크로폰 유닛 및/또는 송수신기 유닛은, 입력 신호의 시간-주파수 표현을 제공하기 위한 TF-변환 유닛을 포함한다. 일 실시예에서, 시간-주파수 표현은, 특정 시간 및 주파수 범위에서 해당 신호의 대응하는 복소값 또는 실수값의 어레이 또는 맵을 포함한다. 일 실시예에서, TF 변환 유닛은, (시간 변화) 입력 신호를 필터링하고, 그리고 각각이 입력 신호의 개별 주파수 범위를 포함하는 다수의 (시간 변화) 출력 신호들을 제공하는 필터 뱅크를 포함한다. 일 실시예에서, TF 변환 유닛은, 시간 변화 입력 신호를 주파수 도메인에서 (시간 변화) 신호로 변환하는 푸리에 변환 유닛을 포함한다. 일 실시예에서, 최소 주파수  $f_{min}$ 로부터 최대 주파수  $f_{max}$ 까지 오디오 프로세싱 디바이스에 의해 고려되는 주파수 범위는, 20Hz에서 20kHz의 통상적인 사람의 가청 주파수 범위의 일부, 예를 들어, 20Hz에서 12kHz까지의 일부를 포함한다. 일 실시예에서, 상기 오디오 프로세싱 디바이스의 순방향 및/또는 분석 경로의 신호는 수  $NI$ 의 주파수 대역들로 분할되고, 여기서  $NI$ 는, 예를 들어, 5보다 크고, 또는 10보다 크고, 또는 15보다 크고, 또는 50보다 크고, 또는 100보다 크고, 또는 500보다 크며, 그 중 적어도 일부는 개별적으로 프로세싱된다. 일 실시예에서, 오디오 프로세싱 디바이스는, 수  $NP$ 의 상이한 주파수 채널들( $NP < NI$ )에서 순방향 및/또는 분석 경로의 신호를 프로세싱하기 위해 적응된다. 주파수 채널들은, 폭이 균일 또는 불균일(예를 들어, 주파수와 함께 폭이 증가)할 수 있고, 오버랩핑되거나 또는 오버랩핑되지 않을 수 있다.

[0065] 일 실시예에서, 오디오 프로세싱 디바이스는, 오디오 프로세싱 디바이스의 현재 물리적 환경(예를 들어, 현재의 음향 환경), 및/또는 오디오 프로세싱 디바이스를 착용하는 사용자의 현재 상태, 및/또는 오디오 프로세싱 디바이스의 현재 상태 또는 동작 모드에 관련된 상태 신호들을 제공하도록 구성된 다수의 검출기들을 포함한다. 대안으로 또는 추가적으로, 하나 이상의 검출기들은 오디오 프로세싱 디바이스와 (무선으로) 통신하는 외부 디바이스의 일부를 형성할 수 있다. 외부 디바이스는, 예를 들어, 다른 청각 보조 디바이스, 원격 제어 및 오디오 전달 디바이스, 전화(예를 들어, 스마트폰), 외부 센서 등을 포함할 수 있다.

[0066] 일 실시예에서, 다수의 검출기들 중 하나 이상은, 전대역 신호(시간 도메인) 상에서 동작한다. 일 실시예에서, 다수의 검출기들 중 하나 이상은, 대역 분할 신호들((시간-)주파수 도메인) 상에서 동작한다.

[0067] 일 실시예에서, 다수의 검출기들은 순방향 경로의 신호의 현재 레벨을 추정하기 위한 레벨 검출기를 포함한다. 일 실시예에서, 미리정의된 기준은, 순방향 경로의 신호의 현재 레벨은 주어진(L-) 임계값 위 또는 아래인지 여부를 포함한다.

[0068] 특정 실시예에서, 오디오 프로세싱 디바이스는, 입력 신호가 (주어진 시점에서) 음성 신호를 포함하는지 여부를 결정하기 위한 음성 검출기(VD)를 포함한다. 음성 신호는, 본 맥락에서, 사람으로부터의 음성 신호를 포함하도록 취해진다. 음성 신호는, 또한, 사람의 스피치 시스템에 의해 발생된 다른 형태들의 발생들(예를 들어, 노래)을 포함할 수 있다. 일 실시예에서, 음성 검출기 유닛은 음성 또는 비음성 환경으로서 사용자의 현재 음향 환경을 분류하도록 적응된다. 이는, 사용자의 환경에서의 사람의 발생들(예를 들어, 스피치)을 포함하는 전기 마이크로폰 신호의 시간 세그먼트들은 식별될 수 있고, 따라서, 다른 사운드 소스들(예를 들어, 인공적으로 생성된 노이즈)만을 포함하는 시간 세그먼트들로부터 분리될 수 있는 이점을 갖는다. 일 실시예에서, 음성 검출기는 음성으로서 또한 사용자의 자신의 음성을 검출하도록 적응된다. 대안적으로, 음성 검출기는 음성의 검출로부

터 사용자 자신의 음성을 제외시키도록 적응된다.

- [0069] 일 실시예에서, 오디오 프로세싱 디바이스는, 주어진 입력 사운드(예를 들어, 음성)가 시스템의 사용자의 음성으로부터 유래하는지 여부를 검출하기 위한 자신의 음성 검출기를 포함한다. 일 실시예에서, 오디오 프로세싱 디바이스의 마이크로폰 시스템은, 사용자 자신의 음성과 다른 사람의 음성 사이를 구분할 수 있고 그리고 음성이 아닌 사운드들을 구분할 가능성이 있도록 되어 있다.
- [0070] 일 실시예에서, 청력 보조 디바이스는, 검출기들(의 적어도 일부)로부터의 입력 신호들 및 가능하게는 다른 입력들에 기초하여 현재 상황을 분류하도록 구성된 분류 유닛을 포함한다. 본 맥락에서, '현재 상황(current situation)'은 다음 중 하나 이상에 의해 정의되는 것으로 간주된다.
- [0071] a) 물리적 환경(예를 들어, 현재의 전자기 환경, 예를 들어, 오디오 프로세싱 디바이스에 의한 수신을 의도하거나 또는 의도하지 않는 (오디오 및/또는 제어 신호들을 포함하는) 전자기 신호들의 발생, 또는 음향 이외의 현재의 환경의 다른 특성들을 포함);
- [0072] b) 현재의 음향 상황(입력 레벨, 피드백 등);
- [0073] c) 사용자의 현재의 모드 또는 상태(움직임, 온도 등); 및
- [0074] d) 청각 보조 디바이스(선택된 프로그램, 마지막 사용자 상호작용 이후에 경과된 시간 등)의 현재 모드 또는 상태, 및/또는 오디오 프로세싱 디바이스와 통신하는 다른 디바이스.
- [0075] 일 실시예에서, 오디오 프로세싱 디바이스는 해당 애플리케이션에 대한 다른 관련 기능, 예를 들어, 압축, 증폭, 피드백 감소 등을 더 포함한다.
- [0076] 일 실시예에서, 오디오 프로세싱 디바이스는, 청각 디바이스, 예를 들어, 보청기, 또는 청각 장비, 또는 사용자의 귀에 위치되거나 또는 사용자의 외이도에 전적으로 또는 부분적으로 위치되도록 적응된 청각 장비와 같은 청각 디바이스를 포함한다.

[0078] 사용:

[0079] 일 양상에서, '실시예들의 상세한 설명' 및 청구범위들에서, 상술한 것과 같은 오디오 프로세싱 디바이스의 사용이 또한 제공된다. 오디오 분배를 포함하는 시스템에서 사용이 제공된다. 일 실시예에서, 사용은, 하나 이상의 청각 장비들, 헤드셋들, 이어폰들, 능동 귀 보호 시스템들 등을 포함하는 시스템들에, 예를 들어, 핸드프리 전화 시스템들, 원격 회의 시스템들, 비상 방송 시스템들, 가라오케 시스템들, 강의실 증폭 시스템들 등에 제공된다.

[0081] 방법:

[0082] 일 양상에서, 타겟 스피치 컴포넌트들 및 노이즈 컴포넌트들로 구성되는 시간 변화 사운드 신호를 나타내는 전기 입력 신호의 시간-주파수 표현  $Y(k,n)$ 의 선형적 신호 대 잡음비  $\zeta(k,n)$ 를 추정하는 방법 -  $k$  및  $n$ 은, 각각, 주파수 대역 인덱스 및 시간 프레임 인덱스이며 - 이 본 출원에 의해 또한 제공된다.

[0083] 상기 전기 입력 신호  $Y(k,n)$ 의 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 를 결정하는 단계와,

[0084] 재귀적 알고리즘에 기초하여 상기 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 로부터 상기 전기적 입력 신호의 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하는 단계와, 그리고

[0085]  $n-1$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n-1)$  및  $n$  번째 시간 프레임에 대한 상기 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 로부터 상기  $n$  번째 시간 프레임에 대한 상기 선형적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하는 단계를 포함한다.

[0086] 본 출원의 다른 양상에서, 타겟 스피치 컴포넌트들 및 노이즈 컴포넌트들로 구성되는 시간 변화 사운드 신호를 나타내는 전기 입력 신호의 시간-주파수 표현  $Y(k,n)$ 의 선형적 신호 대 잡음비  $\zeta(k,n)$ 를 추정하는 방법 -  $k$  및

n은, 각각, 주파수 대역 인덱스 및 시간 프레임 인덱스이며 - 은, 본 출원에 의해 또한 제공된다. 상기 방법은:

- [0087] 상기 전기 입력 신호  $Y(k,n)$ 의 후협적 신호 대 잡음비 추정치  $\gamma(k,n)$ 를 결정하는 단계와,
- [0088] 재귀적 알고리즘에 기초하여 상기 후협적 신호 대 잡음비 추정치  $\gamma(k,n)$ 로부터 상기 전기적 입력 신호의 선협적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 를 결정하는 단계를 포함하고, 상기 재귀적 알고리즘은 적응 시간 상수 또는 지역 차단 주파수를 갖는 지역 통과 필터를 구현한다.
- [0089] '실시예들의 상세한 설명' 및 청구범위들에서, 상술한 디바이스의 구조적 피처의 일부 또는 모두는, 대응하는 프로세스에 의해 적절하게 대체될 때 방법의 실시예들과 결합될 수 있거나 또는 그 반대일 수 있도록 의도된다. 방법의 실시예들은 대응하는 디바이스들과 동일한 이점들을 갖는다.
- [0090] 일 실시예에서, 상기 타겟 스피치 컴포넌트들의 크기들  $\hat{A}(k,n)$ 의 추정치들은, 이득 함수  $G$ 를 곱한 상기 전기 입력 신호  $Y(k,n)$ 로부터 결정되고, 상기 이득 함수  $G$ 는 상기 후협적 신호 대 잡음비 추정치  $\gamma(k,n)$  및 상기 선협적 타겟 신호 대 잡음비 추정치  $\zeta(k,n)$ 의 함수이다.
- [0091] 일 실시예에서, 상기 방법은, 높은 SNR 조건들에 비해 낮은 SNR 조건들에서 평활화를 더 가능하게 하는, SNR-의존 평활화를 제공하는 단계를 포함한다.
- [0092] 일 실시예에서, 상기 방법은, 평활화 파라미터( $\lambda$ ) 및/또는 바이어스 파라미터( $\rho$ ) 및/또는 바이패스 파라미터  $K$ 를 포함한다.
- [0093] 일 실시예에서, 평활화 파라미터( $\lambda$ ) 및/또는 바이어스 파라미터( $\rho$ )는 후협적 SNR  $\gamma$ 에 의존하거나, 또는 전기 입력 신호  $|Y|^2$  및 노이즈 스펙트럼 밀도  $\langle \sigma^2 \rangle$ 에 의존한다. 일 실시예에서, 평활화 파라미터( $\lambda$ ) 및/또는 바이어스 파라미터( $\rho$ ) 및/또는 상기 파라미터  $K$ 는 사용자의 난청, 인지 기술들 또는 스피치 명료도 스코어에 따라 선택된다. 일 실시예에서, 평활화 파라미터( $\lambda$ ) 및/또는 바이어스 파라미터( $\rho$ ) 및/또는 상기 파라미터  $K$ 는, 청각 능력, 인지 기술 또는 스피치 명료도 기술들이 해당 사용자에게 대해 불량해질수록 더 많은 평활화를 제공하도록 선택된다.
- [0094] 일 실시예에서, 상기 방법은 필터 뱅크 오버샘플링을 고려하기 위해 평활화 파라미터( $\lambda$ )를 조정하는 단계를 포함한다.
- [0095] 일 실시예에서, 상기 방법은, 상기 평활화 및/또는 바이어스 파라미터들이 상기 입력이 증가하는지 또는 감소하는지에 의존하는 것을 제공하는 단계를 포함한다.
- [0096] 일 실시예에서, 상기 방법은, 평활화 파라미터( $\lambda$ ) 및/또는 바이어스 파라미터( $\rho$ ) 및/또는 상기 파라미터  $K$ 가 사용자 인터페이스로부터 선택가능한 것을 제공하는 단계를 포함한다. 일 실시예에서, 상기 사용자 인터페이스는 스마트 폰의 APP로서 구현된다.
- [0097] 일 실시예에서, 상기 방법은, 선택된 최소값  $\xi_{\min}^{ML}$ 에 의한 n번째 시간 프레임 최대 우도에 대한 선협적 타겟 신호 대 잡음비  $\zeta(k,n)$ 의 최대 우도 SNR 추정자  $\zeta^{\text{ML}}(k,n)$ 의 사전 평활화를 제공하는 단계를 포함한다. 이는
- [0098] 
$$\frac{|Y_n|^2}{\hat{\sigma}_n^2} < 1$$
- [0099] 인 경우를 대처하는데 사용된다.
- [0100] 일 실시예에서, 재귀적 알고리즘은, 최대 우도 추정치가 바이어스 및 평활화 파라미터들의 계산에서 이전의 프레임의 선협적 추정치를 바이패스하도록 구성된다. 일 실시예에서, 현재의 최대 우도 SNR 추정치  $s_n^{\text{ML}}$ 에서 파라미터  $K$ 를 뺀 것이 이전의 선협적 SNR 추정치  $s_{n-1}$ 보다 큰 경우, 재귀적 알고리즘은, 현재의 최대 우도 SNR 추정치  $s_n^{\text{ML}}$ 가 이전의 프레임의 선협적 추정치  $s_{n-1}$ 를 바이패스하도록 구성된다. 일 실시예에서, 도 4의 맵핑 유닛

MAP에 공급되는 값은 도 4에서 도시된 것처럼  $s_n^{ML} - K$ 이지만, 다른 실시예에서,  $s_n^{ML}$ 은 (조건  $(s_n^{ML} - K > s_{n-1})$ 이 충족될 때) 맵핑 유닛 MAP에 직접 공급된다. 일 실시예에서, 재귀적 알고리즘은 재귀 루프에 위치한 최대 연산자를 포함하여, 최대 우도 SNR 추정치가 파라미터  $K$ 를 통해 바이어스 파라미터 및 평활화 파라미터의 계산에서 이전의 프레임의 선형적 추정치를 바이패스하도록 한다. 이로써, (큰) SNR 온셋들이 즉시 검출될 수 있다 (따라서, 스피치 온셋들의 위험 과다-감쇠가 감소될 수 있다).

[0101] 일 실시예에서, 상기 전기 입력 신호  $Y(k,n)$ 의 후험적 신호 대 잡음비 추정치  $\gamma(k,n)$ 는 제1 및 제2 후험적 신호 대 잡음비의 혼합으로서 생성되는 결합된 후험적 신호 대 잡음비로서 제공된다. (후험적 추정치들과 다른) 다른 결합들이 사용될 수 있다(예를 들어, 잡음 분산 추정치  $\langle \sigma^2 \rangle$ ).

[0102] 일 실시예에서, 2개의 후험적 신호 대 잡음비는 단일 마이크로폰 구성 및 다중 마이크로폰 구성으로부터 각각 생성된다. 일 실시예에서, 제1 후험적 신호 대 잡음비는 제2 후험적 신호 대 잡음비보다 빠르게 생성된다. 일 실시예에서, 결합된 후험적 신호 대 잡음비는 제1 및 제2 후험적 신호 대 잡음비들의 가중된 혼합으로서 생성된다. 일 실시예에서, 동측 보청기의 후험적 신호 대 잡음비에 결합되는 제1 및 제2 후험적 신호 대 잡음비들은, 각각, 바이노럴 보청기 시스템의 동측 보청기 및 반대측 보청기로부터 발생한다.

[0104] 컴퓨터 판독가능 매체

[0105] 일 양상에서, '실시예들의 상세한 설명' 및 청구범위들에서, 데이터 프로세싱 시스템으로 하여금 상술한 방법의 단계들의 적어도 일부(예를 들어, 다수 또는 전부)를 수행하게 하는 프로그램 코드 수단을 포함하는 컴퓨터 프로그램을 저장하는 유형의 컴퓨터-판독가능한 저장 매체가, 상기 컴퓨터 프로그램이 데이터 프로세싱 시스템 상에서 실행될 때, 본 출원에 의해 또한 제공된다.

[0106] 예를 들어, 그와 같은 컴퓨터-판독가능한 매체는, RAM, ROM, EEPROM, CD-ROM 또는 다른 광 디스크 저장 장치, 자기 디스크 저장 장치 또는 다른 자기 저장 장치, 또는 명령어들이나 데이터 구조들의 형태로 원하는 프로그램 코드를 수행하거나 또는 저장하기 위해 사용될 수 있고 그리고 컴퓨터에 의해 액세스될 수 있는 어떤 다른 매체를 포함할 수 있다. 여기에서 사용되는, 디스크들(disk 및 disc)은, CD(Compact Disc), 레이저 디스크, 광 디스크, DVD(Digital Versatile Disc), 플로피 디스크 및 블루-레이 디스크를 포함하고, 여기서 디스크(disk)들은 일반적으로 데이터를 자기적으로 재생하는 반면에 디스크(disc)들은 레이저들로 광학적으로 데이터를 재생한다. 상기의 조합들은 또한 컴퓨터 판독가능한 매체의 범위 내에 포함되어야 한다. 유형의 매체 상에 저장되는 것 이외에, 컴퓨터 프로그램은 또한 유선 또는 무선 링크 또는 네트워크와 같은 전송 매체, 예를 들어, 인터넷을 통해 전송될 수 있고, 그리고 상기 유형의 매체와 다른 위치에서 실행되도록 데이터 프로세싱 시스템에 로딩될 수 있다.

[0108] 컴퓨터 프로그램

[0109] '실시예들의 상세한 설명' 및 청구범위들에서, 프로그램이 컴퓨터에 의해 실행될 때, 상기 컴퓨터로 하여금 상술한 방법(의 단계들)을 수행하게 하는 명령어들을 포함하는 컴퓨터 프로그램 (제품)이, 본 출원에 의해 또한 제공된다.

[0111] 데이터 프로세싱 시스템

[0112] 일 양상에서, '실시예들의 상세한 설명' 및 청구범위들에서, 프로세서로 하여금 상술한 방법의 단계들의 적어도 일부(예를 들어, 다수 또는 전부)를 수행하게 하는 프로세서 및 프로그램 코드 수단을 포함하는 데이터 프로세싱 시스템이 본 출원에 의해 또한 제공된다.

[0114] 청각 시스템

[0115] 또 다른 양상에서, '실시예들의 상세한 설명' 및 청구범위들에서, 상술한 오디오 프로세싱 디바이스를 포함하는

청각 시스템 및 보조 디바이스가 또한 제공된다.

