(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2011-0109333
(43) 공개일자 2011년10월06일

(51) Int. Cl.
G10K 11/175 (2006.01) *G10K 11/04* (2006.01)
G10K 11/16 (2006.01)
(21) 출원번호 10-2010-0029011
(22) 출원일자 2010년03월31일
심사청구일자 2010년03월31일

(71) 출원인
경상대학교산학협력단
경상남도 진주시 가좌동 900
(72) 발명자
정상배
경상남도 진주시 금산면 중천리 330 진주금산진흥더루벤스APT 102-701
이영재
경상남도 창원시 신월동 은아아파트 120동 405호
김수환
경상남도 마산시 양덕동 대림하이빌 104동 204호
(74) 대리인
송윤호, 오세준, 권혁수

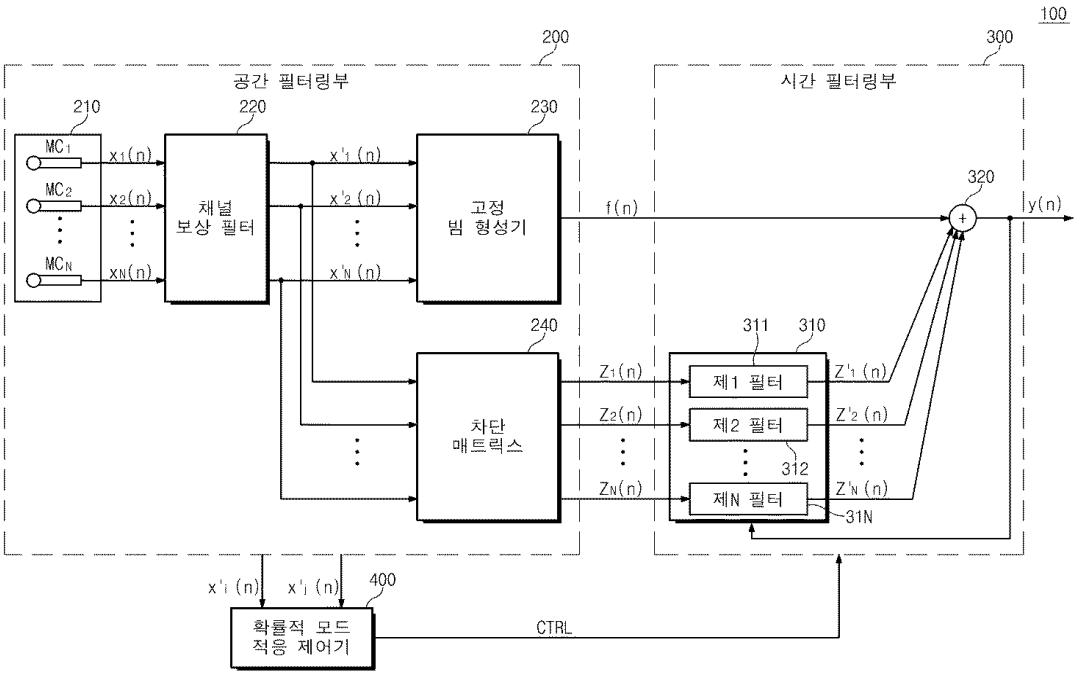
전체 청구항 수 : 총 12 항

(54) 잡음 제거 장치 및 잡음 제거 방법

(57) 요 약

본 발명은 잡음 제거 장치에 관한 것이다. 본 발명의 잡음 제거 장치는 공간 필터링부, 시간 필터링부, 그리고 모드 적응 제어기로 구성된다. 공간 필터링부는 복수의 채널들을 통해 음성 신호들을 수신하고, 수신된 음성 신호들에 기반하여 잡음이 감소된 음성 신호를 생성하고, 그리고 수신된 음성 신호들에 기반하여 복수의 채널들에 대응하는 복수의 잡음 신호들을 생성한다. 시간 필터링부는 복수의 잡음 신호들을 적응적 필터링하여 잡음이 감소된 음성 신호의 잡음 성분에 대응하는 잡음 성분 신호들을 생성하고, 그리고 생성된 잡음 성분 신호들을 이용하여 잡음이 감소된 음성 신호로부터 잡음 성분을 제거한다. 모드 적응 제어기는 수신된 음성 신호들 중 적어도 두 개의 신호들에 기반하여 적어도 두 개의 신호들의 잡음 구간들을 판별하고, 판별 결과에 기반하여 시간 필터링부의 상기 적응적 필터링의 필터링 계수의 학습 모드를 활성화 또는 비활성화 한다. 모드 적응 제어기는 적어도 두 개의 신호들의 특정 구간들이 잡음 구간들일 확률을 산출하고, 산출된 확률에 기반하여 잡음 구간들을 판별한다.

대 표 도



특허청구의 범위

청구항 1

복수의 채널들을 통해 음성 신호들을 수신하고, 상기 수신된 음성 신호들에 기반하여 잡음이 감소된 음성 신호를 생성하고, 그리고 상기 수신된 음성 신호들에 기반하여 상기 복수의 채널들에 대응하는 복수의 잡음 신호들을 생성하도록 구성되는 공간 필터링부;

상기 복수의 잡음 신호들을 적응적 필터링하여 상기 잡음이 감소된 음성 신호의 잡음 성분에 대응하는 잡음 성분 신호들을 생성하고, 그리고 상기 생성된 잡음 성분 신호들을 이용하여 상기 잡음이 감소된 음성 신호로부터 잡음 성분을 제거하도록 구성되는 시간 필터링부; 그리고

상기 수신된 음성 신호들 중 적어도 두 개의 신호들에 기반하여 상기 적어도 두 개의 신호들의 잡음 구간들을 판별하고, 상기 판별 결과에 기반하여 상기 시간 필터링부의 상기 적응적 필터링의 필터링 계수의 학습 모드를 활성화 또는 비활성화 하도록 구성되는 모드 적응 제어기를 포함하고,

상기 모드 적응 제어기는 상기 적어도 두 개의 신호들의 특정 구간들이 상기 잡음 구간들일 확률을 산출하고, 상기 산출된 확률에 기반하여 상기 잡음 구간들을 판별하도록 구성되는 잡음 제거 장치.