- [0116] 일 실시예에서, 상기 시스템은, 오디오 프로세싱 디바이스와 보조 디바이스 사이에 통신 링크를 설정하여, 그와 같은 정보(예를 들어, 제어 및 상태 신호들, 가능하면 오디오 신호들)가 서로 교환될 수 있거나 전달될 수 있도록 제공되게 구성된다.
- [0117] 일 실시예에서, 오디오 프로세싱 디바이스는 청각 디바이스, 예를 들어, 보청기이거나 보청기를 포함한다. 일 실시예에서, 오디오 프로세싱 디바이스는 전화이거나 전화를 포함한다.
- [0118] 일 실시예에서, 보조 디바이스는 (예를 들어, TV 또는 뮤직 플레이어와 같은 엔터테인먼트 디바이스, 이동 전화기와 같은 전화기 장치 또는 PC와 같은 컴퓨터로부터) 다수의 오디오 신호들을 수신하도록 구성되고 그리고 상기 오디오 프로세싱 디바이스에 송신하기 위해 수신된 오디오 신호들(또는 신호들의 조합) 중 적절한 것을 선택 및/또는 결합하도록 구성된다. 일 실시예에서, 보조 디바이스는 오디오 프로세싱 디바이스 또는 청각 디바이스(들)의 기능 및 동작을 제어하는 원격 제어기이거나 또는 원격 제어기를 포함한다. 일 실시예에서, 원격 제어기의 기능은 스마트폰으로 구현되고, 스마트폰(예를 들어, 블루투스 또는 다른 표준화된 또는 독점적인 방식에 기초하여 스마트폰에 대한 적절한 무선 인터페이스를 포함하는 오디오 프로세싱 디바이스(들))을 통해 오디오 프로세싱 디바이스의 기능을 제어하도록 하는 APP의 구동을 상기 스마트 폰은 가능하게 한다.
- [0119] 일 실시예에서, 보조 디바이스는, 다른 오디오 프로세싱 디바이스, 예를 들어, 보청기와 같은 청각 디바이스이다. 일 실시예에서, 청각 시스템은, 바이노럴(binaural) 청각 시스템, 예를 들어, 바이노럴 보청기 시스템을 구현하도록 적응된 2개의 청각 디바이스들을 포함한다.
- [0121] APP
- [0122] 또 다른 양상에서, APP로 언급되는 비-일시적인 애플리케이션이 본 발명에 의해 또한 제공된다. 상기 APP는, '실시예들의 상세한 설명' 및 청구범위들에서, 상술한 청각 디바이스 또는 청각 시스템에 대한 사용자 인터페이스를 구현하기 위해 보조 디바이스에서 실행되도록 구성된 실행가능한 명령어들을 포함한다. 일 실시예에서, APP는, 셀룰러 폰, 예를 들어, 스마트폰 상에서 구현되거나 또는 상기 청각 디바이스 또는 상기 청각 시스템과 통신을 가능하게 하는 다른 휴대용 디바이스 상에서 구현되도록 구성된다.
- [0124] 정의들:
- [0125] 본 명세서에서, '청각 디바이스(hearing device)'는, 예를 들어, 보청기 또는 능동 귀-보호 디바이스 또는 다른 오디오 프로세싱 디바이스와 같은 디바이스에 관한 것으로, 사용자의 주변으로부터 음향 신호들을 수신하고, 대응하는 오디오 신호들을 생성하고, 상기 오디오 신호들의 변경을 가능하게 하며 그리고 사용자의 귀들 중 적어도 하나에 가청 신호들로서 변경이 가능한 오디오 신호들을 제공함으로써 사용자의 청각 능력을 향상시키고, 보강하며 그리고 보호하도록 구성된다. '청각 디바이스'는 또한 오디오 신호들을 전자적으로 수신하고, 상기 오디오 신호들의 변경을 가능하게 하며, 그리고 사용자의 귀들 중 적어도 하나에 가청 신호들로서 변경이 가능한 오디오 신호들을 제공하도록 구성된 이어폰 또는 헤드셋과 같은 디바이스에 관한 것이다. 그와 같은 가청 신호들은, 예를 들어, 사용자의 귀의 바깥쪽들에 방사되는 음향 신호들, 사용자의 머리의 뼈구조를 통해 그리고/또는 귀의 가운데의 부분들을 통해 사용자의 귀의 안쪽들에 기계적 진동들로서 전달되는 음향 신호들과 함께 사용자의 와우 신경(cochlear nerve)에 직접적으로 또는 간접적으로 전달되는 전기 신호들의 형태로 제공될 수 있다.
- [0126] 청각 디바이스는, 어떤 알려진 방식으로, 예를 들어, 방사된 음향 신호들을 외이도(ear canal)로 인도하는 튜브를 갖거나 또는 상기 외이도에 배치되거나 외이도에 가깝게 배치된 라우드스피커를 갖는 귀 뒤에 배치된 유닛으로, 또는 귓바퀴 및/또는 외이도에 전체적으로 또는 부분적으로 배치되는 유닛으로, 또는 두개골에 이식된 고정구에 부착되는 유닛으로, 또는 전체적이거나 또는 부분적으로 이식된 유닛으로 착용되도록 구성될 수 있다. 청각 디바이스는 단일 유닛 또는 서로 전자적으로 통신하는 여러 유닛들을 포함할 수 있다.
- [0127] 보다 일반적으로, 청각 디바이스는 사용자의 주변들로부터 음향 신호를 수신하고 그리고 대응하는 입력 오디오 신호를 제공하기 위한 입력 트랜스듀서, 및/또는 입력 오디오 신호를 전자적으로(즉, 유선 또는 무선으로) 수신하기 위한 수신기, 상기 입력 오디오 신호를 처리하기 위한 (통상적으로 구성가능한) 신호 프로세싱 회로 및 상기 프로세싱된 오디오 신호에 따라 가청 신호를 상기 사용자에게 제공하는 출력 수단을 포함한다. 일부 청각 디바이스들에서, 증폭기가 신호 프로세싱 회로를 구성할 수 있다. 신호 프로세싱 회로는, 통상적으로, 프로그램들

을 실행하고 그리고/또는 상기 프로세싱에서 사용되는 (또는 잠재적으로 사용될) 파라미터들을 저장하고 그리고/또는 청각 디바이스의 기능에 관련된 정보를 저장하고 그리고/또는 예를 들어, 사용자에게 대한 인터페이스 및/또는 프로그래밍 디바이스에 대한 인터페이스와 관련하여 사용하기 위한 정보(예를 들어, 신호 프로세싱 회로에 의해 제공되는 프로세싱된 정보)를 저장하기 위한 하나 이상의 (집적되거나 또는 분리된) 메모리 소자들을 포함한다. 일부 청각 디바이스들에서, 출력 수단은, 예를 들어, 공기로 전파되는 음향 신호를 제공하기 위한 라우드스피커 또는 고체로 전파되거나 액체로 전파되는 음향 신호를 제공하기 위한 바이브레이터와 같은 출력 트랜스듀서를 포함할 수 있다. 일부 청각 디바이스들에서, 출력 수단은, 전기 신호들을 제공하기 위한 하나 이상의 출력 전극들을 포함할 수 있다.

[0128] 일부 청각 디바이스들에서, 진동기는, 고체로 전파되는 음향 신호를 두개골에 피부를 통해 제공하도록 구성될 수 있다. 일부 청각 디바이스들에서, 진동기는 중이(middle ear) 및/또는 내이(inner ear)에 이식될 수 있다. 일부 청각 디바이스들에서, 진동기는, 중이 뼈(middle-ear bone) 및/또는 와우관(cochlea)에 고체로 전파되는 음향 신호를 제공하도록 구성될 수 있다. 일부 청각 디바이스들에서, 진동기는, 예를 들어, 난원창(oval window)을 통해 와우관의 액체(cochlear liquid)에 액체로 전달되는 신호를 제공하도록 구성된다. 일부 청각 디바이스들에서, 출력 전극들은 와우관에 또는 두개골 뼈의 내부에 이식될 수 있고 그리고 전기 신호를 달팽이관의 유모 세포들에, 하나의 이상의 청각 신경에, 청각 피질에 그리고/또는 대뇌 피질의 다른 부분들에 제공하도록 구성될 수 있다.

[0129] '청각 시스템'은 하나 또는 두 개의 청각 디바이스들을 포함하는 시스템에 관한 것이고 그리고 '바이노럴 청각 시스템'은 2개의 청각 디바이스들을 포함하고 그리고 사용자의 귀들 모두에 협조적으로 가청 신호들을 제공하도록 구성되는 시스템에 관한 것이다. 청각 시스템들 또는 바이노럴 청각 시스템들은 하나 이상의 '보조 디바이스들'을 더 포함할 수 있고, 이 '보조 디바이스들'은 청각 디바이스(들)과 통신하고 그리고 청각 디바이스(들)의 기능에 영향을 주고 그리고/또는 청각 디바이스(들)의 기능으로부터 이익을 얻는다. 보조 디바이스들은, 예를 들어, 원격 제어기들, 오디오 게이트웨이 디바이스들, 이동 전화기들(예를 들어, 스마트폰들), 비상 방송 시스템들, 자동차 오디오 시스템들 또는 뮤직 플레이어들일 수 있다. 청각 디바이스들, 청각 시스템들 또는 바이노럴 청각 시스템들은, 예를 들어, 청각 장애인의 청각 기능 상실을 보상하고, 청각이 정상인 사람의 청각 기능을 보강 또는 보호하며 그리고/또는 전자 오디오 신호들을 사람에게 전달하기 위해 사용될 수 있다.

[0130] 본 발명의 실시예들은, 예를 들어, 보청기들, 헤드셋들, 이어폰들, 능동형 뒤 보호 시스템들, 핸즈프리 전화 시스템들, 이동 전화들 등과 같은 애플리케이션들에서 유용할 수 있다.

**도면의 간단한 설명**

[0131] 본 발명의 양상들은 첨부된 도면들과 관련하여 얻어진 다음의 상세한 설명으로부터 가장 잘 이해될 수 있다. 상기 도면들은 명확성을 위해 개략적이고 그리고 단순화되어 있으며, 그리고 이러한 도면들은 청구 범위들의 이해를 향상시키기 위한 세부 사항만을 도시하고 있는 반면에, 다른 부분들은 배제된다. 전체적으로, 동일한 참조 번호들이 동일한 또는 대응하는 부분에 대해 사용된다. 각 양상의 개별적인 특징은 각각 다른 양상의 일부 또는 모든 피처들과 결합될 수 있다. 이들 및 다른 양상들, 피처들 및/또는 기술적 효과는 이하에서 설명되는 도면들을 참조하여 명확해지고 그리고 명료해질 것이다.

도 1a는 단일-채널 노이즈 감소 유닛을 도시하며, 단일 마이크로폰(M)은 타겟 사운드(x) 및 노이즈(v)의 혼합 y(t)를 획득하고, 그리고

도 1b는 다중-채널 노이즈 감소 유닛을 도시하며, 다수의 마이크로폰(들)(M<sub>1</sub>, M<sub>2</sub>)은 타겟 사운드(x) 및 노이즈(v)의 혼합 y(t)를 획득한다.

도 2는 참(true) SNR의 함수로서 최대 우도 추정자  $\xi_n^{ML}$ 의 평균값을 [dB]로 나타내고, 최대 우도 선형적 SNR 추정  $\xi_n^{ML} = \max(\xi_{min}^{ML}, \gamma_n - 1)$ 에서 단방향 정류에 의해 도입된 바이어스를 도시한다.

도 3은 STSA [1] 이득 함수( $\alpha = 0.98$ )에 대한 식 (7)의 수치 평가에 의해 DD\*-알고리즘의 입력-출력 관계 ( $\Delta_{출력} = f(\Delta_{입력})$ )를 도시한다.

도 4는 제안되는 지향성 바이어스 및 평활화 알고리즘(DBSA, 유닛 Po2Pr에 의해 구현됨)의 예시적인 구현의 다이어그램을 도시한다.

도 5는 결정 지향성 접근법에 의해 주어진 파라미터들로부터  $\rho$  및  $\lambda$  가 어떻게 도출될 수 있는지를 도시한다.

도 6a는 STSA 이득 함수,  $\alpha = 0.98$ 에 대한 함수  $10 \log_{10} \Psi \left( \xi_{n-1}, \frac{\xi_n^{ML}}{\xi_{n-1}} \right)$  의 기울기  $\lambda$ 를 도시한다.

도 6b는 STSA 이득 함수,  $\alpha = 0.98$ 에 대한 함수  $10 \log_{10} \Psi \left( \xi_{n-1}, \frac{\xi_n^{ML}}{\xi_{n-1}} \right)$  의 제로 크로싱  $\rho$ 를 도시한다.

도 7은 본 발명에 따른 DBSA 알고리즘(십자형) 및 DD-알고리즘(라인)의 응답들을 도 6a, 도 6b의 적합한 함수들을 사용하여 비교하는 것을 도시하며, 곡선들은 -30dB 내지 +30dB 범위의 선형적 SNR 값들을 5dB 간격으로 나타낸다.

도 8은 필터 뱅크 오버샘플링을 수용하기 위한 DBSA 알고리즘(도 4에 도시됨)의 변형을 도시하며, 재귀 루프에서 추가적인 D-프레임 지연을 삽입하는 목적은 보다 적은 오버 샘플링을 갖는 시스템의 동적 거동을 모방하는 것이다.

도 9a는 본 발명에 따른 오디오 프로세싱 디바이스, 예를 들어, 보청기의 일 실시예를 도시한다.

도 9b는, 예를 들어, 도 9a의 예시적인 오디오 프로세싱 디바이스에서의 사용을 위해 (M=2), 본 발명의 따른 노이즈 감소 시스템의 일 실시예를 도시한다.

도 10은, 2개의 후협적 신호 대 잡음비들(하나는 단일 마이크로폰 채널로부터 생성되고 그리고 다른 하나는 다중-마이크로폰 구성으로부터 생성됨)로부터 결합된 하나의 후협적 신호 대 잡음비의 생성을 도시한다.

도 11은, 귀 또는 사용자의 뒤에 배치된 BTE-부분 및 상기 사용자의 외이도에 배치된 ITE 부분을 포함하는 본 발명에 따른 보청기의 일 실시예를 도시한다.

도 12a는 제안되는 지향성 바이어스 및 평활화 알고리즘(DBSA, 예를 들어, 도 1a, 1b의 유닛 Po2Pr에 의해 구현됨)의 제1의 다른 예시적인 구현의 다이어그램을 도시한다.

도 12b는 제안되는 지향성 바이어스 및 평활화 알고리즘의 제2의 다른 예시적인 구현의 다이어그램을 도시한다.

도 12c는 제안되는 지향성 바이어스 및 평활화 알고리즘의 제3의 다른 예시적인 구현의 다이어그램을 도시한다.

도 13a는 도 12a, 도 12b, 도 12c에 도시된 DBSA 알고리즘들의 실시예들에서 사용하기 위한 온셋 플래그를 제공하는 일반적인 예를 도시한다.

도 13b는 도 12a, 도 12b, 도 12c에 도시된 DBSA 알고리즘들의 실시예들에서 사용 가능하도록 온셋 플래그를 제공하는 이웃하는 주파수 대역들로부터의 입력에 기초한 온셋 검출기(제어기)의 예시적인 실시예를 도시한다.

상기 도면들은 명확성을 위해 개략적이고 그리고 단순화되며, 이 도면들은 본 발명의 이해에 필수적인 세부 사항들만을 도시하는 반면에, 다른 부분들은 배제된다. 전체적으로, 동일한 참조 번호들이 동일하거나 또는 대응하는 부분들에 대해 사용된다.

본 발명의 추가적인 적용가능성의 범위는 이하에 주어진 상세한 설명으로부터 명백해질 것이다. 하지만, 상세한 설명 및 특정 실시예는 본 발명의 바람직한 실시예를 나타내지만, 단지 예시로서 주어진 것으로 이해되어야 한다. 다른 실시예들은 다음의 상세한 설명으로부터 당업자들에게 명백해 질 수 있다.

**발명을 실시하기 위한 구체적인 내용**

[0132] 첨부된 도면들과 관련하여 아래에서 설명되는 상세한 설명은 다양한 구성의 설명으로 의도된다. 상세한 설명은 다양한 개념의 철저한 이해를 제공하기 위한 목적의 구체적인 세부사항들을 포함한다. 하지만, 통상의 기술자에게는, 이러한 개념들이 이러한 특정 세부 사항들없이 실시될 수 있음이 명백할 것이다. 장치 및 방법의 여러 양상들은 다양한 블록들, 기능 유닛들, 모듈들, 컴포넌트들, 회로들, 단계들, 프로세스들, 알고리즘들(집합적으로 "요소들"이라고 언급됨)에 의해 기술된다. 특정 애플리케이션, 설계 제약들 또는 다른 이유들로 인해, 이러한 요소들은 전자 하드웨어, 컴퓨터 프로그램 또는 이들의 어떤 조합을 사용하여 구현될 수 있다.

[0133] 전자 하드웨어는, 마이크로프로세서들, 마이크로제어기들, 디지털 신호 프로세서(DSP)들, 필드 프로그래머블 게이트 어레이(FPGA)들, 프로그래머블 로직 디바이스(PLD)들, 게이트 로직, 개별 하드웨어 회로들, 및 본 명세서 전체에 걸쳐 설명된 다양한 기능을 수행하도록 구성된 다른 적절한 하드웨어를 포함한다. 컴퓨터 프로그램은,

명령어들, 명령어 세트들, 코드, 코드 세그먼트들, 프로그램 코드, 프로그램들, 서브프로그램들, 소프트웨어 모듈들, 애플리케이션들, 소프트웨어 애플리케이션들, 소프트웨어 패키지들, 루틴들, 서브루틴들, 객체들, 실행파일들, 실행 스프레드들, 절차들, 소프트웨어, 펌웨어, 미들웨어, 마이크로코드, 하드웨어 설명 언어 또는 다른 것으로 언급되는 기능들 등을 의미하는 것으로 폭넓게 해석될 것이다.

[0134] 본 출원은 청각 디바이스들, 예를 들어, 보청기의 분야에 관한 것이다.

[0135] 스피치 향상 및 노이즈 감소는 시간-주파수 도메인에서 급변하는 이득을 적용함으로써 획득될 수 있다. 고속 변화 이득을 적용하는 목적은, 잡음에 의해 지배되는 시간-주파수 타일들이 억제되는 동안 스피치에 의해 지배되는 시간-주파수 타일들을 변경하지 않고 유지하는 것이다. 이로써, 향상된 신호의 변조의 발생은 증가하고 그리고 통상적으로, 원래의 스피치 신호의 변조와 유사해져서, 더 높은 음성 명료도를 유도한다.

[0136] 관측된 신호  $y(t)$ 는, 주파수  $k$ (주파수 인덱스  $k$ 는 수식의 간략함을 위해 여기서 생략) 및 시간 프레임  $n$ (예를 들어, 도 1a, 도 1b 참조)에 대응하는 주파수 서브-밴드 신호들  $Y_{kn}(Y(n,k))$ 을 산출하기 위해 분석 필터 뱅크(FBA;  $FBA_1, FBA_2$ )에서 처리되는 타겟 스피치 신호  $x(t)$  및 노이즈  $v(t)$ (예를 들어, 하나의 마이크로 폰 또는 다수의 마이크로폰들에 의해 획득됨)의 합(sum)임을 가정하자. 예를 들어,  $Y_n$ 은 DFT 필터 뱅크로부터 얻어진 복소 계수들을 포함할 수 있다(또는 복소 계수들로 구성될 수 있다). 스펙트럼 스피치 향상 방법들은, 각 시간-주파수 타일에서의 노이즈의 양( $N$ )과 비교된 타겟 신호의 양( $X$ ), 즉 신호-대-잡음(SNR) 비를 추정하는 것에 의존한다. 스펙트럼 노이즈 감소에서, SNR은 통상적으로 2개의 서로 다른 용어들을 사용하여 기술된다:

[0137] 1) 
$$\gamma_n = \frac{|Y_n|^2}{\hat{\sigma}_n^2}$$
 로서 정의되는 후험적 SNR,

[0138] 여기에서,  $\hat{\sigma}_n^2$ 는  $n$ 번째 시간 프레임에서의 잡음 스펙트럼 밀도(잡음 스펙트럼 전력 분산)의 추정치이며, 그리고

[0139] 2) 
$$\xi_n = \frac{\langle |X_n|^2 \rangle}{\hat{\sigma}_n^2}$$
 로서 정의되는 선험적 SNR,

[0140] 여기에서,  $|X_n|^2$ 는 타겟 신호 스펙트럼 밀도이다. 후험적 SNR은 노이즈 전력 스펙트럼 밀도  $\hat{\sigma}_n^2$ 의 추정을 요구하는 반면에, 선험적 SNR은 스피치 스펙트럼 밀도( $|X_n|^2$ ) 및 노이즈 전력 스펙트럼 밀도( $\hat{\sigma}_n^2$ ) 모두에 대한 액세스를 요구한다.

[0141] 선험적 SNR이 이용가능하면, 각각의 유닛에 대해, 시간 및 주파수에서, 타겟 신호의 추정을

[0142] 
$$\hat{X}_n = \frac{\xi_n}{\xi_n + 1} Y_n$$
 로서 발견할 수 있고, 
$$\hat{X}_n = \frac{\xi_n}{\xi_n + 1} Y_n$$
 는 위너 이득 접근법(Wiener gain approach)을 나타낸다.

하지만, 이득 함수들에 대한 다른 SNR이 사용될 수 있다. 용어 '후험적(posteriori)' 및 '선험적(a priori)' 신호-대-잡음비는, 예를 들어, [4]에서 사용된다.

[0143] 도 1a는, 단일 마이크로폰(M)이 타겟 사운드( $x$ ) 및 노이즈( $v$ )의 혼합  $y(t)$ 을 수신하는 단일-채널 노이즈 감소 유닛을 도시하고, 그리고 도 1b는 다수의 마이크로폰들( $M_1, M_2$ )이 타겟 사운드( $x$ ) 및 노이즈( $v$ )의 혼합  $y(t)$ 을 수신하는 다중-채널 노이즈 감소 유닛을 도시한다.

[0144] 본 발명에서, 아날로그-대-디지털 변환 유닛들은, 마이크로폰들로부터 디지털화된 전기 입력 신호들을 제공하기 위해 적절하게 적용되는 것으로 가정한다. 마찬가지로, 디지털 대 아날로그 변환 유닛(들)은, 적절하다면, 출력 신호들(예를 들어, 라우드 스피커에 의해 음향 신호로 변환될 신호들)에 적용된다고 가정된다.

[0145] 혼합(들)은 각 분석 필터 뱅크들(FBA (분석))으로 표시됨, 도 1a 및 도 1b에서 각각  $FBA_1$  (분석),  $FBA_2$  (분석)으로 표시됨에 의해 주파수 도메인으로 변환되고, 그리고 신호  $Y(n,k)$ (도 1a 및 도 1b에서 각각  $Y(n,k)_1$  및  $Y(n,k)_2$ 으로 표시됨)를 획득한다. 각각의 경우에서, 후험적 SNR  $\gamma$ (도 1a 및 도 1b에서 후험적 SNR,  $\gamma_n$ )은, 타겟 신호를 포함하는 전력 스펙트럼 밀도  $|Y_n|^2$ (각 크기 제곱 계산 유닛들  $| \cdot |^2$ 에 의해 제공됨)와 혼합(도 1a, 도 1b의

조합 유닛 ‘./’ 참조) 내의 노이즈 전력 스펙트럼 밀도  $\hat{\sigma}_n^2$  (도 1a, 도 1b 에서  $\langle \sigma^2 \rangle$ 으로 표시되고 그리고 각 노이즈 추정 유닛들 NT에 의해 제공됨)의 추정 사이의 비율로서 발견된다. 하나보다 많은 마이크로폰(예를 들어, 도 1b)의 경우, 혼합 내의 잡음은 마이크로폰 신호들의 선형 조합  $Y(n,k) = w(k)_1 \cdot Y(n,k)_1 + w(k)_2 \cdot Y(n,k)_2$  에 의해 감소될 수 있고, 그리고 나머지 노이즈는, 도 1b에서의 빔 형성기 필터링 유닛(BFU)으로부터의 출력 신호들에 의해 표시된 것처럼, 타겟 신호  $N(n,k) = w(k)_3 \cdot Y(n,k)_1 + w(k)_4 \cdot Y(n,k)_2$ 를 소거하는 것을 목표로 하는 마이크로폰 신호들의 다른 선형 조합 ( $N(n,k)$ )에 의해 더 잘 추정될 수 있다.

[0146] 선형적 신호 대 잡음비(도 1a, 도 1b에서의 선형적 SNR,  $\xi_n$ )는 본 발명에 따른 알고리즘을 구현하는 변환 유닛 Po2Pr에 의해 결정되고, 이는 이하에서 더 설명된다. 상기 선형적 SNR은, 예를 들어, 결과적인 전류 잡음 감소 이득(예를 들어, 위너 이득 함수에 기초함)를 제공하는 선택적 SNR 대 이득 변환 유닛 SNR2G에서의 이득으로 변환될 수 있고, 이는 노이즈 감소 신호  $Y_{NR}(n,k)$ 를 제공하기 위해 조합 유닛 'X'에서 신호  $Y(n,k)$ (도 1a의 입력 신호 및 도 1b에서의 공간적으로 필터링된 신호)에 적용될 수 있다.

[0147] 노이즈 전력 밀도  $\hat{\sigma}_n^2$ 의 추정치(도 1a, 도 1b에서  $\langle \sigma^2 \rangle$ 으로 표시됨)가 이용가능할 때, 후험적 SNR을 직접 발견할 수 있다(도 1a, 도 1b에서의 조합(여기서는 분할) 유닛 ‘./’ 참조). 일반적으로, 타겟 전력 스펙트럼 밀도 ( $A_n^2$ )( $A_n$ 은 알려지지 않은 타겟 크기  $|X_n|$ 의 추정치)에 대해 액세스를 하지 못하므로, 선형적 SNR로 직접 액세스를 할 수 없다. 선형적 SNR을 추정하기 위해, 결정 지시(DD) 알고리즘이 제안된다 [1]:

$$\xi_n = \alpha \frac{\hat{A}_{n-1}^2}{\hat{\sigma}_n^2} + (1 - \alpha) \max(0, \gamma_n - 1) \quad (1)$$

[0149] 여기서,  $\hat{A}_n$ 는 (n번째 시간 프레임에서) 타겟 신호 크기의 추정치이고,  $\hat{\sigma}_n^2$ 는 해당 주파수에서 노이즈 스펙트럼 분산(전력 스펙트럼 밀도)이며, 그리고  $\alpha$ 는 가중치 팩터이다. 상기의 표현식은, 선형적 SNR  $\xi_n$ 의 2개의 추정치들: ( $\gamma - 1 = (|Y|^2 / \sigma^2) - 1 = (|Y|^2 - \sigma^2) / \sigma^2 \sim \xi$ 이기에) 1) 재귀 부분

$$\frac{\hat{A}_{n-1}^2}{\hat{\sigma}_n^2}$$

[0151] (일반적으로,  $\hat{A}_n$ 은  $\xi_n$ 에 의존) 및 2) 비-재귀 부분  $\max(0, \gamma_n - 1)$ 의 선형적인 조합이다. 가중치 파라미터  $\alpha$ 는 통상적으로 간격 0.94 내지 0.99에서 선택되지만, 명백하게,  $\alpha$ 는 프레임 속도 및 가능한 다른 파라미터들에 의존할 수 있다. 노이즈 추정치  $\hat{\sigma}_n^2$ 는, 예를 들어, 음성 활성화도 검출기 및 레벨 추정자를 사용하여(어떤 음성도 검출되지 않을 때 노이즈 레벨을 추정; 주파수 하위-대역들에서 작업), 스펙트럼 노이즈 추정자, 예를 들어, 노이즈 트랙커(예를 들어, [2] EP2701145A1 [3])로부터 이용가능하다고 가정된다. 스피치 크기 추정치  $\hat{A}_n$ 는, 스피치 추정자를 사용하여 획득되고, 이중 몇몇은 이용가능하다. 일반적으로, 스피치 추정자는 대응하는 이득 함수 G에 의해 표시될 수 있다.

$$\hat{A}_n = G(\xi_n, \gamma_n) |Y_n| \quad (2)$$

[0153] 이득 함수는, 스피치 및 노이즈 프로세스들에 관하여, 최소화될 비용 함수 또는 목적, 그리고 통계적 과정에 따라 선택될 수 있다. 잘 알려진 예들은, STSA 이득 함수 [1], LSA [4], MOSIE [5], 위너, 및 스펙트럼 감산 이득 함수들 [5], [7]이다. STSA(STSA = 최소-평균 제곱 오차 단기간 스펙트럼 진폭 추정자), LSA 및 MOSIE는 (추정된) 선형적 SNR  $\xi_n$  및 후험적 SNR,

$$\gamma_n = \frac{|Y_n|^2}{\hat{\sigma}_n^2} \quad \text{모두에 의존할 수 있는 반면에,}$$

[0155] 위너 및 스펙트럼 감산 이득 함수는 1차원이며 그리고  $\xi_n$ 에만 의존한다. [5]에 기술된 것처럼,  $\hat{A}_n$ 는 MOSIE 추정자로서 알려진 다음의 방정식을 사용하여 추정될 수 있다:

$$[0156] \hat{A}_n = \sqrt{\frac{\xi_n}{\mu + \xi_n}} \left[ \frac{\Gamma\left(\mu + \frac{\beta}{2}\right) \Phi\left(1 - \mu - \frac{\beta}{2}, 1; -v_n\right)}{\Gamma(\mu) \Phi(1 - \mu, 1; -v_n)} \right]^{1/\beta} \sqrt{\hat{\sigma}_n^2}, \quad (3)$$

[0157] 여기서,  $\Gamma(\cdot)$ 는 감마 함수이고,  $\Phi(a, b; x)$ 는 합류 초기하 함수(confluent hypergeometric function)이고 그리고

$$[0158] v_n = \frac{\xi_n}{\mu + \xi_n} \gamma_n \text{ 이다.}$$

[0159] (2)와 (3)을 결합하면,

$$[0160] G(\xi_n, \gamma_n) = \sqrt{\frac{\xi_n}{\mu + \xi_n}} \left[ \frac{\Gamma\left(\mu + \frac{\beta}{2}\right) \Phi\left(1 - \mu - \frac{\beta}{2}, 1; -v_n\right)}{\Gamma(\mu) \Phi(1 - \mu, 1; -v_n)} \right]^{1/\beta} \frac{1}{\sqrt{\gamma_n}} \text{ 으로 쓸 수 있다.}$$

[0161] LSA 추정자(예를 들어, [4] 참조)는  $\beta = 0.001$  및  $\mu = 1$ 를 가지면 잘 접근될 수 있다(예를 들어, [5] 참조). 결정-지향성 접근법에 의해 추정된 선형적 SNR은, 따라서,  $\hat{A}_n$ 을 얻기 위한 선택된 추정자와 함께 평활화 인자  $\alpha$ 에 의존하는 최대( $0, \gamma_n - 1$ )의 평활화된 버전이다.

[0162] 상기에서 언급한 것처럼,  $\alpha$ 는 프레임 속도에 따라 달라질 수 있다. 일 실시예에서, [1]에서 원래 제안된 것 같은 결정 지향성 접근법은 매 8ms 마다 시프트된 프레임들로 설계된다. 보청기들에서, 프레임들은 통상적으로 매우 높은 프레임 속도(예를 들어, 모든 밀리 초)로 업데이트된다. 필터 बैं크의 이러한 높은 오버샘플링 팩터는, (예를 들어, 스피치 온셋들을 더 잘 유지하기 위해) 시스템으로 하여금 훨씬 빠르게 반응하게 한다. 가능한 더 빠른 반응 시간의 이러한 이점은, 더 높은 프레임 속도에 따라  $\alpha$ 를 단지 조정함으로써 완전히 달성될 수 없다. 대신, 더 높은 오버샘플링 팩터를 이용하는 것이 더 나은 방법을 제안한다.

[0163] DD-알고리즘(1)은 재귀 함수로서 재공식화될 수 있다.

$$[0164] \xi_n = f(\xi_{n-1}, \gamma_{n-1}, \gamma_n) = \alpha G\left(\xi_{n-1}, \frac{|Y_{n-1}|^2}{\hat{\sigma}_{n-1}^2}\right)^2 \frac{|Y_{n-1}|^2}{\hat{\sigma}_n^2} + (1 - \alpha) \max\left(0, \frac{|Y_n|^2}{\hat{\sigma}_n^2} - 1\right) \quad (4)$$

[0165] 첫번째 단순화로서, 약간 수정된 알고리즘을 고려하는데, DD\*라고 언급할 것이다. DD\*에서의 재귀는, 현재의 프레임 관측들 및 이전의 선형적 추정치에 의해서만 변경된다.

$$[0166] \xi_n = F(\xi_{n-1}, \gamma_n) = \alpha G\left(\xi_{n-1}, \frac{|Y_n|^2}{\hat{\sigma}_n^2}\right)^2 \frac{|Y_n|^2}{\hat{\sigma}_n^2} + (1 - \alpha) \max\left(0, \frac{|Y_n|^2}{\hat{\sigma}_n^2} - 1\right) \quad (5)$$

[0167] 이러한 수정에 의한 선형적 추정치들에 대한 영향은 수치적 시뮬레이션(이후의 부분들 참조)에 의해 정량화될 수 있는데, 여기서 이러한 영향은 일반적으로 작지만, 들을 수 있는 것으로 발견된다. 사실, 이득 함수에서 후험적 SNR에 대한 가장 최근의 프레임 전력을 사용하는 것은, 스피치 온셋들에서 SNR 추정에 유리하게 보인다.

[0168] 이제, SNR 값을 가장 높은 우도로 표현하는 최대 우도 SNR 추정자를 고려하여 보자; 노이즈 및 스피치 프로세스들은 상관없는 가우시안 프로세스들이고, 그리고 스펙트럼 계수들은 시간 및 주파수에 대해 독립적이라는 표준 가정을 여기에 제시한다[1]. 그때, 최대 우도 SNR 추정자  $\xi_n^{ML}$ 는 다음과 같이 주어진다.

$$[0169] \xi_n^{ML} = \max\left(\xi_{min}^{ML}, \frac{|Y_n|^2}{\hat{\sigma}_n^2} - 1\right). \quad (6)$$

[0170] 최대 우도 추정자는, 그 평균값이 실제 값과 다르므로 중앙 추정자가 아님을 알 수 있다. 이 경우, 중앙 추정자의 예는,

[0171]  $\frac{|Y_n|^2}{\sigma_n^2} - 1$  이고, 이는 음의 값들을 가질 수 있다.

[0172] 도 2는 참(true) SNR의 함수로서 최대 우도 추정자  $\xi_n^{ML}$ 의 평균값을 [dB]로 나타내고, 최대 우도 선형적 SNR 추정  $\xi_n^{ML} = \max(\xi_{min}^{ML}, \gamma_n - 1)$ 에서 단방향 정류에 의해 도입된 바이어스를 도시한다. 타겟 신호는 가정된 가우시안이다. 노이즈 전용 입력에 대해, 가정된 SNR은  $\xi_n^{ML} = e^{-1} \approx -4.3$  dB과 동일하다( $\xi_{min}^{ML} = 0$ 으로 가정, 또한 [5] 참조), 도 1에서 바이어스 참조. DD 접근법의 한 가지 효과는 이 바이어스에 대한 보상을 제공하는 것이다.

[0174] 입력-출력 관계

[0175] 다음에서, 식 (5)의 DD\* 알고리즘의 기능적 근사가 제안된다. 수학적 편의를 위해 다음과 같이 가정하며,

[0176]  $\frac{|Y_n|^2}{\sigma_n^2} \geq 1$ ,

[0177] 그리고 그와 같은 근사를 도출한다. 이 가정은,  $\xi_n = \max(0, \gamma_n - 1)$ 이  $\xi_n = \gamma_n - 1$  및  $\gamma_n = \xi_n + 1$ 으로 단순화되기 때문에, 비-재귀 부분을 단순화한다. 이 가정의 (결과에 대한) 영향은 실제로 미미하다는 것을 알 수 있다. 따라서, 사례들을 무시하고,

[0178]  $\frac{|Y_n|^2}{\sigma_n^2} < 1$  인 경우,

[0179] 식 (5)의 DD\* 알고리즘은 다음의  $\xi_n^{ML}$ 의 함수로서 서술될 수 있다

[0180] 
$$\xi_n = \Psi \left( \xi_{n-1}, \frac{\xi_n^{ML}}{\xi_{n-1}} \right) \xi_{n-1}. \quad (7)$$

[0181] 
$$\frac{\xi_n}{\xi_{n-1}} = \Psi \left( \xi_{n-1}, \frac{\xi_n^{ML}}{\xi_{n-1}} \right)$$
 이기에,

[0182] 함수  $\Psi$ 는, 현재  $\xi_n^{ML}$ 와 이전의 선형적 SNR 추정치  $\xi_{n-1}$  사이의 비(ratio)의 함수로서 선형적 추정치에서의 상대적인 변화를 설정한다. 따라서,

[0183] 
$$\Psi \left( \xi_{n-1}, \frac{\xi_n^{ML}}{\xi_{n-1}} \right) = \alpha G(\xi_{n-1}, \xi_n^{ML} + 1)^2 (\xi_n^{ML} + 1) + (1 - \alpha) \max(0, \xi_n^{ML}). \quad (8)$$

의 관계를 갖는다.

[0184] SNR 비들을 대수(dB) 스케일로 나타냄으로써, 위의 관계는 DD\*-알고리즘에 의해 표현된 비-선형 입력-출력 관계를 나타낸다.

[0185] 도 3은 STSA [1] 이득 함수( $\alpha = 0.98$ )에 대한 식 (7)의 수치 평가에 의해 DD\*-알고리즘의 입력-출력 관계 ( $\Delta_{출력} = f(\Delta_{입력})$ )를 도시한다. 선형적 SNR 추정치들이 낮을 때(예를 들어, -30dB라고 표시된 곡선), 중간 입력에 대한 출력의 작은 변화들이 있기에 평활화가 효과적이다. 더욱이, 비-제로 횡단 제로 크로싱(zero crossing)에 의해 보여지는 편향이 도입되어 평균의 추정된 선형적 SNR이 평균 최대 우도 SNR 추정치보다 낮은 결과를 초래한다. 용어 "바이어스(bias)"는 예측값  $E()$ 와 "참(true)" 참조 값 간의 차이를 반영하기 위해 종종 사용되지만,

상기 용어는 여기에서 예측값  $E(\xi_n^{ml})$  과  $E(\xi_n)$  사이의 차이를 반영하기 위해 사용된다. 도 3은, 현재의 최대-우도 추정치  $\xi_n^{ML}$  와 이전의 선형적 SNR 추정치  $\zeta_{n-1}$ (그리고 이전의 선형적 SNR 추정치  $\zeta_{n-1}$ (입력) 사이의 차이(또는 비율)의 지식으로부터 현재의 선형적 SNR 추정치  $\zeta_n$  와 이전의 선형적 SNR 추정치  $\zeta_{n-1}$ (출력) 사이의 차이(또는 비율)의 결정을 가능하게 하는 그래픽 관계를 제공한다.

[0186] 도 3은, 이러한 관계식이 2개의 눈에 띄는 효과를 보여준다. 낮은 선형적 SNR 값들(예를 들어,  $\zeta_{n-1}=-30$  dB라고 표시된 곡선)에 대해, 출력 변화들은 입력 변화들보다 작아서, 최대 우도 SNR 추정치  $\xi_n^{ML}$  의 저역-통과 필터링/평활화를 효율적으로 구현한다. 높은 선형적 SNR 값들( $\zeta_{n-1}=+30$  dB)에 대해, DD\* 선형적 SNR 추정치  $\xi_n$ 는  $\xi_n^{ML}$ 에서의 변화만큼 변화하여, 매우 작은 양의 평활화를 초래한다. 두번째로, 낮은 선형적 SNR 값들에 대한 곡선들의 제로 크로싱들은, 최대 10dB까지

[0187] 
$$\frac{\xi_n^{ML}}{\xi_{n-1}}$$

[0188] 의 양의 dB 값들로 시프트된다. 이는, 낮은 SNR 영역들에 대해, 선형적 SNR 추정치  $\xi_n$ 가  $\xi_n^{ML}$ 의 평균값보다 약 10dB 낮은 값으로 설정되어야 함을 의미한다.