청구항 2

제 1 항에 있어서,

상기 모드 적응 제어기는

상기 적어도 두 개의 신호들의 선형 예측 잔차 신호를 산출하고, 상기 산출된 선형 예측 잔차 신호에 기반하여 상기 적어도 두 개의 신호들의 상호 상관도를 산출하고, 상기 산출된 상호 상관도에 기반하여 우도를 산출하고, 그리고 상기 산출된 우도에 기반하여 상기 잡음 구간들을 판별하도록 구성되는 잡음 제거 장치.

청구항 3

제 2 항에 있어서,

상기 모드 적응 제어기는 상기 산출된 우도가 미리 설정된 값 보다 큰 구간들을 상기 잡음 구간들로 판별하도록 구성되는 잡음 제거 장치.

청구항 4

제 2 항에 있어서,

상기 모드 적응 제어기는 상기 산출된 우도가 미리 설정된 값 보다 클 때 상기 학습 모드를 활성화하고, 그리고 상기 산출된 우도가 상기 미리 설정된 값 보다 작을 때 상기 학습 모드를 비활성화하도록 구성되는 잡음 제거 장치.

청구항 5

제 1 항에 있어서,

상기 수신된 음성 신호들의 음성 구간들을 판별하고, 상기 수신된 음성 신호들의 음성 구간 신호들 및 상기 수신된 신호들의 비음성 구간 신호들에 기반하여 상기 수신된 음성 신호들을 조절하도록 구성되는 채널 보상 필터를 더 포함하는 잡음 제거 장치.

청구항 6

제 5 항에 있어서,

상기 채널 보상 필터는

상기 음성 구간 신호들의 통계 및 상기 비음성 구간 신호들의 통계에 기반하여 상기 수신된 음성 신호들을 조절하도록 구성되는 잡음 제거 장치.

청구항 7

제 5 항에 있어서,

상기 채널 보상 필터는

상기 수신된 음성 신호들 중 기준 채널에 대응하는 음성 구간 신호, 상기 수신된 음성 신호들 중 특정 채널에 대응하는 음성 구간 신호, 상기 기준 채널에 대응하는 비음성 구간 신호, 그리고 상기 특정 채널에 대응하는 비음성 구간 신호에 기반하여 상기 특정 채널을 통해 수신되는 신호를 조절하도록 구성되는 잡음 제거 장치.

청구항 8

제 5 항에 있어서,

상기 채널 보성 필터는 상기 복수의 채널들 사이의 차이로 인해 상기 수신된 음성 신호들에서 발생하는 차이를 보상하도록 구성되는 잡음 제거 장치.

청구항 9

제 5 항에 있어서,

상기 모드 적응 제어기는

상기 채널 보상 필터의 출력들 중 적어도 두 개의 출력들에 기반하여 상기 잡음 구간들을 판별하도록 구성되는 잡음 제거 장치.

청구항 10

적어도 두 개의 채널들을 통해 음성 신호들을 수신하는 단계;

상기 수신된 음성 신호들에 기반하여, 상기 수신된 음성 신호들의 특정 구간들이 잡음 구간들일 확률을 산출하는 단계;

상기 산출된 확률에 기반하여, 상기 수신된 음성 신호들의 잡음 구간들을 판별하는 단계;

상기 수신된 음성 신호들의 잡음 성분을 적응적 필터링하는 단계; 그리고

상기 필터링 결과를 이용하여 상기 수신된 음성 신호들로부터 잡음 성분을 제거하는 단계를 포함하고,

상기 적응적 필터링 시의 필터링 계수의 학습 모드는 상기 판별된 잡음 구간들에서 활성화되는 잡음 제거 방법.

청구항 11

제 10 항에 있어서,

상기 수신된 음성 신호들의 잡음 구간들을 판별하는 단계는

상기 수신된 음성 신호들의 선형 예측 잔차 신호를 산출하는 단계;

상기 산출된 선형 예측 잔차 신호에 기반하여 상호 상관도를 산출하는 단계;

상기 산출된 상호 상관도에 기반하여 우도를 산출하는 단계; 그리고

상기 산출된 우도에 기반하여 상기 수신된 음성 신호들의 잡음 구간들을 판별하는 단계를 포함하는 잡음 제거 방법.

청구항 12

제 10 항에 있어서,

상기 수신된 음성 신호들의 음성 구간들을 판별하는 단계; 그리고

상기 수신된 음성 신호들의 상기 판별된 음성 구간들 및 비음성 구간들에 기반하여, 상기 수신된 음성 신호들의 채널들의 특성에 따른 차이를 보상하는 단계를 더 포함하고;

상기 잡음 구간들은 상기 차이가 보상된 음성 신호들에 기반하여 판별되는 잡음 제거 방법.

명 세 서

기 술 분 야

[0001] 본 발명은 음성 신호의 잡음 제거 장치 및 잡음 제거 방법에 관한 것이다.

배 경 기 술

[0002] 전자 장치는 사용자로부터 신호를 수신하는 다양한 사용자 인터페이스들을 제공한다. 예를 들면, 전자 장치는 키보드, 마우스, 버튼 등과 같은 사용자 인터페이스들을 제공한다. 기술이 발전하면서, 전자 장치에 제공되는 사용자 인터페이스는 더욱 다양해지고 있다. 예를 들면, 터치 패널, 인체 감응 센서 등과 같은 다양한 사용자 인터페이스가 전자 장치에 제공된다.

[0003] 기술의 발전과 함께 새롭게 연구되고 있는 사용자 인터페이스 중 하나는 음성 인식 장치이다. 음성 인식 장치는 사용자의 음성으로부터 정보를 추출하여 사용한다.

[0004] 통상적인 환경에서, 다양한 형태의 잡음들이 존재한다. 예를 들면, 자동차 소음, 전자 장치의 소음, 주변 사람의 음성, 방송 매체의 방송 등과 같은 다양한 잡음들이 통상적인 환경에 존재한다. 잡음이 존재하면, 음성 인식 장치의 음성 인식률이 저하된다. 음성 인식 장치의 음성 인식률을 향상시키기 위하여, 다양한 형태의 잡음 제거 장치 및 잡음 제거 방법들이 연구되고 있다.

발명의 내용

해결하려는 과제

[0005] 본 발명의 목적은 향상된 동작 성능을 갖는 잡음 제거 장치 및 잡음 제거 방법을 제공하는 데에 있다.

과제의 해결 수단

[0006] 본 발명의 실시 예에 따른 잡음 제거 장치는, 복수의 채널들을 통해 음성 신호들을 수신하고, 상기 수신된 음성 신호들에 기반하여 잡음이 감소된 음성 신호를 생성하고, 그리고 상기 수신된 음성 신호들에 기반하여 상기 복수의 채널들에 대응하는 복수의 잡음 신호들을 생성하도록 구성되는 공간 필터링부; 상기 복수의 잡음 신호들을 적응적 필터링하여 상기 잡음이 감소된 음성 신호의 잡음 성분에 대응하는 잡음 성분 신호들을 생성하고, 그리고 상기 생성된 잡음 성분 신호들을 이용하여 상기 잡음이 감소된 음성 신호로부터 잡음 성분을 제거하도록 구성되는 시간 필터링부; 그리고 상기 수신된 음성 신호들 중 적어도 두 개의 신호들에 기반하여 상기 적어도 두 개의 신호들의 잡음 구간들을 판별하고, 상기 판별 결과에 기반하여 상기 시간 필터링부의 상기 적응적 필터링의 필터링 계수의 학습 모드를 활성화 또는 비활성화 하도록 구성되는 모드 적응 제어기를 포함하고, 상기 모드 적응 제어기는 상기 적어도 두 개의 신호들의 특정 구간들이 상기 잡음 구간들일 확률을 산출하고, 상기 산출된 확률에 기반하여 상기 잡음 구간들을 판별하도록 구성된다.

[0007] 실시 예로서, 상기 모드 적응 제어기는, 상기 적어도 두 개의 신호들의 선형 예측 잔차 신호를 산출하고, 상기 산출된 선형 예측 잔차 신호에 기반하여 상기 적어도 두 개의 신호들의 상호 상관도를 산출하고, 상기 산출된 상호 상관도에 기반하여 우도를 산출하고, 그리고 상기 산출된 우도에 기반하여 상기 잡음 구간들을 판별하도록 구성된다.

[0008] 실시 예로서, 상기 모드 적응 제어기는 상기 산출된 우도가 미리 설정된 값 보다 큰 구간들을 상기 잡음 구간들로 판별하도록 구성된다.

[0009] 실시 예로서, 상기 모드 적응 제어기는 상기 산출된 우도가 미리 설정된 값 보다 클 때 상기 학습 모드를 활성화하고, 그리고 상기 산출된 우도가 상기 미리 설정된 값 보다 작을 때 상기 학습 모드를 비활성화하도록 구성된다.

[0010] 실시 예로서, 상기 수신된 음성 신호들의 음성 구간들을 판별하고, 상기 수신된 음성 신호들의 음성 구간 신호들 및 상기 수신된 신호들의 비음성 구간 신호들에 기반하여 상기 수신된 음성 신호들을 조절하도록 구성되는 채널 보상 필터를 더 포함한다.

[0011] 실시 예로서, 상기 채널 보상 필터는 상기 음성 구간 신호들의 통계 및 상기 비음성 구간 신호들의 통계에 기반하여 상기 수신된 음성 신호들을 조절하도록 구성된다.

[0012] 실시 예로서, 상기 채널 보상 필터는, 상기 수신된 음성 신호들 중 기준 채널에 대응하는 음성 구간 신호, 상기 수신된 음성 신호들 중 특정 채널에 대응하는 음성 구간 신호, 상기 기준 채널에 대응하는 비음성 구간 신호, 그리고 상기 특정 채널에 대응하는 비음성 구간 신호에 기반하여 상기 특정 채널을 통해 수신되는 신호를 조절하도록 구성된다.

[0013] 실시 예로서, 상기 채널 보성 필터는 상기 복수의 채널들 사이의 차이로 인해 상기 수신된 음성 신호들에서 발생하는 차이를 보상하도록 구성된다.

[0014] 실시 예로서, 상기 모드 적응 제어기는 상기 채널 보상 필터의 출력들 중 적어도 두 개의 출력들에 기반하여 상기 잡음 구간들을 판별하도록 구성된다.

[0015] 본 발명의 실시 예에 따른 잡음 제거 방법은, 적어도 두 개의 채널들을 통해 음성 신호들을 수신하는 단계; 상기 수신된 음성 신호들에 기반하여, 상기 수신된 음성 신호들의 특정 구간들이 잡음 구간들일 확률을 산출하는 단계; 상기 산출된 확률에 기반하여, 상기 수신된 음성 신호들의 잡음 구간들을 판별하는 단계; 상기 수신된 음성 신호들의 잡음 성분을 적응적 필터링하는 단계; 그리고 상기 필터링 결과를 이용하여 상기 수신된 음성 신호들로부터 잡음 성분을 제거하는 단계를 포함하고, 상기 적응적 필터링 시의 필터링 계수의 학습 모드는 상기 판별된 잡음 구간들에서 활성화된다.