[0189] 도 3은 현재의 최대-우도 추정치  $\xi_n^{ML}$  와 이전의 선형적 SNR 추정치  $\zeta_{n-1}$ (그리고 이전의 선형적 SNR 추정치  $\zeta_{n-1}$ (입력) 사이의 차이(또는 비율)의 지식으로부터 현재의 선형적 SNR 추정치  $\zeta_n$  와 이전의 선형적 SNR 추정치  $\zeta_{n-1}$ (출력) 사이의 차이(또는 비율)의 결정을 가능하게 하는 그래픽 관계를 제공한다.

[0190] 아래에서 논의될 평활화 파라미터( $\lambda_{DD}$ ) 및 바이어스 파라미터( $\rho$ )의 값들은, 선형적 SNR  $\zeta_{n-1}=-30$  dB,  $\zeta_{n-1}= 0$  dB 및  $\zeta_{n-1}=+30$  dB에 관련된 그래프들에 대해 도 3에 도시된 그래프들로부터 관측될 수 있다. 바이어스 파라미터  $\rho$ 는 가로축과 함께 그래프의 제로 크로싱으로 발견된다. 평활화 파라미터  $\lambda_{DD}$ 는 제로 크로싱에서 해당 그래프의  $\alpha(\cdot)$ 로서 표시된 기울기로서 발견된다. 이러한 값들은, 예를 들어, 추출되고 그리고 선형적 SNR의 관련 값들에 대한 표에 저장된다. 예를 들어, 도 4의 맵핑 유닛 MAP 참조.

[0191] 도 4는 변환 유닛 Po2Pr에서 구현되는 제안된 지향성 바이어스 및 평활화 알고리즘(DBSA)의 예시적인 구현의 다이어그램을 도시한다.

[0193] 지향성 바이어스 및 평활화 알고리즘(DBSA)

[0194] 도 4는, 제안된 지향성 바이어스 및 평활화 알고리즘(DBSA, 유닛 Po2Pr에서 구현됨)의 다이어그램을 도시하고, 이의 목적은 DD 접근법의 구성가능한 대안의 구현을 제공하는 것으로, DD의 3개의 주요 효과들을 포함한다

- [0195] 1. SNR-의존 평활화로서, 음악적인 노이즈를 감소시켜 낮은 SNR 조건들에서 더 많은 평활화를 가능하게 한다.
- [0196] 2. 낮은 SNR 조건들에 대해  $\xi_n^{ML}$  와 비교된 음(negative)의 바이어스로서, 노이즈 전용 기간들에서 음악적인 노이즈의 청취를 감소시킨다.
- [0197] 3. 재귀 바이어스로서, 저-고 및 고-저 SNR 조건들로부터 빠른 스위칭을 가능하게 한다.

[0198] DBSA 알고리즘은 dB 도메인에서 SNR 추정치들로 동작하고; 따라서,

[0199] 
$$s_n^{ml} = 10\log_{10}(\xi_n^{ml}),$$

[0200] 및

[0201]  $s_n = 10\log_{10}(\xi_n)$  을 도입한다.

[0202] 제안된 알고리즘의 실시예의 중심부분은 단위 DC-이득 및 적응 시간 상수를 갖는 1차 IIR 저역 통과 필터이다. 2개의 함수들  $\hat{\lambda}(s_n)$  및  $\hat{\rho}(s_n)$ 은, 추정된 SNR의 재귀 함수로서, 평활화의 양 및 SNR 바이어스의 양을 제어한다.

[0203] 다음에서, 위에서 기술된 DD 시스템의 입력-출력 관계를 모방하기 위해 제어 기능들을 유도할 것이다.  $s_n$  및  $s_n^{ML}$ 을 선형적 및 최대 우도 SNR을 dB로 표현하도록 하고, 최대 동작(현재에서  $\kappa \rightarrow \infty$ 로 두기)을 무시하면, 입력-출력 관계는

[0204] 
$$s_n - s_{n-1} = (s_n^{ML} + \rho(s_{n-1}) - s_{n-1})\lambda(s_{n-1}) \quad (9)$$
 로 정의된다.

[0205] 따라서, DBSA를 DD\* 접근법과 동일시하는 것은 근사치와 동일하다.

[0206] 
$$s_n - s_{n-1} = (s_n^{ML} + \rho(s_{n-1}) - s_{n-1})\lambda(s_{n-1}) \approx 10\log_{10}\Psi\left(\xi_{n-1}, \frac{\xi_n^{ML}}{\xi_{n-1}}\right) \quad (10)$$

[0207] (10)에서 DBSA를 완전히 지정하기 위해서는, 바이어스 함수  $\hat{\rho}(s_n)$  및 평활화 함수  $\hat{\lambda}(s_n)$ 가 특정되어야 한다. 우리의 목표는 DD\* 접근법의 행위를 모방하는 것이기에, 예를 들어, 함수

[0208] 
$$10\log_{10}\Psi\left(\xi_{n-1}, \frac{\xi_n^{ML}}{\xi_{n-1}}\right)$$

[0209] ( $\xi_n^{ML}$ 의 함수로서 평가됨)의 제로-크로싱 위치 및 이러한 위치에서의 기울기를 측정할 수 있고, 그리고 동일한 값들을 갖도록 함수들  $\hat{\rho}(s_n)$  및  $\hat{\lambda}(s_n)$ 를 선택할 수 있다. 따라서, 바이어스 함수  $\hat{\rho}(s_n)$ 에 대해,

[0210] 
$$10\log_{10}\Psi\left(\xi_{n-1}, \frac{\xi}{\xi_{n-1}}\right) = 0\text{dB}$$
 를 만족시키는  $\xi$ 를 갖는

[0211] 
$$10\log_{10}\frac{\xi}{\xi_{n-1}}$$
의 값과 동일한 것을 선택한다.

[0212] 마찬가지로, 평활화 함수  $\hat{\lambda}(s_{n-1})$ 는 0 dB 크로싱의 위치에서(즉,  $s_n^{ML} - s_n = -\rho(s_{n-1})$ 일 때) 도 3의 곡선들의 ( $s_n^{ML}$ 에 대한) 기울기와 동일하게 설정될 수 있다.

[0213] 도 4는 DD-접근법의 대안인 지향성 바이어스 및 평활화 알고리즘(DBSA)의 구현을 도시한다. 도 4의 우측 상부에 있는 점선 상자는, 단위-이득 및 가변 평활 계수  $\hat{\lambda}$ (도 4에서는  $\hat{\lambda}_{n-1}$ )를 갖는 1차 IIR 저역 통과 필터를 나타낸다. 이 부분은, 1차 IIR 저역-통과 필터에 입력들을 제공하는 조합 유닛 '+'(신호  $s_n^{ml}$ - $\rho_{n-1}$  제공) 및 맵핑 유닛 MAP(평활 및 바이어스 파라미터들  $\hat{\lambda}$ ,  $\hat{\rho}$ 를 각각 제공)과 함께, 아래의 식 (10)을 구현한다(도 4의 표시 '식 (10)으로부터'를 참조). 2개의 맵핑 함수들  $\hat{\lambda}(s)$  및  $\hat{\rho}(s)$ (맵핑 유닛 MAP 참조)은, 추정된 선형적 SNR(도 4의  $s_{n-1}$  ( $\zeta_{n-1}$ ))의 재귀 함수로서, 평활화( $\hat{\lambda}$ ) 및 바이어스( $\hat{\rho}$ )의 양을 제어한다. n번째 시간 프레임의 선형적 신호 대 잡음비의 최대 우도 값  $\zeta_n^{ml}$ 을 제공하는 도 4의 좌측 부분은, 상기의 식 (6)을 구현한다(도 4의 표시 '식 (6)으로부터' 참조). 선형적 신호 대 잡음비의 최대 우도 값  $\zeta_n^{ml}$ 은 'dB' 단위로 대수 영역으로 변환된다. 맵핑 유닛 MAP는, 예를 들어, 선형적 SNR  $\zeta$ 의 관련된 값들에 대해(예를 들어,  $\zeta$ 의 더 큰 범위에 대해 및/또는 더 큰 수의 값들에 대해, 예를 들어, 매 5dB에 대해 하나의 커브, 또는 매 dB에 대해 하나의 커브), 도 3으로부터 추출된 평활화 및 바이어스 파라미터들  $\hat{\lambda}$  및  $\hat{\rho}$ 의 값들(또는 균등한 데이터 자료)(도 4의 표시 '도 3으로부터')을 갖는 룩업 테이블을 포함하는 메모리로서 구현된다. 맵핑 유닛(MAP)의 메모리에 저장하기 위한 관련된 평활화

및 바이어스 파라미터들( $\lambda$  및  $\rho$ )의 (오프-라인) 계산에 대한 알고리즘의 구현이 도 5에서 도시된다. 도 4의 실시예는, 유닛 BPS에 의해 구현되는 선형적 SNR의 현재 최대 우도값  $s_n^{ml}(\zeta_n^{ml})$ 의 더 큰 값들에 대한 바이패스 브랜치를 포함한다. 바이패스 유닛(BPS)은, 조합 유닛 '+' 및 최대 조작 유닛 'max'를 포함한다. 조합 유닛 '+'은 입력으로서 파라미터  $K$ 를 갖는다.  $K$ 의 값은 현재 최대 우도 값  $s_n^{ml}$ 으로부터 감소되고 그리고 결과 값  $s_n^{ml} - K$ 는 선형적 SNR의 이전 값  $s_{n-1}$ 과 함께 최대 단위 max로 공급된다. 따라서, 선형적 SNR의 현재 최대 우도 값  $s_n^{ml}(\zeta_n^{ml})$ 의 상대적으로 더 큰 값들( $s_{n-1} + K$ 보다 큼)은, 맵핑 유닛 MAP에 대한 입력에 즉시 영향을 줄 수 있다. 일 실시예에서, 파라미터  $K$ 는 주파수 종속적이다(즉, 예를 들어, 상이한 주파수 채널들  $K$ 에 대해 상이하다).

[0214] 도 5는 바이어스 파라미터  $\rho$  및 평활화 파라미터  $\lambda$ 가 결정 지향성 접근법(식 5 참조)의 파라미터들로부터 유도될 수 있는 방법을 도시한다. 도 5는 도 4의 맵핑 유닛(MAP)에 관련 데이터를 생성하기 위한 알고리즘의 일 실시예를 도시한다. 상기 알고리즘은, 선형적 SNR의 현재 최대 우도 값  $s_n^{ml}(\zeta_n^{ml})$  및 이전의 선형적 SNR 값  $s_{n-1}$ 으로부터 바이어스 파라미터  $\rho$  및 평활화 파라미터  $\lambda$ 를 결정한다.  $\rho$  및  $\lambda$ 의 단일 맵핑을 갖는 것과는 달리, 입력이 증가 또는 감소하는지 여부에 따라  $\rho$  및  $\lambda$ 의 서로 다른 세트들을 갖도록 선택할 수 있다. 이는  $\rho$  및  $\lambda$ 에 대한 상이한 공격 및 릴리스 값을 갖는 것과 대응한다. 그와 같은 파라미터들의 집합은 서로 다른 공격 및 릴리스 시간들에 대응하는  $\alpha$ 의 서로 다른 값들로부터 도출될 수 있다(그리고 이후 맵핑 유닛(MAP)에 저장될 수 있다). 이후에 서술된 바와 같이, LSA 접근법[4]에서 사용되는 것과 상이한 프레임 속도(또는 프레임 길이)를 고려한 평활화 파라미터의 보상은 바람직하게 구현된다(그 결과 맵핑 유닛에 저장된 평활화 파라미터  $\lambda$ 의 값들은 직접 적용가능하다). 이것은, 예를 들어, 도 8에 관련하여 아래에서 더 논의된다.

[0215] 도 6a는 기울기  $\lambda$ 를 도시하고, 그리고 도 6b는, 두 가지의 경우들에서  $\alpha = 0.98$ 을 사용하여, STSA 이득 함수 [1]에 대한 함수

[0216] 
$$10\log_{10}\Psi\left(\xi_{n-1}, \frac{\xi_n^{ML}}{\xi_{n-1}}\right)$$

[0217] 의 제로 크로싱  $\rho$ 을 도시한다. 도 7은, 도 6a 및 도 6b의 피팅 함수를 사용하여, 본 발명에 따른 DBSA 알고리즘(크로스들) 및 DD-알고리즘(라인들)의 응답들의 비교를 도시하고, 여기서, 곡선은 -30dB 내지 +30dB 범위의 선형적 SNR 값을 5dB 간격으로 나타낸다.

[0218] 도 6은 수치적 평가로부터의 결과들을 도시하고, 그리고 도 7은 DD\*-알고리즘 및 DBSA 알고리즘의 입력-출력 응답들 사이의 비교를 도시한다. 이 차이는, 이후 섹션의 시뮬레이션에서 볼 수 있듯이 대부분의 경우들에서, 매우 작게 나타난다.

[0220] 낮은 관측된 SNR의 경우

[0221] 이제,  $\frac{|Y_n|^2}{\hat{\sigma}_n^2} < 1$ 인 경우를 고려해보자. DBSA에서, 이 경우는 영향을 제한하는 최소 값  $\xi_{min}^{ML}$ 에 의해 획득된다.

식 (2),  $\hat{A}_n = G\left(\xi_n, \frac{|Y_n|^2}{\hat{\sigma}_n^2}\right) |Y_n|$ 를 다시 생각해보면, 위너 이득 함수의 거듭제곱으로서 표현될 수 있는 이득 함수들의 클래스는 일반적으로  $\frac{|Y_n|^2}{\hat{\sigma}_n^2} \rightarrow 0$ 일 때,  $\hat{A}_n \rightarrow 0$ 을 가질 수 있다. 이 속성은, DD-알고리즘 바이어스를 상당히 크고 그리고 음으로 만들며, 이는  $\xi_{min}^{ML}$ 의 상대적으로 낮은 값을 갖도록 DBSA에서 모방될 수 있다.

[0222] 반면에, STSA, LSA 및 MOSIE 이득 함수들에 대해, 제한에서 비-제로  $\hat{A}_n$ 에서 발생하는

[0223]  $\frac{|Y_n|^2}{\hat{\sigma}_n^2} \rightarrow 0$  일 때,

[0224] 0dB보다 큰 이득이 발생한다. 이 효과는 어느 정도 더 큰  $\xi_{\min}^{ML}$ 에 의해 처리될 수 있다. 실제로, DD\* 접근법과 DBSA 간의 나머지 차이는 무시할 수 있도록 만들 수 있다.

[0226] 수치적 문제들

[0227] 일부 경우들에서(통상적으로, 낮은 선형적인 SNR 값들에 대해서), 함수

[0228] 
$$10 \log_{10} \Psi \left( \xi_{n-1}, \frac{\xi_n^{ML}}{\xi_{n-1}} \right)$$

[0229] 가 제로 크로싱을 갖지 않음을 알 수 있다. 이는, 시스템이 생성할 수 있는 실제적인 선형적 SNR 값들의 범위에서의 한계를 반영한다. 하나의 특정 예는, 이득 함수

[0230] 
$$G(\xi_n, \frac{|Y_n|^2}{\hat{\sigma}_n^2})$$

[0231] 가 일부 최소 이득 값  $G_{\min}$ 에 의해 제한될 때 발생한다. 식 (5)에 이 최소값을 삽입하면,

[0231] 
$$\Psi \left( \xi_{n-1}, \frac{\xi_n^{ML}}{\xi_{n-1}} \right) \geq \frac{\alpha G_{\min}}{\xi_{n-1}}$$
 인 것을 용이하게 알 수 있다.

[0232] 따라서,  $\xi_{n-1}$  가 충분히 낮으면, 함수  $\Psi$ 는 1보다 더 클 것이고, 이는 다시 함수  $10 \log_{10} \Psi$ 에 대한 제로 크로싱이 전혀 존재하지 않음을 의미한다. 수치적인 구현은 이러한 상황을 검출할 필요가 있고 그리고  $\rho_{(s_n)}$  및  $\lambda_{(s_n)}$ 에 대한 일부 타당한 룩업 테이블 값들을 동일하게 특정할 필요가 있다. 사용된 정확한 값들은, 초기 상태에서 수렴하는 동안 샘플링될 가능성이 높기 때문에 실제로는 중요하지 않을 것이다.

[0234] 최대 연산자 및 기타

[0235] 도 4에서, 최대 연산자는 재귀 루프에 위치하며, 최대 우도 SNR 추정치가 바이어스 및 평활화 파라미터의 계산에서(파라미터  $\mathbf{K}$ 를 통해) 이전의 프레임의 선형적인 추정치를 바이패스하는 것을 가능하게 한다. 이러한 요소에 대한 이유는, SNR 온셋들의 탐지를 돕고 그리고 스피치 온셋들의 과도한 감쇠 위험을 줄이는 것이다. DD 접근법 식(1)에서, 항  $(1 - \alpha)$ 는 현재 프레임의 큰 온셋이 음의 바이어스를 빨리 줄이는 것을 가능하게 하며, 최대값은 파라미터  $\mathbf{K}$ 에 의해 제어되는 것처럼 이러한 동작을 모방한다. 따라서, 팩터  $\mathbf{K}$ 를 사용하여 평활화를 우회할 수 있다.  $\mathbf{K}$ 를 증가시키으로써, 스피치 온셋들을 더 잘 유지할 수 있다. 반면에, 증가된  $\mathbf{K}$ 는 또한 노이즈 플로어를 증가시킬 수 있다. 하지만, 증가된 노이즈를 갖는 노이즈 플로어는 높은 감쇠량을 적용할 때만 영향을 미칠 것이다. 따라서, 선택된  $\mathbf{K}$ 의 값은 선택된 최대 감쇠에 의존한다.

[0236] 최대 연산자(도 4, 5 및 8에서의 'max') 대신에, 보다 일반적인 선택 방식이 (갑작스런) SNR-변화들(예를 들어, 온셋들)을 식별하기 위해 사용될 수 있다. 예를 들어, 도 12a, 도 12b 및 도 12c에서 실시예들에서의 '선택' 유닛 참조. 그와 같은 일반적인 방식들은, 예를 들어, 음향 환경에서의 사건들(변화들)(예를 들어, 노이즈 소스들(예를 들어, 바람 소리)의 급격한 출현 또는 제거, 또는 스피치 소스들, 예를 들어, 자신의 목소리와 같은 다른 음향 소스들에서 급격한 변화들)의 고려를 포함할 수 있고(도 13a 참조), 그리고/또는 고려된 주파수 대역들 주변의 다수의 주파수 대역들에 걸친 신호 변화들의 고려(예를 들어, 모든 주파수 대역을 평가하고 그리고 해당 주파수 대역에 대한 결과적인 온셋 플래그를 제공하기 위한 로직 기준을 적용)를 포함할 수 있다(도 13b 참조).