[0016] 실시 예로서, 상기 수신된 음성 신호들의 잡음 구간들을 판별하는 단계는, 상기 수신된 음성 신호들의 선형 예측 잔차 신호를 산출하는 단계; 상기 산출된 선형 예측 잔차 신호에 기반하여 상호 상관도를 산출하는 단계; 상기 산출된 상호 상관도에 기반하여 우도를 산출하는 단계; 그리고 상기 산출된 우도에 기반하여 상기 수신된 음성 신호들의 잡음 구간들을 판별하는 단계를 포함한다.

[0017] 실시 예로서, 상기 수신된 음성 신호들의 음성 구간들을 판별하는 단계; 그리고 상기 수신된 음성 신호들의 상기 판별된 음성 구간들 및 비음성 구간들에 기반하여, 상기 수신된 음성 신호들의 채널들의 특성에 따른 차이를 보상하는 단계를 더 포함하고; 상기 잡음 구간들은 상기 차이가 보상된 음성 신호들에 기반하여 판별된다.

발명의 효과

[0018] 본 발명에 의하면, 수신된 음성 신호들의 잡음 구간 확률에 기반하여 적응적 필터링의 학습 모드가 제어된다. 또한, 음성 신호의 음성 구간의 통계 및 잡음 구간의 통계에 따라 채널 특성이 보상된다. 따라서, 향상된 동작 성능을 갖는 잡음 제거 장치 및 잡음 제거 방법이 제공된다.

도면의 간단한 설명

[0019] 도 1은 본 발명의 실시 예에 따른 잡음 제거 장치를 보여주는 블록도이다.

도 2는 도 1의 음성 수신기의 출력 신호들을 보여주는 파형도이다.

도 3은 도 1의 확률적 모드 적응 제어기의 동작 방법을 보여주는 순서도이다.

도 4는 도 1의 채널 보상 필터의 동작 방법을 보여주는 순서도이다.

도 5는 도 1 내지 도 4를 참조하여 설명된 잡음 제거 장치의 실험 환경을 보여주는 블록도이다.

도 6은 도 5의 실험 환경에서 도 1의 잡음 제거 장치를 실험한 제 1 결과를 보여주는 테이블이다.

도 7은 도 5의 실험 환경에서 도 1의 잡음 제거 장치를 실험한 제 2 결과를 보여주는 테이블이다.

도 8은 도 1 내지 도 7을 참조하여 설명된 잡음 제거 장치를 포함하는 음성 인식 시스템을 보여주는 블록도이다.

도 9는 도 8의 음성 인식 시스템을 포함하는 컴퓨팅 시스템을 보여주는 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0020] 이하에서, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 정도로 상세히 설명하기 위하여, 본 발명의 실시 예를 첨부된 도면을 참조하여 설명하기로 한다.

[0021] 도 1은 본 발명의 실시 예에 따른 잡음 제거 장치(100)를 보여주는 블록도이다. 도 1을 참조하면, 잡음 제거 장치(100)는 공간 필터링부(200), 시간 필터링부(300), 그리고 확률적 모드 적응 제어기(400)를 포함한다. 공간 필터링부(200)는 잡음 제거 장치(100)의 공간적 특성에 따라 잡음을 제거하도록 구성된다. 시간 필터링부(300)는 음성 신호의 시간적 특성에 따라 잡음을 제거하도록 구성된다.

[0022] 공간 필터링부(200)는 음성 수신기(210), 채널 보상 필터(220), 고정 빔 형성기(230), 그리고 차단 매트릭스(240)를 포함한다.

[0023] 음성 수신기(210)는 복수의 마이크들(MC_1~MC_N)을 포함한다. 음성 수신기(210)는 복수의 마이크들(MC_1~MC_N)을 통해 음성을 수신하여 제 1 내지 제 N 음성 신호들($x_1(n)$~$x_N(n)$)을 출력한다.

[0024] 예시적으로, 음성 수신기(210)의 출력 신호들이 도 2에 도시되어 있다. 도 2에서, 가로 축은 시간을 나타내며, 세로 축은 각 채널별 진폭을 나타낸다. 예시적으로, 도 2에서, 제 1 내지 제 4 채널들(CH1~CH4)에 대응하는 음성 신호들이 도시되어 있다. 그러나, 음성 수신기(210)는 제 1 내지 제 4 채널들(CH1~CH4)에 대응하는 신호를 출력하는 것으로 한정되지 않는다. 예시적으로, 제 1 시간(t1) 내지 제 2 시간(t2) 동안 목표 음성이 수신되고, 나머지 시간 동안 잡음이 수신된다. 즉, 제 1 시간(t1) 내지 제 2 시간(t2)은 음성 구간이고, 나머지 시간은 잡음 구간이다.

[0025] 다시 도 1을 참조하면, 음성 수신기(210)의 출력 신호들($x_1(n)$~$x_N(n)$)은 채널 보상 필터(220)로 전달된다.

[0026] 채널 보상 필터(220)는 음성 수신기(210)로부터 음성 신호들($x_1(n)$~$x_N(n)$)을 수신한다. 채널 보상 필터(220)는 채널들 사이의 차이에 의해 음성 신호들($x_1(n)$~$x_N(n)$)에서 발생되는 차이들을 보상하도록 구성된다. 예시적으로, 특정 음성 소스로부터 음성 수신기(210)에 음성이 수신될 때, 음성 소스 및 마이크들(MC_1~MC_N) 사이의 거리 및 채널 환경은 상이하다. 상이한 거리 및 채널 환경에 따라, 음성 수신기(210)의 출력 신호들($x_1(n)$~$x_N(n)$) 사이에 차이가 발생할 것이다.