[0238] 필터 뱅크 오버샘플링

- [0239] 필터 뱅크 파라미터들은 DD 접근법의 결과에 큰 영향을 미친다. 오버샘플링은, 선형적 SNR 추정치에 도입되는 평활화의 영향 및 바이어스의 양에 직접적인 영향을 미치기에, 고려해야할 주요 파라미터이다.
- [0240] DD 접근법에서 필터 뱅크 오버샘플링을 교정하는 방법은 문헌에서 상세하게 설명되지 않는다. 원래의 공식 [1]에서, 256 포인트 FFT가, 4배 오버샘플링에 대응하는 192 샘플들 오버랩 및 8kHz의 샘플 속도를 갖는, 해닝 윈도우로 사용된다. 일반적으로, 2배 오버샘플링(50% 프레임 오버랩)이 일반적이다. 식 [1] 및 그에 대한 참조들을 참고. 하지만, 보정기들 및 다른 저-지연(low-latency) 애플리케이션들은, 16배 이상의 오버 샘플링은 비현실적이지 않다.
- [0241] 모든 점이 동일하다면, 오버샘플링은, DBSA 방법과 함께 DD-접근법의 재귀적 효과들을 감소시킨다. "무한(infinite)" 오버샘플링의 한계에서, 재귀 바이어스는, 점근적 편향 함수로 대체된다.
- [0242] 오버샘플링 보상에 대한 하나의 가능한 접근법은, 오버샘플링에 비례하는 인자에 의해 DD/DBSA 추정을 다운샘플링하여, 다수의 프레임들에 대해 선형적 추정치를 일정하게 유지하는 것이다. 이 접근법의 단점은, 이득 점프들이 도입되어, 오버샘플링된 필터 뱅크와 함께 사용될 때, 음질이 저하될 수 있다는 것이다. 오버샘플링을 사용하면, 동등한 합성 필터들이 더 짧아지고, 그리고 이득 점프들에 의해 도입되는 컴벌루션 잡음의 감소에 충분하지 않을 수 있다.
- [0243] DBSA 방법을 사용하면, 시간적 거동(즉, SNR 추정치들의 평활화 및 온셋들에 대한 응답성들)은 지향성 재귀 평활화, 지향성 재귀 바이어스의 조합에 의해 제어된다. 보다 계산적으로 요구되지만 이론상으로는 필터 뱅크 오버샘플링을 다루는 더 정밀한 방법은, 도 8에서 도시된 것처럼 재귀 루프 내의 고차 지연 소자(순환 버퍼)에 의한 것이다.
- [0244] 도 8는 필터 뱅크 오버샘플링을 수용하기 위한 DBSA 알고리즘(도 4에 도시됨)의 변형을 도시하며, 재귀 루프에서 추가적인 D-프레임 지연을 삽입하는 목적은오버샘플링 적은 시스템의 동적 거동을 모방하는 것이다.
- [0245] 도 4, 5 및 8에 예시된 DBSA 알고리즘의 실시예와 비교하면, 도 12a, 12b 및 12c에 도시된 실시예들은, 예를 들어, 온셋 플래그(Onset flag)에 의해 제어될 수 있는 선택 연산자(선택)에 의해 대체된다는 점에서 다르다. 로컬 주파수 채널  $k$ 에만 영향을 미치는 최대 연산자와는 달리, 온셋 플래그는, 예를 들어, 미리정의된 또는 적응형(예를 들어, 로직) 기법(예를 들어, 도 1a 참조) 및/또는 다른 주파수 채널들을 함께 포함하는 것(예를 들어, 도 13b 참조)에 따라 자격이 부여된 '제어 입력들(control inputs)'의 수에 의존할 수 있다. 일 실시예에서, 바이어스 파라미터  $K$ 는 주파수 종속적이다(즉, 예를 들어, 상이한 주파수 채널들  $k$ 에 대해 상이하다).
- [0246] 도 12a는 제안되는 지향성 바이어스 및 평활화 알고리즘(DBSA, 예를 들어, 도 1a, 1b 및 도 9b의 유닛 Po2Pr에 의해 구현됨)의 제1의 다른 예시적인 구현의 다이어그램을 도시한다. 로컬 주파수 채널  $k$ 에만 영향을 주는 최대 연산자와는 달리, 온셋 플래그는 다른 주파수 채널들에도 의존할 수 있다(예를 들어, 도 13b 참조). 온셋 플래그의 이점은 (온셋이 다수의 주파수 채널들에 동시에 영향을 미친다고 가정하면) 높은 SNR을 갖는 소수의 주파수 채널에서 검출된 온셋 정보가 낮은 SNR을 갖는 주파수 채널로 전파될 수 있다는 점이다. 이로써 온셋 정보는 저-SNR 주파수 채널들에서 빠르게 적용될 수 있다. 일 실시예에서, 광대역 온셋 검출기는 주어진 주파수 채널  $k$ 에 대한 온셋 플래그와 함께(또는 온셋 플래그를 결정하기 위한 기준에 대한 입력으로서) 사용될 수 있다. 대안으로, 예를 들어, 복수의  $K$ 개의 주파수 채널들(예를 들어, 해당  $k$ 에서의 채널 및 각 측상의 인접하는 채널들(예를 들어,  $k-1$ ,  $k+1$ , 도 13b 참조)에서의 바이어스 보정된 최신 SNR 값  $s_n^{ML} - K$ (의 최대 우도 ('선형적') 추정치가 이전의 ('선형적') SNR 값  $s_{n-1}$ 보다 높으면, 이는 온셋의 표시이다. 바로 이웃하는 채널들과 다른 주파수 채널들 및/또는 다른 온셋 표시들은 주어진 주파수 채널  $k$ 에 대한 온셋 플래그의 결정에서 고려될 수 있다. 일 실시예에서, 특정 주파수 채널  $k$ 에서의 온셋 플래그는, 로컬 온셋들이 적어도  $q$  채널들에서 검출되었는지에 따라 결정되며,  $q$ 는 1과  $K$  사이의 수이다.
- [0247] 도 12b는 제안되는 지향성 바이어스 및 평활화 알고리즘(DBSA, 예를 들어, 도 1a, 도 1b 및 도 9b의 유닛 Po2Pr에 의해 구현됨)의 제2의 다른 예시적인 구현의 다이어그램을 도시한다. SNR에 의존하는 것 이외에도,  $\hat{\lambda}$  및  $\hat{\rho}$ 는 또한, SNR이 증가 또는 감소하는지 여부에 따라 달라질 수 있다.  $s_n^{ML} + \hat{\rho}_{n-1-s_{n-1}} > 0$ 에 의해 표시된 것처럼, SNR이 증가하면,  $\hat{\lambda}(s)$  및  $\hat{\rho}(s)$ ,  $\hat{\lambda}_{atk}(s)$ ,  $\hat{\rho}_{atk}(s)$ 의 하나의 세트를 선택하고, 그리고  $s_n^{ML} + \hat{\rho}_{n-1-s_{n-1}} < 0$ 에 의해 표시된 것처럼, SNR이 감소하면,  $\hat{\lambda}(s)$  및  $\hat{\rho}(s)$ ,  $\hat{\lambda}_{rel}(s)$ ,  $\hat{\rho}_{rel}(s)$ 의 다른 세트를 선택한다. 평활화 파라미터들

$\lambda(s)$  및  $\rho(s)$ 의 예시적인 코스트들이 도 6a 및 도 6b에 각각 도시된다.

- [0248] 더욱이, 다른 바람직한 실시예들에서, "선택(select)" 유닛은 검출된 온셋에만 의존하지 않을 수 있다. 선택 유닛은, 검출된 자신의 목소리 또는 바람 소리 또는 언급된 (또는 다른) 검출기들의 어떤 조합(예를 들어, 도 13a 참조)에 의존할 수도 있다.
- [0249] 도 12c는 제안되는 지향성 바이어스 및 평활화 알고리즘(DBSA, 예를 들어, 도 1a, 도 1b 및 도 9b의 유닛 Po2Pr에 의해 구현됨)의 제3의 다른 예시적인 구현의 다이어그램을 도시한다. SNR에 의존하는 것 이외에,  $\lambda$  및  $\rho$ 는 또한, SNR이 증가 또는 감소하고 있다는 또 다른 표시에 의존할 수 있다.  $s_n^m - s_{n-1} > 0$ 에 의해 표시된 것처럼 SNR이 증가하면,  $\lambda$  및  $\rho$ ,  $\lambda_{atk}$ ,  $\rho_{atk}$ 의 하나의 세트를 선택하고, 그리고  $s_n^m - s_{n-1} < 0$ 에 의해 표시된 것처럼 SNR이 감소하면,  $\lambda$  및  $\rho$ ,  $\lambda_{rel}$ ,  $\rho_{rel}$ 의 다른 세트를 선택한다.
- [0250] 도 13a는 도 12a, 도 12b, 도 12c에 도시된 DBSA 알고리즘들의 실시예들에서 사용하기 위한 온셋 플래그를 제공하는 일반적인 예를 도시한다. 오디오 프로세싱 디바이스, 예를 들어, 보청기는 오디오 프로세싱 디바이스 주위의 음향 장면의 변화의 온셋의 다수의 표시자들(신호들  $IX_1, \dots, IX_{ND}$ )을 제공하는 다수(ND) 검출기들 또는 표시기들( $IND_1, \dots, IND_{ND}$ )을 포함할 수 있고, 오디오 프로세싱 디바이스는 상기 오디오 프로세싱 디바이스의 순방향 경로에 의해 고려되는 신호의 SNR의 변화를 초래할 수 있다. 그와 같은 표시기들은, 예를 들어, 시간 변화 입력 사운드  $s(t)$ 에서의 갑작스러운 변화들(예를 들어, 도 9a 참조), 예를 들어, 시간 변화 입력 사운드에서의 변조를 검출하는 일반적인 온셋 검출기, 바람 소리 검출기, 목소리 검출기, 예를 들어, 자신의 목소리 검출기 등 및 이들의 조합을 포함할 수 있다. 표시기들( $IND_1, \dots, IND_{ND}$ )로부터의 출력들( $IX_1, \dots, IX_{ND}$ )은 제어기(제어)에 공급되고, 이 제어기는 주어진 주파수 채널  $k$ 에 대해 발생한 온셋 표시자(신호 온셋 플래그)를 제공하는 알고리즘을 구현한다. 그와 같은 방식의 특정 구현(또는 부분적인 구현)은 도 13b에서 도시된다.
- [0251] 도 13b는 도 12a, 도 12b, 도 12c에 도시된 DBSA 알고리즘들의 실시예들에서 사용가능하기 위해 온셋 플래그를 제공하는 이웃하는 주파수 대역들로부터의 입력에 기초한 제어기(제어)의 예시적인 실시예를 도시한다. 도시된 방식은,  $s_n^m(k') - K > s_{n-1}(k')$  가 고려되는 주파수 대역  $k$  주위의 다수의 주파수 대역들  $k'$  또는 다수의 주파수 대역들 중 하나 또는 '3개중 2개'에 대해 수행되는지 여부에 의해 표시되는 SNR의 시간에 대한 변화들을 평가(예를 들어,  $k' = k-1, k$ , 및  $k+1$ 에 대한 표현을 평가)하거나, 또는 모든 주파수 대역들  $k=1, \dots, K$ (또는 선택된 범위, 예를 들어, 스피치 및/또는 노이즈가 발생할 것으로 예측되는 것)에 대한 표현을 평가하고 그리고 문제되는 주파수 대역에 대한 발생한 온셋 플래그를 제공하기 위한 로직 기준을 적용하는 표시자들을 포함하는 입력 표시자 신호들( $IX_p, \dots, IX_q$ )을 제공한다. 일 실시예에서, 주어진 채널  $k$ 에 바로 인접한 대역들만이 고려된다. 즉, 각 채널에 대해 온셋 플래그를 제공하는데 3개의 채널들이 포함된다. 일 실시예에서, 그와 같은 방식은 도 13a와 관련하여 언급된 다른 검출기들로부터의 입력과 결합된다. 일 실시예에서, 식  $s_n^m(k') - K > s_{n-1}(k')$  또는 다른 유사한 식들이 해당 채널 주위의 다수의 주파수 채널들, 예를 들어, 모든 채널들에 대해 평가되고, 그리고 발생한 온셋 플래그를 제공하는 방식은 입력 표시자들( $IX_p, \dots, IX_q$ )에 적용된다. 바이어스 상수  $k$ 는 주파수에 대해 일정할 수 있거나 또는 채널마다 상이할 수 있거나 또는 일부 채널들에 대해 상이할 수 있다.
- [0253] 제안된 구현의 이점들
- [0254] 제안된 구현은 결정 지향성 접근법에 비해 다음과 같은 이점들을 가진다:
- [0255] • 필터 뱅크 오버샘플링을 고려하기 위해 평활화 파라미터를 조정할 수 있는데, 이는 보청기와 같은 대기 시간이 짧은 애플리케이션에서 구현하는데 중요하다.
- [0256] • 선택된 이득 함수에 따라 평활화 및 바이어스를 갖는 대신에, 평활화  $\lambda(s)$  및 바이어스  $\rho(s)$ 는 2개의 맵핑 함수들의 파라미터화에 의해 집적 제어된다. 이는 노이즈 감소 및 음질 간의 원하는 트레이드 오프를 위해 맵핑 함수들 각각을 튜닝하는 것을 가능하게 한다. 예를 들어, 타겟 에너지는 바이어스를 강하게 강조함으로써 더 잘

유지될 수 있다. 또한, 파라미터들은 관심있는 SNR의 특정 범위를 처리하기 위해 설정될 수 있다. 일부 사용자들은 낮은 SNR 영역들에서 (변동 이득의 관점에서) 노이즈 감소로부터 주로 이득을 얻고, 더 높은 신호 대 잡음비들로서 노이즈 감소를 필요로하지 않기 때문에, 그와 같은 파라미터 집합은 개별 사용자들 별로 상이하게 선택될 수 있다. 반면에, 다른 사용자들은 높은 신호 대 잡음비 영역에서 노이즈 감소를 요구할 수 있고, 그리고 낮은 신호 대 잡음비들에서 일정한 감쇠를 요구할 수 있다.

[0257] • 제안된 시스템으로의 확장으로서, 평활화 및 바이어스 파라미터들은, 입력이 증가하는지 또는 감소하는지에 따라 달라질 수 있다. 즉, 2개의 파라미터들의 상이한 공격 및 릴리스 값들을 사용할 수 있다.

[0258] • 현재 프레임 관측들 및 이전의 선형적 추정에만 의존하는 결정 지향성 접근법의 변화는 스피치 온셋들에서의 SNR 추정에 대해 유익한 것으로 보인다.

[0259] • 마찬가지로, 파라미터  $K$ 에 의해 제어되는 최대 연산자는 스피치 온셋들을 과도하게 감쇠시키는 위험을 줄이기 위해 사용될 수 있다. 선택된 값은 선택한 최대 감쇠량에 따라 달라질 수 있다.

[0260] • 선택된 최소 값  $\xi_{\min}^{ML}$ 에 의한  $\xi_n^{ML}$ 의 사전-평활화는  $\frac{|Y_n|^2}{\hat{\sigma}_n^2} < 1$ 에 대처하기 위해 사용된다.

[0261] • 노이즈 추정기는 단일 채널 입력들뿐 아니라 다중 채널에 의존할 수 있거나, 또는 두가지 모두에 의존할 수 있거나 또는 바이노럴 입력(binaural input)들에 의존할 수 있다(예를 들어, 도 10 참조). BDSA 파라미터들은, 노이즈 추정기가 단일 채널 입력 또는 다중-채널 입력들에 의존하는지에 따라 상이하게 조정될 수 있다.

[0262] 도 9a는, 본 발명에 따른 오디오 프로세싱 디바이스(APD), 예를 들어, 보청기의 일 실시예를 도시한다. 시간 변화형 입력 사운드  $s(t)$ 는, 타겟 신호 컴포넌트  $x(t)$  및 노이즈 신호 컴포넌트  $v(t)$ 의 혼합을 포함하도록 추정되고, 상기  $s(t)$ 는 오디오 프로세싱 디바이스에 의해 픽업되어 처리되고 그리고 처리된 후 오디오 신호로서 사용자에게 제공된다. 도 9의 오디오 프로세싱 디바이스(여기서는, 보청기)는 다수의 입력 유닛들( $IU_j$ ,  $j=1, \dots, M$ )을 포함하고, 각 입력 유닛은 시간-주파수 표현( $k,m$ )으로서 사운드  $s(t)$ 를 나타내는 전기 입력 신호  $S_i$ 를 제공한다. 도 9a의 실시예에서, 각 입력 유닛  $IU_i$ 는 (입력 유닛 ( $IU_i$ )에서 수신된) 환경으로부터의 입력 사운드  $s_i$ 를 전기 시간-도메인 신호  $s'_i$ ,  $i=1, \dots, M$ 으로 변환하는 입력 트랜스듀서  $IT_i$ 를 포함한다. 입력 유닛  $IU_i$ 는, 전기 시간-도메인 신호  $s'_i$ 를 다수의 주파수 서브-대역 신호들( $k=1, \dots, K$ )로 변환하는 분석 필터 뱅크  $FBA_i$ 를 더 포함함으로써, 전기 입력들 신호를 시간-주파수 표현  $S_i(k,m)$ 으로 제공한다. 보청기는, 다수의 전기 입력 신호들  $S_i$ ,  $i=1, \dots, M$ 에 기초하여 노이즈 감소된 신호 YNR을 제공하는, 다중-입력 노이즈 감소 시스템(NRS)를 더 포함한다. 다중-입력 노이즈 감소 시스템(NRS)는 다중-입력 빔 형성기 필터링 유닛(BFU), 포스트 필터 유닛(PSTF) 및 제어 유닛(CONT)을 포함한다. 다중-입력 빔 형성기 필터링 유닛(BFU)(및 제어 유닛(CONT))은 다수의 전기 입력 신호들  $S_i$ ,  $i=1, \dots, M$ 를 수신하고, 그리고 신호들  $Y$  및  $N$ 을 제공한다. 제어 유닛(CONT)은, 복소 가중치들  $W_{ij}$ 이 저장되는 메모리(MEM)를 포함한다. 복소 가중치  $W_{ij}$ 는 (신호  $W_{ij}$ 를 통해 BFU에 공급되는) 빔 형성기 필터링 유닛 BFU의 가능한 미리 정의된 고정된 빔 형성기들을 정의한다(예를 들어, 도 9b 참조). 제어 유닛(CONT)은, 주어진 입력 신호(예를 들어, 입력 신호의 주어진 시간-주파수 유닛)가 음성을 포함하는지(또는 음성에 의해 지배되는지)를 추정하기 위한 하나 이상의 음성 활성 검출기들(VAD)을 더 포함한다. 각각의 제어 신호들(V-N1 및 V-N2)은 빔 형성기 필터링 유닛(BFU) 및 포스트 필터링 유닛(PSTF)에 각각 공급된다. 제어 유닛(CONT)은, 입력 유닛들( $IU_i$ )로부터 다수의 전기 입력 신호들  $S_i$ ,  $i=1, \dots, M$ , 및 빔 형성기 필터링 유닛(BFU)으로부터 신호  $Y$ 를 수신한다. 신호  $Y$ 는 타겟 신호 컴포넌트의 추정치를 포함하고, 그리고 신호  $N$ 은 노이즈 신호 성분의 추정치를 포함한다. 상기 (단일 채널) 포스트 필터링 유닛(PSTF)은, (공간적으로 필터링된) 타겟 신호 추정치  $Y$  및 (공간적으로 필터링된) 노이즈 신호 추정치  $N$ 을 수신하며, 그리고 노이즈 신호 추정치  $N$ 으로부터 추출된 노이즈의 지식에 기초하여 (추가로) 노이즈 감소된 타겟 신호 추정치 YNR을 제공한다. 보청기는, 노이즈 감소된 신호를 (추가로) 프로세싱하고 그리고 프로세싱된 신호(ES)를 제공하는 신호 프로세싱 유닛(SPU)을 더 포함한다. 신호 프로세싱 유닛(SPU)는, 예를 들어, 사용자의 청각 장애를 보상하기 위해, 노이즈 감소된 신호(YNR)의 레벨 및 주파수 중속 형성을 적용하도록 구성될 수 있다. 보청기는, 프로세싱된 주파수 서브-대역 신호(ES)를 시간-도메인 신호(es)로 변환하기 위한 합성 필터 뱅크(FBS)를 더 포함하고, 시간-도메인 신호(es)는, 사운드인 수용가능한 신호로서 자극들( $t$ )을 사용자에게 제공하는 출력 유닛(OT)에 공급된다. 도 9a의 실시예에

서, 출력 유닛은 신호로서 상기 처리된 신호(es)를 사용자에게 나타내기 위한 라우드스피커를 포함한다. 보청기의 입력 유닛으로부터 출력 유닛까지의 순방향 경로는, 여기에서, 시간-주파수 도메인(다수의 주파수 서브-대역들  $FB_k$ ,  $k=1, \dots, K$ 에서 처리됨)에서 동작된다. 다른 실시예에서, 보청기의 입력 유닛으로부터 출력 유닛으로의 순방향 경로는 시간 도메인에서 동작될 수 있다. 보청기는, 사용자 인터페이스, 및 사용자들 입력들과 검출기 입력들이 노이즈 감소 시스템(NRS), 예를 들어, 빔 형성기 필터링 유닛(BFU)에 의해 수신되도록 하는 하나 이상의 검출기들을 더 포함할 수 있다. 빔 형성기 필터링 유닛(BFU)의 적응 기능이 제공될 수 있다.

[0263] 도 9b는, 예를 들어, 도 9a(M=2인 경우)의 예시적인 오디오 프로세싱 디바이스, 예를 들어, 본 발명에 따른 보청기에서 사용을 위해, 노이즈 감소 시스템(NRS)의 일 실시예의 블록도를 도시한다. 도 9a의 노이즈 감소 시스템의 예시적인 실시예는 도 9b에서 더 상세하게 개시된다. 도 9b는 본 발명에 따른 적응형 빔 형성기 유닛(BFU)의 일 실시예를 도시한다. 상기 빔 형성기 필터링 유닛은, 제1 (무-지향성) 빔 형성기 및 제2 (타겟 상쇄) 빔 형성기(도 9b의 고정 BF 0 및 고정 BF C로 표시되고 그리고 대응하는 빔 패턴으로 상징됨)를 포함한다. 제1 및 제2 고정 빔 형성기는 제1 및 제2 전기 입력 신호( $S_1$  및  $S_2$ )의 선형 조합으로서 빔 형성 신호(O 및 C)를 제공하고, 각각의 빔 패턴을 나타내는 제1 및 제2 세트의 복소 가중 상수( $W_{o1}(k)^*$ ,  $W_{o2}(k)^*$  및  $(W_{c1}(k)^*$ ,  $W_{c2}(k)^*$ )는 메모리 유닛(MEM)에 저장된다(도 9a의 제어 유닛(CONT) 내의 메모리 유닛(MEM) 및 신호( $W_{ij}$ ) 참조). \*은 복소 결합(complex conjugation)을 나타낸다. 빔 형성기 필터링 유닛(BFU)은 적응적으로 결정된 빔 패턴을 나타내는 적응 상수( $\beta_{ada}(k)$ )를 제공하는 적응형 빔 형성기(적응형 BF, ADBF)를 더 포함한다. 빔 형성기 필터링 유닛(BFU)의 고정된 빔 형성기 및 적응형 빔 형성기를 결합함으로써, 타겟 신호(Y)의 발생된 (적응형) 추정치가  $Y=O-\beta_{ada}C$ 로서 제공된다. 빔 형성기 필터링 유닛(BFU)은, 입력 신호(여기서는, O 또는 입력 신호들  $S_i$  중 하나)가 음성 콘텐츠(예를 들어, 스피치)를 포함하는지(또는 포함할 확률이 얼마인지)를 나타내는 제어 신호(V-N1)(예를 들어, 신호 O 또는 입력 신호들  $S_i$  중 하나에 기초함)를 제공하는 음성 활성화 검출기(VAD1)를 더 포함하고, 음성 활성화 검출기(VAD1)에 의해 어떤 음성/스피치도 표시되지 않는(또는 표시될 확률이 낮은) 시간 세그먼트 동안 적응형 빔 형성기가 (여기서는, 타겟 상쇄형 빔 형성기(C)에 기초하여) 노이즈 추정치( $\langle \sigma_c^2 \rangle$ )를 갱신하도록 한다.

[0264] 빔 형성기 필터링 유닛으로부터 발생된 (공간적으로 필터링되거나 또는 빔형성된) 타겟 신호 추정치(Y)는, 따라서,

[0265] 
$$Y(k) = O(k) - \beta_{ada}(k) \cdot C(k)$$

[0266] 
$$Y(k) = (W_{o1}^* \cdot S_1 + W_{o2}^* \cdot S_2) - \beta_{ada}(k) \cdot (W_{c1}^* \cdot S_1 + W_{c2}^* \cdot S_2)$$
 으로 표현될 수 있다.

[0267] 하지만, 발생한 신호를 얻기 위해 사용된 상이한 빔 형성기들을 계산하는 것보다 각 마이크로폰 신호에 적용된 실제 발생한 가중치들을 계산하는 것이 단지 유리할 수 있다.

[0268] 도 9b의 포스트 필터링 유닛(PSTF)의 실시예는, 입력 신호들(Y)(공간적으로 필터링된 타겟 신호 추정치) 및  $\langle \sigma_c^2 \rangle$ (노이즈 전력 스펙트럼 추정치)를 수신하고, 그리고 그에 기초한 출력 신호  $Y_{BF}$ (노이즈 감소된 타겟 신호 추정치)를 제공한다. 포스트 필터링 유닛(PSTF)은, 빔 형성기 필터링 유닛으로부터 수신된 노이즈 전력 스펙트럼 추정치  $\langle \sigma_c^2 \rangle$ 를 개선하고 그리고 개선된 노이즈 전력 스펙트럼 추정치  $\langle \sigma^2 \rangle$ 를 제공하는 노이즈 감소 보정 유닛(N-COR)을 포함한다. 이러한 개선은, 공간적으로 필터링된 타겟 신호 추정치(Y)(신호 V-N2 참조)에서 무-음성 시간-주파수(no-voice time-frequency) 유닛들의 존재를 표시하기 위해 음성 활성화 검출기(VAD 2)의 사용으로부터 발생한다. 포스트 필터링 유닛(PSTF)은, 타겟 신호 전력 스펙트럼 추정치  $|Y|^2$  및 후협적 신호 대 잡음비  $\gamma = |Y|^2 / \langle \sigma^2 \rangle$ 를 각각 제공하는 크기 제공( $| \cdot |^2$ ) 및 제산( $\cdot / \cdot$ ) 처리 유닛들을, 각각 더 포함한다. 포스트 필터링 유닛(PSTF)은, 후협적 신호 대 잡음비 추정치  $\gamma$ 를, 본 발명에 따른 알고리즘을 구현하는 선협적 신호 대 잡음비 추정치  $\zeta$ 로 변환하는 변환 유닛(Po2Pr)을 더 포함한다. 포스트 필터링 유닛(PSTF)은, 선협적 신호 대 잡음비 추정치  $\zeta$ 를 공간적으로 필터링된 타겟 신호 추정치(여기서는, 승산 유닛 'X'에 의한 것)에 적용될 대응하는 이득  $G_{NR}$ 으로 변환하도록 구성되는 변환 유닛(SNR2G)를 더 포함하며, 노이즈 감소된 타겟 신호 추정치(YBF)를 발생

시켜 제공한다. 주파수 및 시간 인덱스들(k 및 n)은 단순화를 위해 도 9b에서 도시되지 않는다. 하지만, 대응하는 시간 프레임들은 프로세싱된 신호들(예를 들어,  $|Y_n|^2$ ,  $\langle \sigma_n^2 \rangle$ ,  $\gamma_n$ ,  $\zeta_n$ ,  $G_{NR,n}$ , 등)에 대해 이용가능한 것으로 추정된다.

[0269] 다중-입력 빔 형성기 필터링 유닛(BFU) 및 단일 채널 포스트 필터링 유닛(PSTF)을 포함하는 다중-입력 노이즈 감소 시스템은, 예를 들어, 본 발명에서 제안된 변경들을 갖는 [2]에서 논의된 것처럼 구현될 수 있다.

[0270] 노이즈 전력 스펙트럼  $\langle \sigma^2 \rangle$ 은 2개의 마이크로폰 빔 형성기(타겟 상쇄형 빔 형성기 C)에 기초하여 도 9b의 실시예에 존재하지만, 대신에, 단일 채널 노이즈 추정치, 예를 들어, 변조의 분석(예를 들어, 음성 활성화 검출기)에 기초할 수 있다.

[0271] 도 10은, 각각의 분석 필터 뱅크들(FBA<sub>1</sub> 및 FBA<sub>2</sub>)에 전기적으로 접속된 마이크로폰들(M<sub>1</sub> 및 M<sub>2</sub>)을 포함하고 그리고 도 1b와 관련하여 서술된 것처럼, 각각의 혼합된 전기 입력 주파수 서브-대역 신호들  $Y(n,k)_1$ ,  $Y(n,k)_2$ 을 제공하는 입력 스테이지(예를 들어, 보청기)를 도시한다. 제1 및 제2 마이크로폰 신호들에 기초한, 전기 입력 신호들  $Y(n,k)_1$ ,  $Y(n,k)_2$ 은, 예를 들어, 상기에서 도 1b에 관련하여 논의된 것처럼, 다중-입력 후협적 SNR  $\gamma_{n,m}$ (n번째 시간 프레임에 대해)을 제공하기 위해 후협적 신호 대 노이즈 계산 유닛(APSNR-M)에 다중-입력(여기서는 2)으로 공급된다. 2개의 전기 입력 신호들  $Y(n,k)_1$ ,  $Y(n,k)_2$  중 하나, 또는 제3의 상이한 전기 입력 신호(예를 들어, 빔 형성 신호 또는 제3 마이크로폰(예를 들어, 대측성(contra-lateral) 보청기의 마이크로폰 또는 개별 마이크로폰)에 기초한 신호)는, 예를 들어, 상기의 도 1a에 관련하여 논의된 것처럼, 단일-입력 후협적 SNR  $\gamma_{n,s}$ (n번째 시간 프레임에 대해)를 제공하는 후협적 신호 대 노이즈 계산 유닛(APSNR-S)에 단일-입력으로 공급된다. 2개의 후협적 SNR들  $\gamma_{n,m}$  and  $\gamma_{n,s}$ 는, 2개의 후협적 신호 대 잡음비들로부터 결합된(발생한) 후협적 신호 대 잡음비  $\gamma_{n,res}$ 의 생성을 위해 믹싱 유닛(MIX)에 공급된다. 2개의 독립적인 후협적 추정치들의 조합은 일반적으로, 상기 추정치들의 각각보다 더 양호한 추정치를 제공할 것이다. 다중채널 추정치  $\gamma_{n,m}$ 는 통상적으로 단일 채널 추정치  $\gamma_{n,s}$ 보다 더 신뢰할 수 있기 때문에, 다중채널 추정치는 단일 입력 채널 노이즈 추정치에 비해 평활화가 덜 필요할 것이다. 따라서, 다중 마이크로폰 후협적 SNR 추정치  $\gamma_{n,m}$  및 단일 마이크로폰 후협적 SNR 추정치  $\gamma_{n,s}$ 의 평활화를 위해 평활화 파라미터들  $\rho$  (바이어스),  $\lambda$  (평활화), 및  $K$  (바이어스) (FIG. 3, 4 참조)의 상이한 세트들이 요구된다. 결과적인 후협적 SNR 추정치  $\gamma_{n,res}$ 를 제공하기 위해 2개의 추정치들의 혼합은, 예를 들어, 2개의 추정치들  $\gamma_{n,m}$ ,  $\gamma_{n,s}$ 의 가중된 합으로 제공될 수 있다.

[0272] 바이노럴 보청기 시스템의 일 실시예에서, 대측부(contralateral side) 상의 후협적 SNR, 선협적 SNR, 노이즈 추정치 또는 이득은, 동측부(ipsilateral side) 상의 보청기에 전송되고 그리고 사용된다.

[0273] 동측상의 보청기로부터의 후협적 추정치외에도, 선협적 추정은 대측부 상의 보청기로부터의 후협적 추정, 선협적 추정, 노이즈 추정(또는 이득 추정)에 또한 의존할 수 있다. 다시, 개선된 선협적 SNR 추정치는 상이한 독립적인 SNR 추정치들을 결합함으로써 달성될 수 있다.

[0274] 도 11은 귀 또는 사용자의 뒤에 위치한 BTE-부분 및 사용자의 외이도에 위치한 ITE 부분을 포함하는 본 발명에 따른 보청기의 일 실시예를 도시한다.

[0275] 도 11은 귓바퀴의 뒤에 위치되도록 구성되는 BTE-부분(BTE) 및 사용자의 외이도(ear canal)에 위치되도록 구성되는 출력 트랜스듀서(예를 들어, 라우드스피커/수신기, SPK)를 포함하는 부분(ITE)을 포함하는, 귀(RITE) 타입 보청기에서 수신기로 형성되는 예시적인 보청기(HD)(도 9a에서 도시된 보청기(HD)를 예시화 함)를 도시한다. BTE 부분(BTE) 및 ITE 부분(ITE)은 연결 소자(IC)에 의해 접속된다(예를 들어, 전기적으로 접속된다). 도 11의 보청기의 실시예에서, BTE 부분(BTE)은 환경으로부터(도 11의 시나리오에서, 사운드 소스 S로부터) 입력 사운드 신호( $S_{BTE}$ )를 나타내는 전기 입력 오디오 신호를 각각 제공하는 2개의 입력 트랜스듀서들(여기서는 마이크로폰들)( $M_{BTE1}$ ,  $M_{BTE2}$ )을 포함한다. 도 11의 보청기는 각각의 직접 수신된 보조 오디오 및/또는 정보 신호들을 제공하기 위한 2개의 무선 수신기들(WLR1, WLR2)을 더 포함한다. 보청기(HD)는, 다수의 전자 컴포넌트들이

장착되고, 해당 애플리케이션들(아날로그, 디지털 수동 컴포넌트들 등)에 따라 기능적으로 구획되지만 서로 결합되고 전기 컨덕터들(Wx)을 통해 입력 및 출력 유닛들에 결합되는 구성가능한 신호 프로세싱 유닛(SPU), 빔 형성기 필터링 유닛(BFU) 및 메모리 유닛(MEM)을 포함하는 기관(SUB)을 더 포함한다. 언급된 기능 유닛들(및 다른 컴포넌트들)은 해당 애플리케이션에 따라(예를 들어, 크기, 전력 소모, 아날로그 대 디지털 프로세싱 등을 고려하여) 회로들 및 컴포넌트들에 위치될 수 있다. 예를 들어, 언급된 기능 유닛들(및 다른 컴포넌트들)은 하나 이상의 집적 회로들에 집적되거나 또는 하나의 이상의 집적 회로들의 조합 또는 하나 이상의 개별 전기 컴포넌트들(예를 들어, 인덕터, 커패시터 등)의 조합으로서 구현될 수 있다. 구성가능한 신호 프로세싱 유닛(SPU)은, 사용자에게 제공되도록 의도된 강화된 오디오 신호(도 9a의 신호 ES 참조)를 제공한다. 도 11의 보청기의 실시예에서, ITE 부분(ITE)은 전기 신호(도 9a에서 es)를 음향 신호로 변환하기 위한(귀 드럼(ear drum)에서 음향 신호(SED)를 제공하거나 또는 음향 신호(S<sub>ED</sub>)에 기여하기 위한) 라우드스피커(수신기)(SPK)의 형태의 출력 유닛을 포함한다. 일 실시예에서, ITE 부분은, 외이도에서의 또는 외이도 내의 환경으로부터 입력 사운드 신호(S<sub>ITE</sub>)를 나타내는 전기 입력 오디오 신호를 제공하기 위한 입력 트랜스듀서(예를 들어, 마이크로폰)(M<sub>ITE</sub>)를 포함하는 입력 유닛을 더 포함한다. 다른 실시예에서, 보청기는 BTE-마이크로폰들(M<sub>BTE1</sub>, M<sub>BTE2</sub>) 만을 포함할 수 있다. 또 다른 실시예에서, 보청기는 BTE-부분 및/또는 ITE-부분에 위치한 하나 이상의 입력 유닛들과 결합하여 외이도 이외의 곳에 위치한 입력 유닛(IT<sub>3</sub>)을 포함할 수 있다. ITE-부분은 사용자의 외이도 내의 ITE 부분을 안내하고 그리고 위치시키는, 안내 소자, 예를 들어, 돔(DO)을 더 포함한다.

[0276] 도 11에서 예시된 보청기(HD)는 휴대형 디바이스이고 그리고 BTE-부분 및 ITE-부분의 전자 컴포넌트들에 전원을 공급하기 위한 배터리(BAT)를 더 포함한다.

[0277] 보청기(HD)는, 보청기 디바이스를 착용하는 사용자의 국부적인 환경에서 다수의 음향 소스들 중 타겟 음향 소스를 개선하도록 구성된 지향성 마이크로폰 시스템(빔 형성기 필터링 유닛(BFU))을 포함한다. 일 실시예에서, 지향성 시스템은, 마이크로폰 신호의 특정 부분(예를 들어, 타겟 부분 및/또는 노이즈 부분)이 어떤 방향으로부터 발생하는지 검출(예를 들어, 적응적으로 검출)하고, 그리고/또는 현재의 타겟 방향에 관련하여 사용자 인터페이스(예를 들어, 원격 제어기 또는 스마트폰)로부터 입력들을 수신하도록 구성된다. 메모리 유닛(MEM)은, 빔 형성된 신호(Y)를 정의하는 것과 함께(도 9a, 도 9b 참조), 본 발명에 따라 미리정의되거나 또는 고정된(또는 적응적으로 결정된 '고정된') 빔 패턴들을 정의하는 미리정의된 (또는 적으로 결정된) 복소수 주파수 상수들을 포함한다.

[0278] 도 11의 보청기는 본 발명에 따른 보청기 및/또는 바이노럴 보청기 시스템의 일부를 구성하거나 또는 형성할 수 있다.

[0279] 본 발명에 따른 보청기(HD)는, 예를 들어, 스마트폰 또는 다른 휴대용(또는 고정된) 전자 디바이스에서 APP로서 구현되는 보조 디바이스(AUX), 예를 들어, 원격 제어기로 구현되는 도 11에서 도시된 것과 같은 사용자 인터페이스(UI)를 포함할 수 있다. 도 11의 실시예에서, 사용자 인터페이스(UI)의 스크린은, *평활화 빔형성 APP*을 나타낸다. 빔포밍 노이즈 감소 시스템의 신호 대 잡음비의 현재의 평활화를 제어하거나 또는 영향을 미치는 파라미터들, 여기서는, 파라미터들  $\rho$  (바이어스),  $\lambda$  (평활화)(참조 도 3, 4에 관련하여 논의됨)는 *평활화 빔형성 APP*(부제: '방향성, 평활화 파라미터들 구성')에 의해 제어될 수 있다. 바이어스 파라미터  $\rho$  는, 슬라이더를 통해, 최소값(예를 들어, 0)과 최대값(예를 들어, 10dB) 사이의 값으로 설정될 수 있다. 현재의 설정된 값(여기서는, 5dB)은 구성가능한 값들의 범위에 걸쳐 있는 (회색 음영의) 바(bar) 상의 슬라이더 위치에 있는 스크린 상에 표시된다. 마찬가지로, 평활화 파라미터  $\lambda$  는, 슬라이더를 통해, 최소값(예를 들어, 0)과 최대값(예를 들어, 1) 사이의 값으로 설정될 수 있다. 현재의 설정된 값(여기서는, 0.6)은 구성가능한 값들의 범위에 걸쳐 있는 (회색 음영의) 바 상의 슬라이더 위치에 있는 스크린 상에 표시된다. 스크린 하단의 화살표들은, APP의 선행 및 진행 스크린을 변경시키도록 하고, 그리고 2개의 화살표들 사이의 원형 점 상의 탭은 다른 APP들 또는 디바이스의 피쳐들의 선택을 가능하게 하는 메뉴를 불러온다. 스무딩과 관련된 파라미터들  $\rho$  및  $\lambda$  은 사용자에게 반드시 보이지 않을 수도 있다.  $\rho$ ,  $\lambda$  의 세트들은 제3 파라미터로(예를 들어, 작은 소음-큰 소음 감소 바 또는 환경 검출기를 통한 설정)부터 도출될 수 있다.