[0027] 채널 보상 필터(220)는 음성 소스 및 마이크들(MC_1~MC_N) 사이의 거리 및 채널 환경의 차이를 보상하도록 구성된다. 채널 보상 필터(220)는 음성 신호들($x_1(n)$~$x_N(n)$)을 조절하여 보상된 신호들($x'_1(n)$~$x'_N(n)$)을 출력한다. 보상된 신호들($x'_1(n)$~$x'_N(n)$)은 고정 빔 형성기(230) 및 차단 매트릭스(240)에 전달된다.

[0028] 고정 빔 형성기(230)는 채널 보상 필터(220)로부터 보상된 신호들($x'_1(n)$~$x'_N(n)$)을 수신한다. 고정 빔 형성기(230)는 수신된 신호들($x'_1(n)$~$x'_N(n)$)에 기반하여 빔 포밍(beam foraming)을 수행하도록 구성된다. 예를 들면, 고정 빔 형성기(230)는 수신된 신호들($x'_1(n)$~$x'_N(n)$)에 기반하여 잡음이 감소된 신호(f(n))를 생성한다.

[0029] 예를 들면, 고정 빔 형성기(230)는 수신된 신호들($x'_1(n)$~$x'_N(n)$) 중 음성 구간 신호들의 위상을 보상하여 합산한다. 음성 소스 및 마이크들(MC_1~MC_N) 사이의 거리 및 채널의 차이로 인해, 수신된 신호들($x'_1(n)$~$x'_N(n)$)의 위상들은 상이할 것이다. 도 2에 도시된 바와 같이, 수신된 신호들($x'_1(n)$~$x'_N(n)$)의 제 1 시간(t1) 및 제 2 시간(t2) 사이의 구간들은 음성 구간이며, 나머지 구간들은 잡음 구간이다. 고정 빔 형성기(230)는 수신된 신호들($x'_1(n)$~$x'_N(n)$)의 음성 구간 신호들의 위상이 일치하도록 보상할 것이다. 위상이 보상된 신호들이 합산되면, 위상의 차이로 인해 잡음 성분 중 일부가 상쇄될 것이다. 따라서, 합산된 신호의 신호 대 잡음 비(SNR, Signal to Noise Ratio)는 입력 신호들($x'_1(n)$~$x'_N(n)$)의 SNR 보다 감소할 것이다. 잡음이 감소된 신호(f(n))는 시간 필터링부(300)에 전달된다.

[0030] 차단 매트릭스(240)는 채널 보상 필터(220)로부터 보상된 신호들($x'_1(n)$~$x'_N(n)$)을 수신한다. 차단 매트릭스(240)는 수신된 신호들($x'_1(n)$~$x'_N(n)$)로부터 잡음 신호들($Z_1(n)$~$Z_N(n)$)을 검출하여 출력하도록 구성된다. 예를 들면, 차단 매트릭스(240)는 수신된 신호들($x'_1(n)$~$x'_N(n)$) 중 인접한 채널들에 대응하는 신호들을 차감하여 잡음 신호를 검출할 것이다. 인접한 채널들에 대응하는 신호들의 음성 성분들의 위상 차이는 인접하지 않은 채널들에 대응하는 신호들의 음성 성분들의 위상 차이보다 상대적으로 작다. 따라서, 인접한 채널들에 대응하는 신호들이 차감될 때, 음성 성분이 감소된다. 즉, 차감된 신호는 잡음 신호인 것으로 이해될 수 있다. 검출된 잡음 신호들($Z_1(n)$~$Z_N(n)$)은 시간 필터링부(300)에 전달된다.

[0031] 시간 필터링부(300)는 적응적 필터 회로(310) 및 합산기(320)를 포함한다.

[0032] 적응적 필터 회로(310)는 제 1 내지 제 N 필터들(311~31N)을 포함한다. 제 1 내지 제 N 필터들(311~31N)은 공간 필터링부(200)의 차단 매트릭스(240)로부터 잡음 신호들($Z_1(n)$~$Z_N(n)$)을 각각 수신한다. 제 1 내지 제 N 필터들(311~31N)은 수신된 잡음 신호들($Z_1(n)$~$Z_N(n)$)을 필터링하도록 구성된다. 필터링된 신호들($Z'_1(n)$~$Z'_N(n)$)은 잡음이 감소된 신호(f(n))의 잡음 성분에 대응한다. 이하에서, 적응적 필터링된 신호들($Z'_1(n)$~$Z'_N(n)$)을 잡음 성분 신호들($Z'_1(n)$~$Z'_N(n)$)이라 부르기로 한다. 잡음 성분 신호들($Z'_1(n)$~$Z'_N(n)$)은 합산기(320)로 전달된다.

[0033] 합산기(320)는 공간 필터링부(200)의 고정 빔 형성기(230)로부터 잡음이 감소된 신호(f(n))를 수신하고, 제 1 내지 제 N 필터들(311~31N)로부터 잡음 성분 신호들($Z'_1(n)$~$Z'_N(n)$)을 수신한다. 합산기(320)는 잡음이 감소된 신호(f(n)) 및 잡음 성분 신호들($Z'_1(n)$~$Z'_N(n)$)을 합산한다. 상술한 바와 같이, 잡음 성분 신호들($Z'_1(n)$~$Z'_N(n)$)은 잡음이 감소된 신호(f(n))의 잡음 성분에 대응한다. 따라서, 잡음이 감소된 신호(f(n)) 및 잡음 성분 신호들($Z'_1(n)$~$Z'_N(n)$)이 합산되면, 잡음이 제거된 신호(y(n))가 생성된다.