[0280] 보조 디바이스 및 보청기는, 예를 들어, 무선, 통신 링크(도 11에서의 점선 화살표 WL2)를 통해 현재 선택된 평활화 파라미터들을 나타내는 데이터의 상기 보청기와의 통신을 가능하게 하도록 구성된다. 통신 링크(WL2)는,

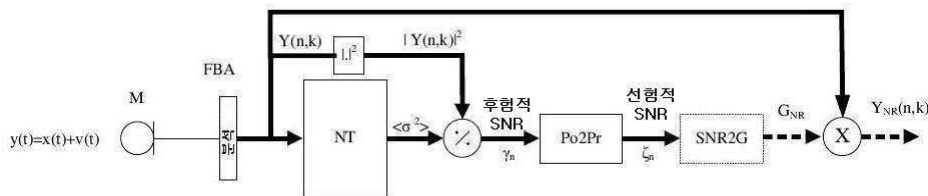
예를 들어, 원거리 통신, 예를 들어, 블루투스 또는 블루투스 저 에너지 (또는 유사한 기술)에 기초할 수 있고, 보청기(HD) 및 보조 디바이스(AUX) 내의 적절한 안테나 및 송수신기 회로에 의해 구현될 수 있으며, 보청기 내의 송수신 유닛(WLR<sub>2</sub>)에 의해 표시될 수 있다. 통신 링크는 단방향 통신(예를 들어, APP 대 보청기) 또는 양방향 통신(예를 들어, 오디오 및/또는 제어 또는 정보 신호들)을 제공하도록 구성될 수 있다.

- [0281] 대응하는 프로세스에 의해 적절하게 대체될 때, 상세한 설명 및/또는 청구범위에서, 전술한 디바이스들의 구조적 피쳐들은 본 방법의 단계와 조합될 수 있다.
- [0282] 사용된 바와 같이, 단수 형태의 단어("하나의" 및 "상기")은, 다르게 명시적으로 언급되지 않는 한, 복수의 형태들도 포함하도록(즉, "적어도 하나"의 의미를 갖도록) 의도된다. 본 명세서에서 사용될 때, 용어들 "포함하다" 및/또는 "포함하는"은 명시된 피쳐들, 정수들, 단계들, 동작들, 요소들 및/또는 컴포넌트들의 존재를 특정하지만, 하나 이상의 다른 피쳐들, 정수들, 단계들, 동작들, 요소들, 컴포넌트들 및/또는 이들의 그룹들을 배제하지 않는 것으로 또한 이해될 것이다. 하나의 소자가 다른 소자에 "접속(connect)"되거나 또는 "결합(couple)"되는 것으로 언급될 때, 하나의 소자가 다른 소자에 직접 접속되거나 결합될 수 있지만, 다르게 명시적으로 언급되지 않으면, 간접 소자들이 또한 존재할 수 있음을 또한 이해될 것이다. 또한, 여기에서 사용되는 "접속된" 또는 "결합된"은 무선으로 접속되거나 또는 결합될 수 있다. 여기에서 사용된 바와 같이, 용어 "및/또는"은, 하나 이상의 관련된 항목들의 일부 그리고 모든 조합들을 포함한다. 어떤 개시된 방법의 단계들은 다르게 명시적으로 언급되지 않으면, 여기에서 언급된 정확한 순서에 제한되지 않는다.
- [0283] 본 명세서 전체에 걸친, "하나의 실시예" 또는 "일 실시예" 또는 "하나의 양상" 또는 포함된 피쳐들에 관한 언급은, 상기 실시예와 관련하여 서술된 특정 피쳐, 구조 또는 특성은 본 발명의 적어도 하나의 실시예에 포함되었음을 의미"할 수 있다". 더욱이, 특정 피쳐들, 구조들 또는 특성들은 본 발명의 하나 이상의 실시예들에 적합하게 결합될 수 있다. 이전의 설명은, 여기에서 설명된 다양한 양상들을 실시할 수 있도록 제공된다. 이러한 양상들에 대한 다양한 변형들이 통상의 기술자에게 명백할 것이고, 본 명세서에 정의된 일반적인 원리들은 다른 양상들에 적용될 수 있다.
- [0284] 청구 범위들은 여기에서 도시된 양상들로 제한되도록 의도되지 않으며, 청구 범위의 언어와 일치하는 전체 범위가 주어져야 하며, 단수의 요소에 대한 언급은, 구체적으로 그렇게 언급되지 않으면, "단 하나"를 의미하도록 의도하는 것이 아닌, 오히려 "하나 이상"을 의미하도록 의도된다. 다르게 명시적으로 언급되지 않으면, 용어 "일부"는 하나 이상을 언급한다.
- [0285] 따라서, 범위는 다음의 청구 범위에 의해 판단되어야 한다.
- [0287] 참조들
- [0288] [1] Ephraim, Y.; Malah, D., "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator", IEEE Transactions on Acoustics, Speech and Signal Processing, vol.32, no.6, pp. 1109- 1121, Dec 1984
- [0289] URL:<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1164453&isnumber=26187>
- [0290] [2] EP2701145A1
- [0291] [3] Martin, R., "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics", IEEE Transactions on Speech and Audio Processing, vol.9, no.5, pp. 504-512, Apr 2001
- [0292] [4] Ephraim, Y.; Malah, D., "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator", IEEE Transactions on Acoustics, Speech and Signal Processing, vol.33, no.2, pp. 443- 445, Apr 1985
- [0293] URL:<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1164550&isnumber=26190>
- [0294] [5] Breithaupt, C.; Martin, R., "Analysis of the Decision-Directed SNR Estimator for Speech Enhancement With Respect to Low-SNR and Transient Conditions", IEEE Transactions on Audio, Speech, and Language Processing, vol.19, no.2, pp. 277-289, Feb. 2011
- [0295] URL:<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5444986&isnumber=5609232>

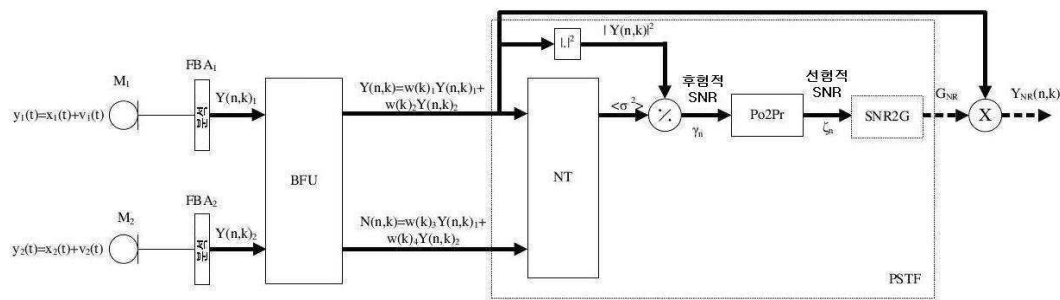
- [0296] [6] Cappe, O., "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor," Speech and Audio Processing, IEEE Transactions on , vol.2, no.2, pp. 345-349, Apr 1994
- [0297] URL:<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=279283&isnumber=6926>
- [0298] [7] Loizou, P. (2007). Speech Enhancement: Theory and Practice, CRC Press, Boca Raton: FL

도면

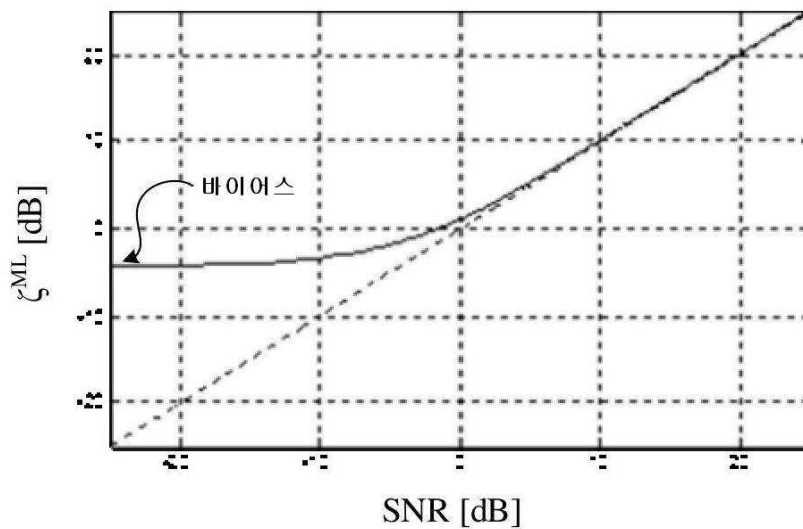
도면1a



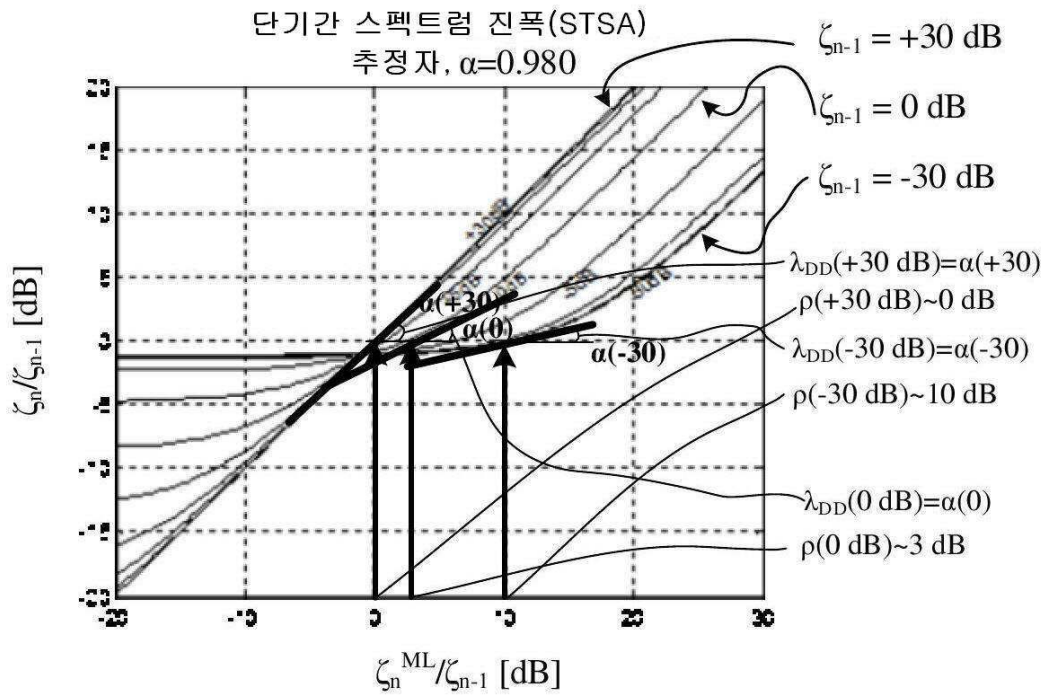
도면1b



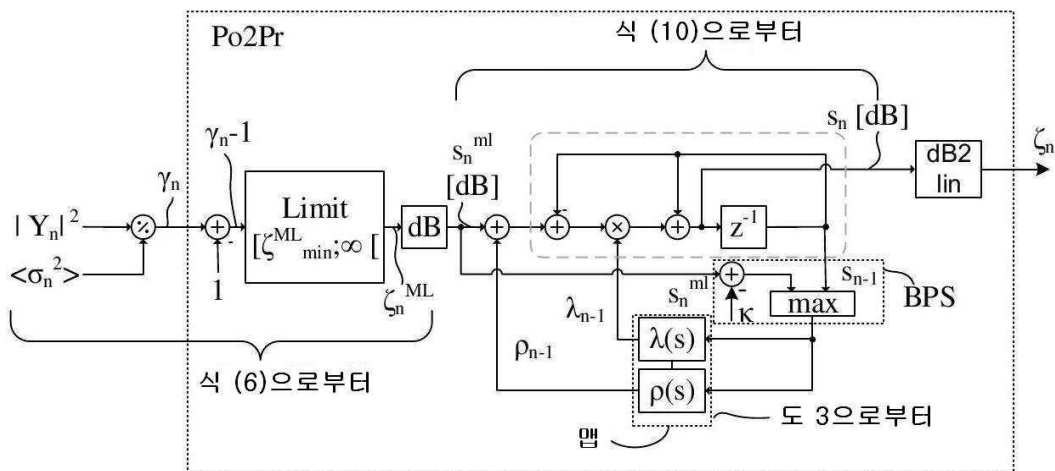
도면2



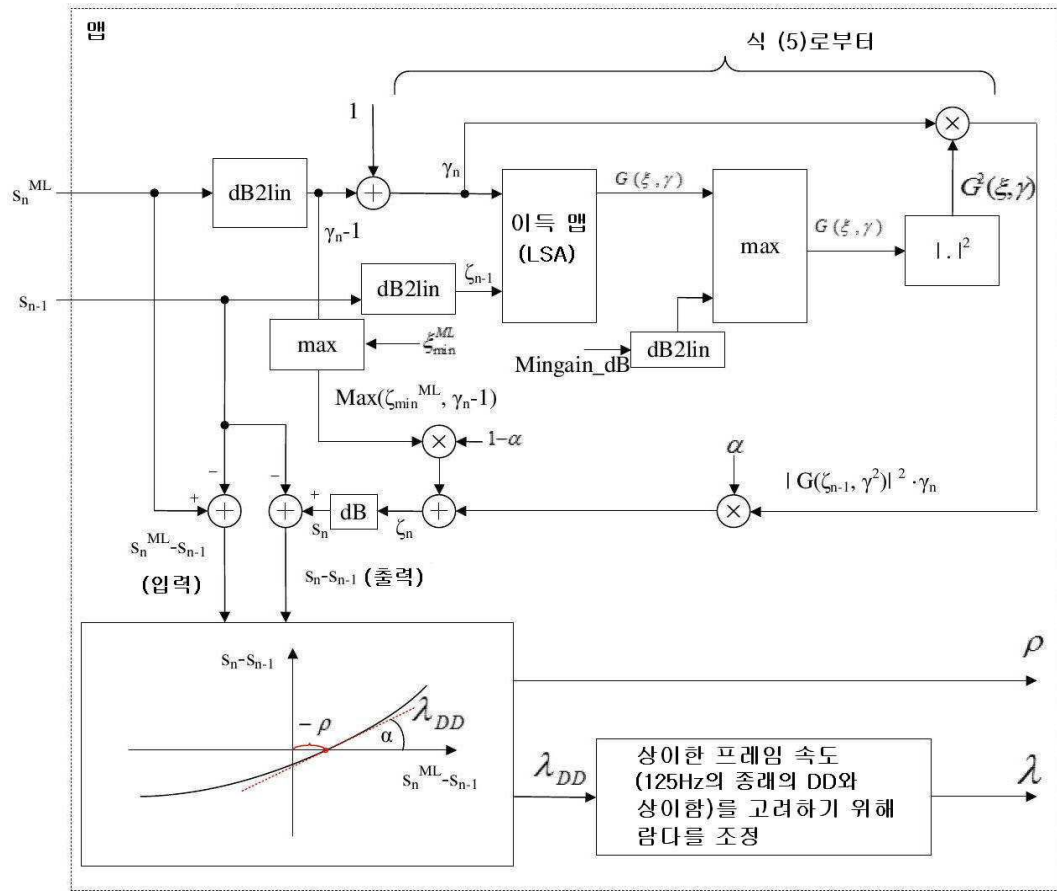
도면3



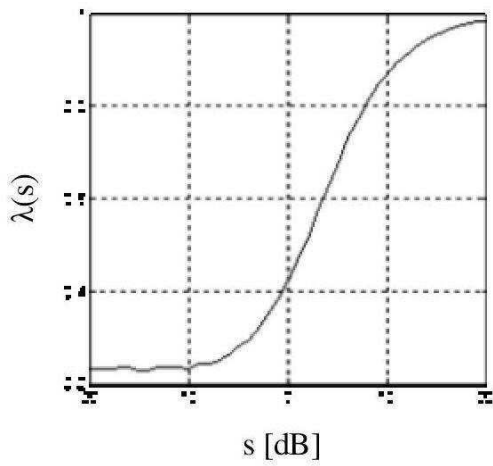
도면4



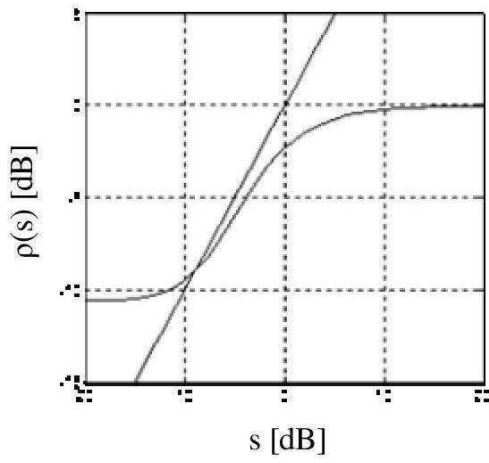
도면5



도면6a

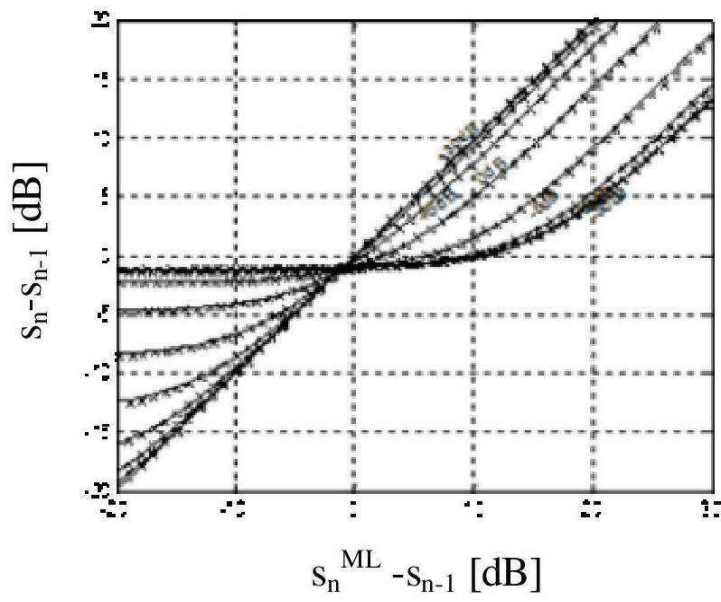


도면6b

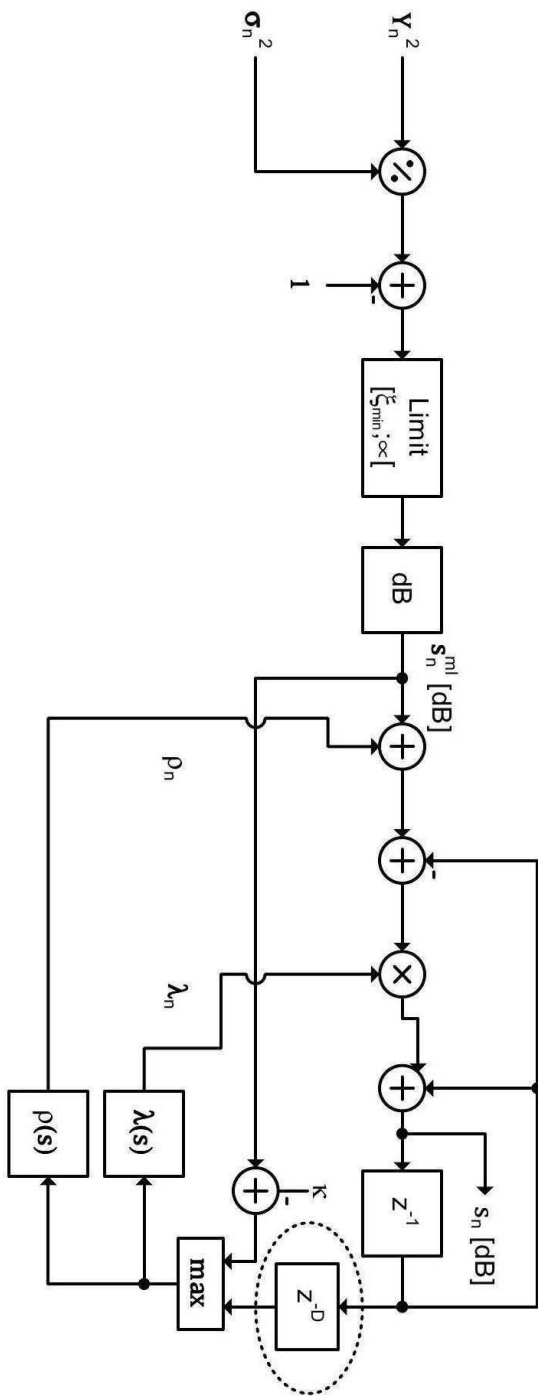


도면7

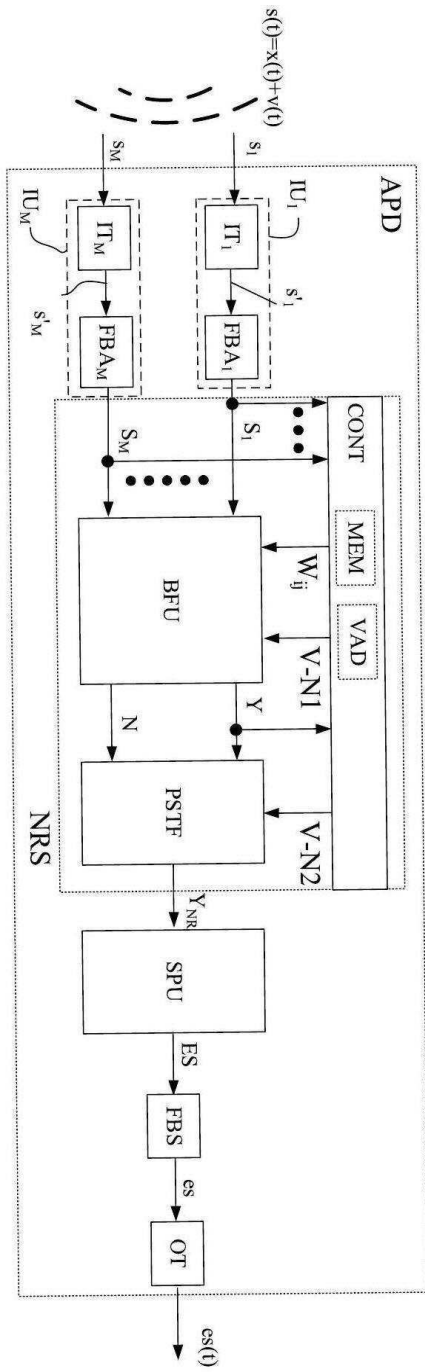
STSA,  $\alpha=0.980$



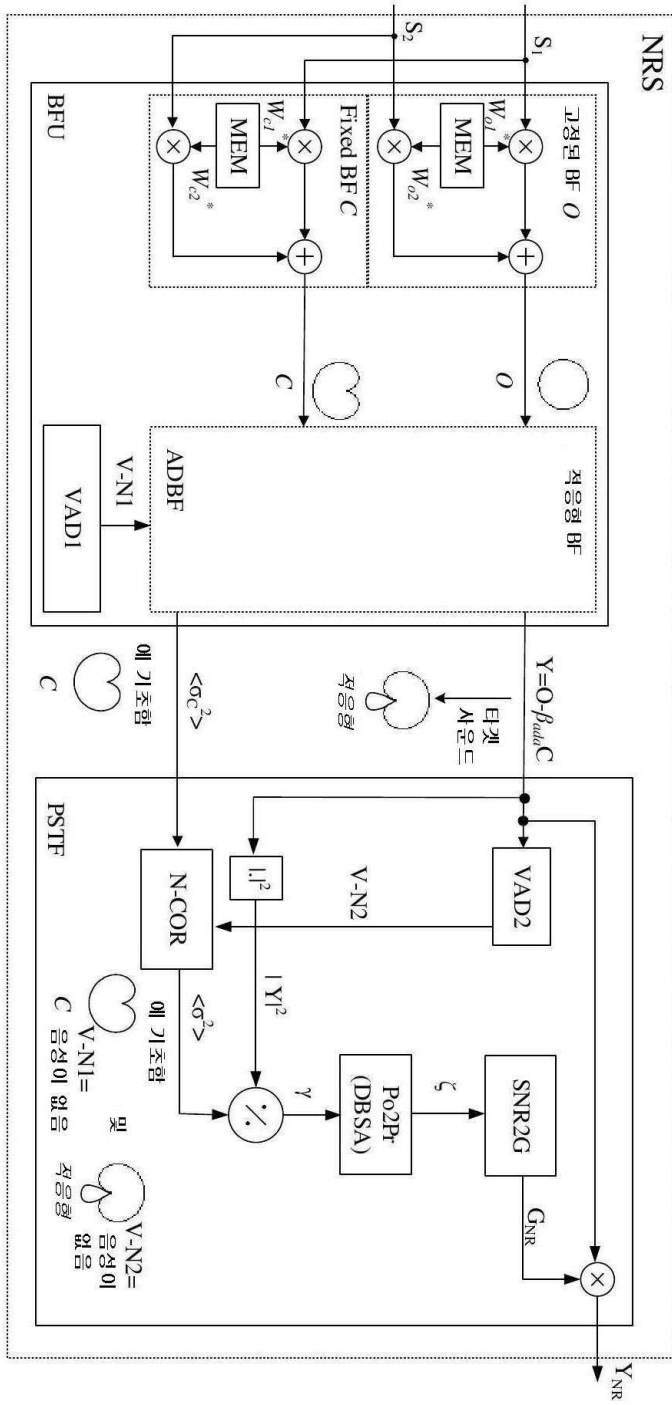
도면8



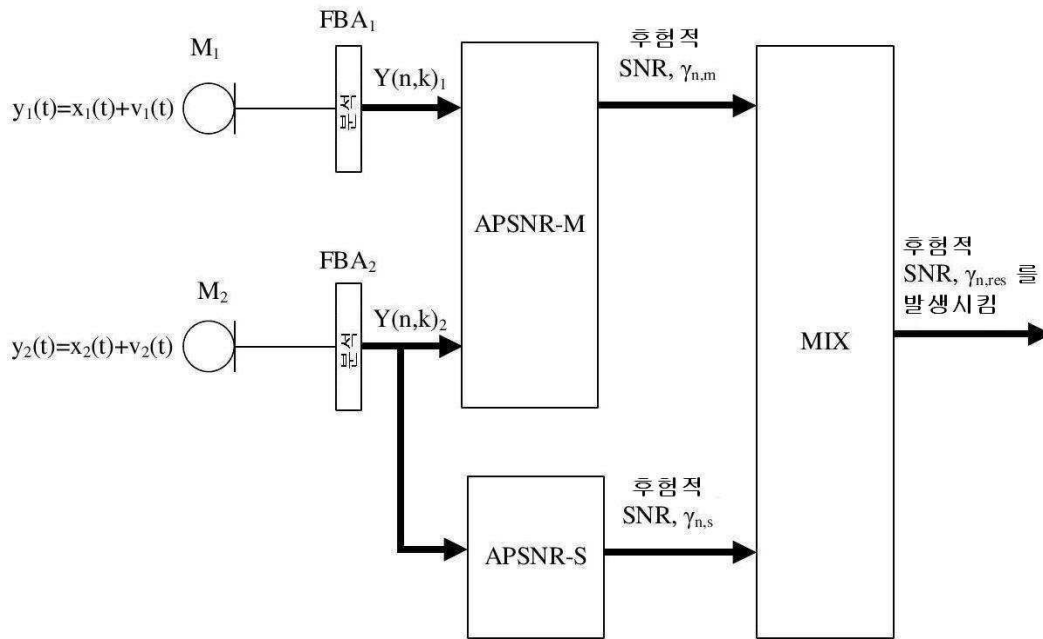
도면9a



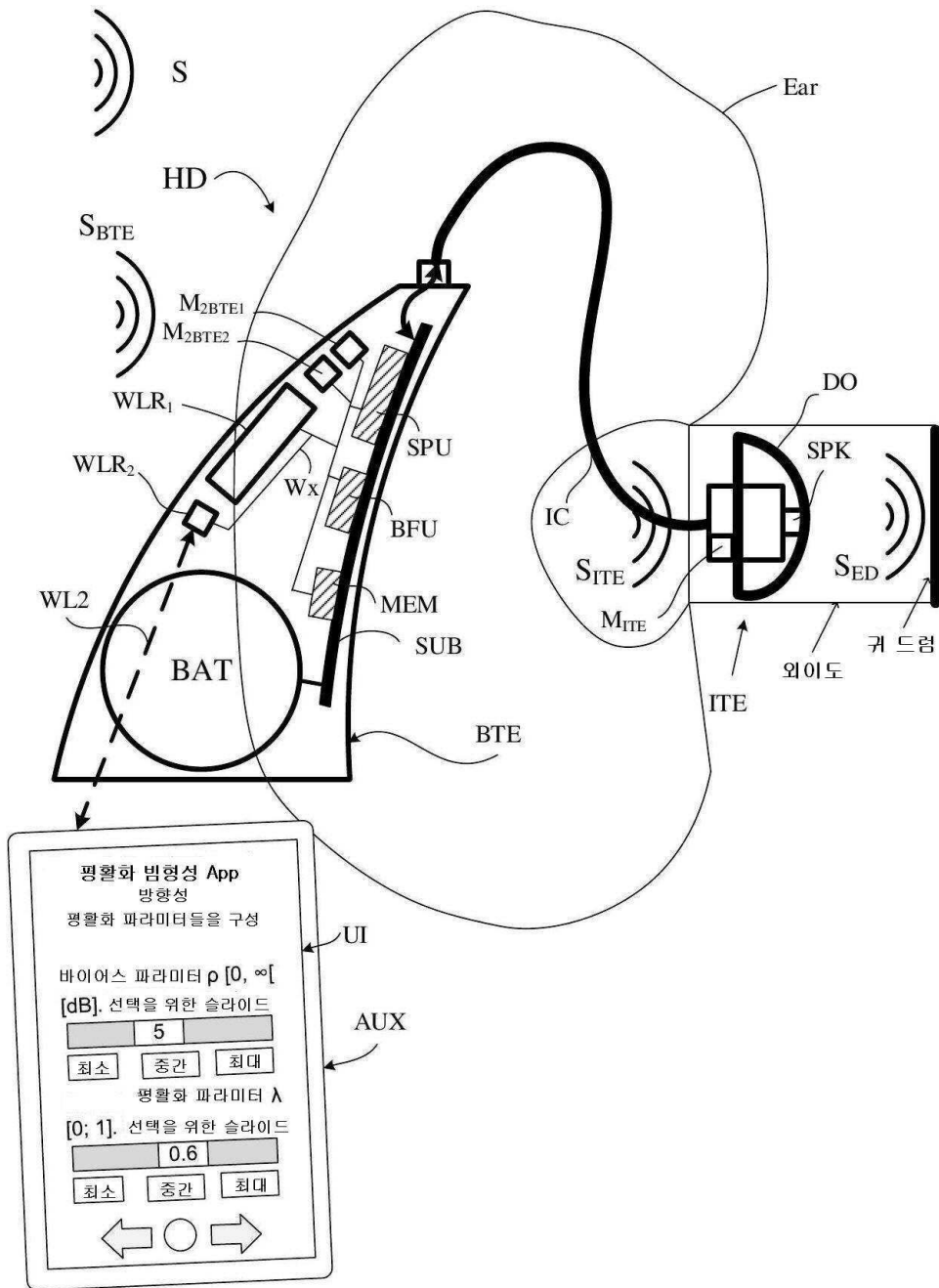
도면9b



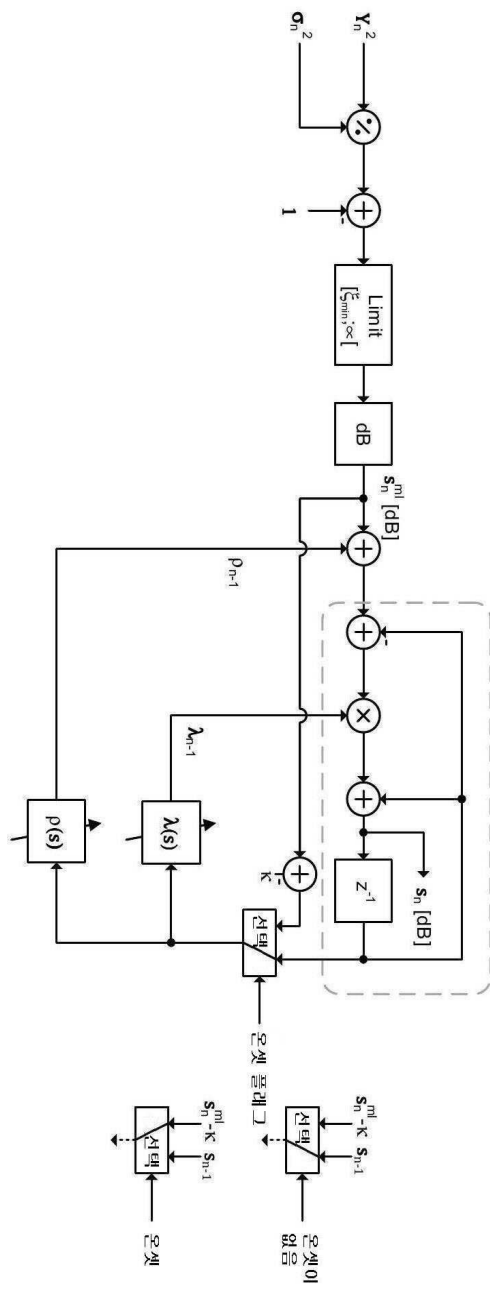
도면10



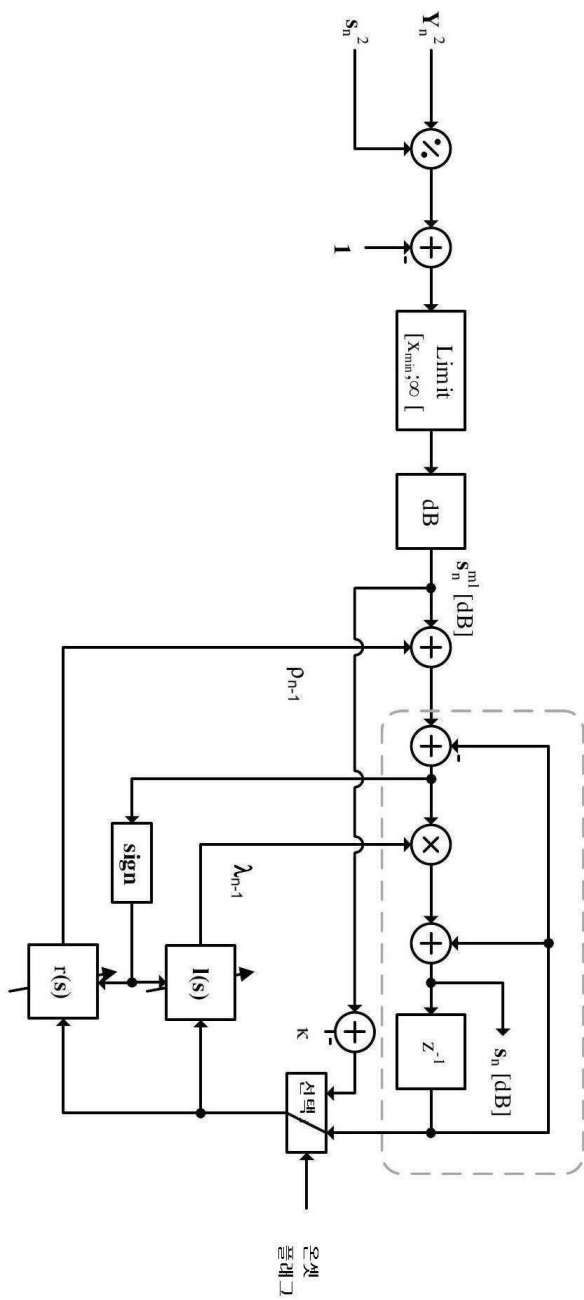
도면11



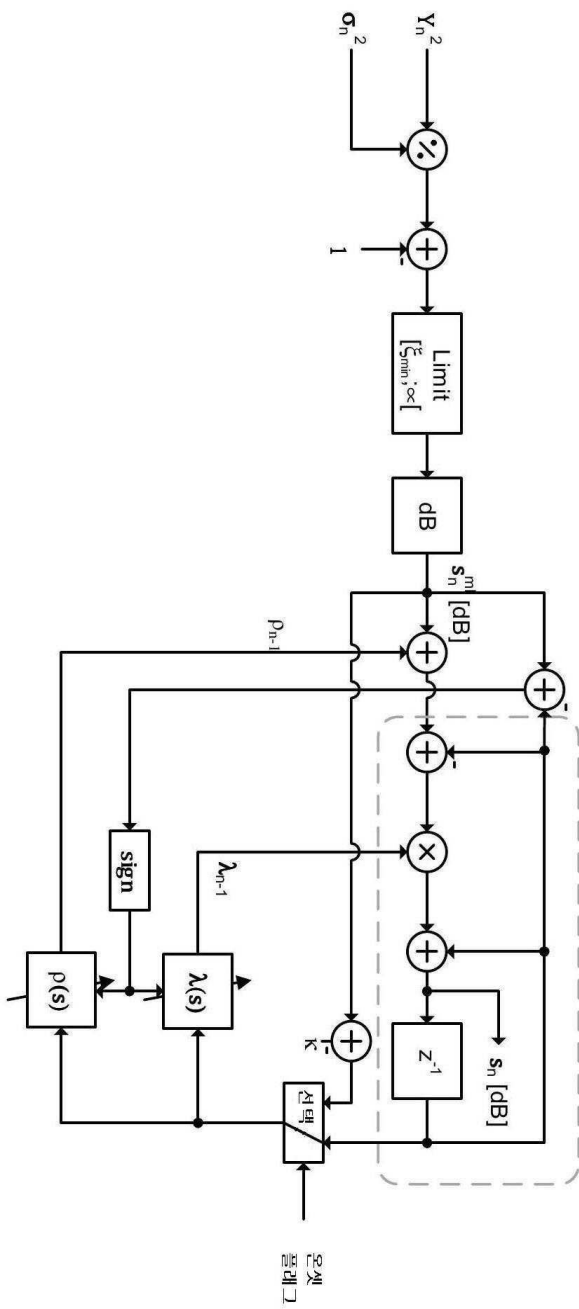
도면12a



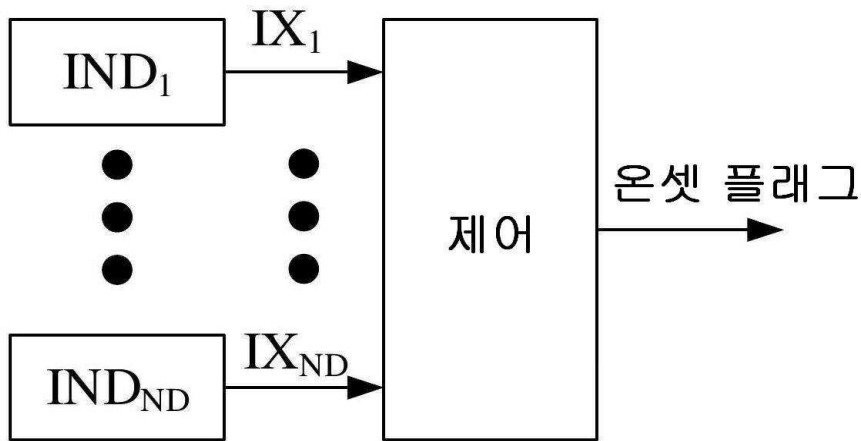
도면12b



도면12c



도면13a



도면13b