[0034] 예시적으로, 제 1 내지 제 N 필터들(311~31N)은 잡음 신호들($Z_1(n)$~$Z_N(n)$)을 적응적 필터링(adaptive filtering)하도록 구성된다. 그리고, 제 1 내지 제 N 필터들(311~31N)은 활성화 및 비활성화가 가능한 학습 모드를 갖는다. 학습 모드가 활성화될 때, 제 1 내지 제 N 필터들(311~31N)은 잡음이 제거된 신호(y(n))에 기반하여 필터링 계수를 조절하도록 구성된다. 예를 들면, 제 1 내지 제 N 필터들(311~31N)은 정규화된 최소 평균 제곱근(normalized Least Mean Square) 알고리즘에 기반하여 필터링 계수를 갱신하도록 구성된다. 예를 들면, 제 1 내지 제 N 필터들(311~31N)은 잡음이 제거된 신호(y(n)) 중 미리 설정된 시간 윈도우에 대응하는 신호에 기반하여 필터링 계수를 갱신하도록 구성된다.

[0035] 학습 모드가 비활성화될 때, 제 1 내지 제 N 필터들(311~31N)은 필터링 계수를 유지하도록 구성된다.

[0036] 제 1 내지 제 N 필터들(311~31N)의 학습 모드는 확률적 모드 적응 제어기(400)의 제어에 따라 활성화 또는 비활성화 된다.

[0037] 확률적 모드 적응 제어기(400)는 공간 필터링부(200)로부터 적어도 두 개의 음성 신호들을 수신하도록 구성된다. 예를 들면, 확률적 모드 적응 제어기(400)는 공간 필터링부(200)의 채널 보상 필터(220)로부터 적어도 두 개의 보상된 신호들($x'_i(n)$, $x'_j(n)$)을 수신하도록 구성된다. 수신된 신호들($x'_i(n)$, $x'_j(n)$)에 기반하여, 확률적 모드 적응 제어기(400)는 수신된 신호들($x'_i(n)$, $x'_j(n)$)의 잡음 구간을 판별하도록 구성된다. 확률적 모드 적응 제어기(400)는 판별 결과에 따라 제어 신호(CTRL)를 제어한다.

[0038] 예를 들면, 잡음 구간으로 판별된 구간 동안, 확률적 모드 적응 제어기(400)는 제 1 내지 제 N 필터들(311~31N)의 학습 모드가 활성화되도록 제어 신호(CTRL)를 제어할 것이다. 비잡음 구간으로 판별된 구간 동안, 확률적 모드 적응 제어기(400)는 제 1 내지 제 N 필터들(311~31N)의 학습 모드가 비활성화 되도록 제어 신호(CTRL)를 제어할 것이다.

[0039] 상술한 바와 같은 잡음 제거 장치(100)에서, 잡음 제거 동작 성능은 보상된 신호들($x'_1(n)$~$x'_N(n)$)의 잡음 구간을 판별하는 방법에 따라 변화한다. 예를 들면, 확률적 모드 적응 제어기(400)에 의해 판별된 잡음 구간 동안 제 1 내지 제 N 필터들(311~31N)의 학습 모드가 활성화 되므로, 확률적 모드 적응 제어기(400)에 의해 판별된 잡음 구간의 정확성이 잡음 제거 장치(100)의 동작 성능을 결정한다.

[0040] 본 발명의 실시 예에 따른 잡음 제거 장치(100)의 확률적 모드 적응 제어기(400)는 보상된 신호들($x'_1(n)$~$x'_N(n)$)의 잡음 구간을 확률에 기반하여 판별하도록 구성된다. 예를 들면, 잡음 제거 장치(100)는 보상된 신호들($x'_1(n)$~ $x'_N(n)$) 중 적어도 두 개의 신호들을 수신하고, 수신된 신호들의 잡음 구간의 확률을 산출한다. 확률적 모드 적응 제어기(400)는 산출된 확률에 기반하여 잡음 구간을 판별하도록 구성된다.

[0041] 예시적으로, 확률적 모드 적응 제어기(400)는 수신된 신호들을 프레임 단위로 처리할 것이다. 예를 들면, 확률적 모드 적응 제어기(400)는 수신된 신호들의 특정 구간이 잡음 구간일 확률을 산출하고, 산출된 확률에 기반하여 특정 구간이 잡음 구간인지의 여부를 판별할 것이다.

[0042] 도 3은 도 1의 확률적 모드 적응 제어기(400)의 동작 방법을 보여주는 순서도이다. 도 1 및 도 3을 참조하면,

S110 단계에서, 확률적 모드 적응 제어기(400)는 미리 설정된 수의 채널 보상 필터(220)의 출력들을 수신한다. 예시적으로, 확률적 모드 적응 제어기(400)는 채널 보상 필터(220)의 출력들 중 제 i 채널에 대응하는 보상 신호($x'_i(n)$) 및 제 j 채널에 대응하는 보상 신호($x'_j(n)$)을 수신할 것이다. 예시적으로, 제 i 채널 및 제 j 채널은 음성 수신기(210)의 마이크들(MC_1~MC_N) 중 미리 설정된 위치의 마이크들(MC_i, MC_j)에 대응할 것이다.

[0043] 예시적으로, 확률적 모드 적응 제어기(400)는 채널 보상 필터(220)의 출력들 중 두 개를 수신하는 것으로 한정되지 않는다. 확률적 모드 적응 제어기(400)는 채널 보상 필터(220)의 출력들 중 적어도 두 개를 수신하도록 구성될 수 있다.

[0044] 예시적으로, 확률적 모드 적응 제어기(400)는 채널 보상 필터(220)의 출력들을 수신하는 것으로 한정되지 않는다. 예를 들면, 확률적 모드 적응 제어기(400)는 음성 신호들($x_1(n)$~$x_N(n)$) 중 적어도 두 개를 수신하도록 구성될 수 있다. 예를 들면, 잡음 제거 장치(100)에서 채널 보상 필터(220)가 제거되는 경우, 확률적 모드 적응 제어기(400)는 음성 신호들($x_1(n)$~$x_N(n)$) 중 적어도 두 개를 수신하도록 구성될 것이다.

[0045] S120 단계에서, 확률적 모드 적응 제어기(400)는 수신된 출력들의 선형 예측 잔차 신호를 추정한다. 예를 들면, 확률적 모드 적응 제어기(400)는 수신된 출력들의 제 t 프레임의 선형 예측 잔차 신호를 추정할 것이다. 예시적으로, 제 i 채널에 대응하는 선형 예측 잔차 신호는 수학식 1에 기반하여 산출된다.

수학식 1

[0046]
$$e_i(n) = x'_i(n) - \sum_{k=1}^{P} a_i(k)x'_i(n-k)$$

[0047] 여기에서, $a_i(k)$는 제 i 보상 신호($x'_i(n)$)에서 추정한 선형 예측 계수이며, P는 선형 예측 차수이다.

[0048] 예시적으로, $a_i(k)$는 제 i 보상 신호($x'_i(n)$)의 비정상 잡음 성분에 대응하는 계수이다. 즉, 수학식 1에서, $\sum_{k=1}^{P} a_i(k)x'_i(n-k)$는 제 i 보상 신호($x_i(n)$)의 비정상 잡음 성분에 대응한다. 예를 들면, 비정상 잡음 성분은 음성의 반사에 의한 잔향을 포함할 것이다. 즉, 선형 예측 잔차 신호($e_i(n)$)는 제 i 보상 신호($x'_i(n)$)로부터 비정상 잡음 성분을 제거된 정상 잡음 성분일 것이다.

[0049] 확률적 모드 적응 제어기(400)는 제 i 및 제 j 보상 신호들($x'_i(n)$, $x'_j(n)$)에 대응하는 제 i 및 제 j 선형 예측 잔차 신호들($e_i(n)$, $e_j(n)$)을 각각 추정할 것이다. 확률적 모드 적응 제어기(400)가 적어도 두 개의 보상 신호들을 수신하는 경우, 확률적 모드 적응 제어기(400)는 적어도 두 개의 선형 예측 잔차 신호들을 추정할 것이다.

[0050] S130 단계에서, 확률적 모드 적응 제어기(400)는 추정된 제 i 및 제 j 선형 예측 잔차 신호들($e_i(n)$, $e_j(n)$)에 기반하여 상호 상관도를 산출한다. 예시적으로, 확률적 모드 적응 제어기(400)는 제 t 프레임에 대응하는 상호 상관도를 산출할 것이다. 상호 상관도는 수학식 2에 기반하여 산출된다.

수학식 2

[0051]
$$c_t = \frac{\sum_{n=0}^{L-1} e_{i,t}(n)e_{j,t}(n)}{\sqrt{\sum_{n=0}^{L-1} e^2{}_{i,t}(n)}\sqrt{\sum_{n=0}^{L-1} e^2{}_{j,t}(n)}}$$

[0052] 여기에서, $e_{i,t}(n)$ 및 $e_{j,t}(n)$ 은 제 t 프레임에 대응하는 제 i 및 제 j 선형 예측 잔차 신호들($e_i(n)$, $e_j(n)$)을 각각 나타내고, L은 프레임 길이를 나타낸다.

[0053] 산출된 상호 상관도(c_t)는 제 i 및 제 j 채널에 대응하는 정상 잡음들 사이의 상호 관련도를 나타낼 것이다.

[0054] S140 단계에서, 확률적 모드 적응 제어기(400)는 산출된 상호 상관도(c_t)에 기반하여 로그 우도를 산출한다. 예시적으로, 확률적 모드 적응 제어기(400)는 제 t 프레임에 대응하는 로그 우도를 산출할 것이다. 로그 우도는 수학식 3에 기반하여 산출된다.

수학식 3

[0055]
$$LL_t = \ln\left(\frac{1}{\sigma_{noise}\sqrt{2\pi}}\exp\left(-\frac{(c_t-\mu_{noise})^2}{2\sigma_{noise}^2}\right)\right) = G - \frac{(c_t-\mu_{noise})^2}{2\sigma_{noise}^2}$$

[0056] 여기에서, μ_{noise}는 잡음구간 상호 상관도의 평균을 나타내고, σ_{noise}는 잡음구간 상호 상관도의 표준편차를 나타낸다.

[0057] G는 수학식 4와 같이 표현된다.

수학식 4

[0058]
$$G = \ln\left(\sigma_{noise}\sqrt{2\pi}\right)$$

[0059] 예시적으로, 확률적 모드 적응 제어기(400)는 제 i 및 제 j 보상 신호들($x'_i(t)$, $x'_j(t)$)의 초기 구간을 잡음 구간으로 간주하고, 잡음 구간의 상호 상관도의 평균(μ_{noise}) 및 표준편차(σ_{noise})를 산출할 것이다. 그리고, 산출된 잡음 구간의 상호 상관도의 평균(μ_{noise}) 및 표준편차(σ_{noise})에 기반하여, 로그 우도(LL_t, 예를 들면, 정규 분포)를 추정할 것이다. 로그 우도(LL_t)는 제 t 프레임이 잡음 구간일 확률을 나타낸다.

[0060] S150 단계에서, 확률적 모드 적응 제어기(400)는 산출된 로그 우도(LL_t)가 제 1 임계값보다 큰 지 판별한다. 산출된 로그 우도(LL_t)가 제 1 임계값보다 크면, 확률적 모드 적응 제어기(400)는 제 t 프레임을 잡음 구간으로 판별한다. 이후에, 확률적 모드 적응 제어기(400)는 제 1 내지 제 N 필터들(311~31N)의 필터링 계수 학습 모드가 활성화되도록 제어 신호(CTRL)를 제어한다.

[0061] 산출된 로그 우도(LL_t)가 제 1 임계값보다 작으면, 확률적 모드 적응 제어기(400)는 제 t 프레임을 비잡음 구간, 즉 음성 구간으로 판별한다. 이후에, 확률적 모드 적응 제어기(400)는 제 1 내지 제 N 필터들(311~31N)의 필터링 계수 학습 모드가 비활성화되도록 제어 신호(CTRL)를 제어한다.

[0062] 상술한 바와 같이, 확률적 모드 적응 제어기(400)는 특정 프레임이 잡음 구간일 확률에 기반하여 특정 프레임이 잡음 구간인지 비잡음 구간인지 판별한다. 따라서, 확률적 모드 적응 제어기(400)가 잡음 구간을 판별하는 정확도가 향상된다.

[0063] 도 4는 도 1의 채널 보상 필터(220)의 동작 방법을 보여주는 순서도이다. 예시적으로, 채널 보상 필터(220)가 제 i 채널에 대응하는 필터링 계수를 획득하는 방법이 도시되어 있다. 도 1 및 도 4를 참조하면, S210 단계에서, 채널 보상 필터(220)는 제 i 채널 신호를 수신한다. 예를 들면, 채널 보상 필터(220)는 음성 수신부(210)로부터 제 i 음성 신호($x_i(n)$)를 수신할 것이다. 예시적으로, 채널 보상 필터(220)는 수신된 제 i 음성 신호($x_i(n)$)를 미리 설정된 구간 단위로 처리할 것이다.

[0064] S220 단계에서, 채널 보상 필터(220)는 수신된 제 i 음성 신호($x_i(n)$)를 푸리에(Fourier) 변환한다. S220 단계

가 수행되면, 수신된 제 i 음성 신호($x_i(n)$)의 주파수 응답이 획득될 것이다. 예를 들면, 수신된 제 i 음성 신호($x_i(n)$)의 특정 구간에 대응하는 주파수 응답이 획득될 것이다.

[0065] S230 단계에서, 채널 보상 필터(220)는 수신된 제 i 음성 신호($x_i(n)$)의 특정 구간이 음성 구간인지 판별한다. 예시적으로, 채널 보상 필터(220)는 수신된 제 i 음성 신호($x_i(n)$)의 특정 구간의 에너지에 기반하여 특정 구간이 음성 구간인지 판별할 것이다. 예를 들면, 채널 보상 필터(220)는 수신된 제 i 음성 신호($x_i(n)$)의 특정 구간의 에너지가 제 2 임계값보다 클 때, 특정 구간을 음성 구간으로 판별할 것이다.

[0066] 예를 들면, 채널 보상 필터(220)는 수신된 제 i 음성 신호($x_i(n)$)의 초기 구간(예를 들면, 200msec)을 초기 잡음 구간으로 간주할 것이다. 특정 구간의 에너지가 초기 잡음 구간의 에너지보다 10dB 내지 20dB 크면, 채널 보상 필터(220)는 특정 구간을 음성 구간으로 판별할 것이다.

[0067] 특정 구간이 음성 구간으로 판별되면, S240 단계에서 채널 보상 필터(220)는 수신된 제 i 음성 신호($x_i(n)$)의 특정 구간의 주파수 응답을 이용하여 음성 통계를 산출한다. 예를 들면, 산출되는 음성 통계는 제 i 음성 신호($x_i(n)$)가 수신되는 동안 누적될 것이다. 음성 통계는 수학식 5에 기반하여 산출될 것이다.

수학식 5

[0068] $$V1 = E\left[\left|X_i(k)\right|^2\right]$$

[0069] $$V2 = E\left[X_R(k)X_i^*(k)\right]$$

[0070] 여기에서, $X_i(k)$는 제 i 음성 신호($x_i(n)$)의 특정 구간의 주파수 응답이고, $X_R(k)$는 기준 채널에 대응하는 음성 신호($x_R(n)$)의 특정 구간의 주파수 응답이다. 예시적으로, 기준 채널은 음성 수신기(210)의 미리 설정된 위치의 마이크에 대응할 것이다.

[0071] 특정 구간이 비음성 구간, 즉 잡음 구간으로 판별되면, S250 단계에서 채널 보상 필터(220)는 수신된 제 i 음성 신호($x_i(n)$)의 특정 구간의 주파수 응답을 이용하여 잡음 통계를 산출한다. 예를 들면, 산출되는 잡음 통계는 제 i 음성 신호($x_i(n)$)가 수신되는 동안 누적될 것이다. 잡음 통계는 수학식 6에 기반하여 산출될 것이다.

수학식 6

[0072] $$N1 = E\left[\left|V_i(k)\right|^2\right]$$

[0073] $$N2 = E\left[V_i(k)V_R(k)\right]$$

[0074] 여기에서, $V_i(k)$는 제 i 음성 신호($x_i(n)$)의 특정 구간의 주파수 응답이고, $V_R(k)$는 기준 채널에 대응하는 음성 신호($x_R(n)$)의 특정 구간의 주파수 응답이다.

[0075] S260 단계에서, 채널 보상 필터(220)는 제 i 음성 신호($x_i(n)$)의 입력이 종료되는지 판별한다. 제 i 음성 신호($x_i(n)$)의 입력이 종료되지 않으면, S210 단계에서, 채널 보상 필터(220)는 제 i 음성 신호($x_i(n)$)의 다음 구간을 수신할 것이다. 제 i 음성 신호($x_i(n)$)의 입력이 종료되면, S270 단계가 수행된다.

[0076] S270 단계에서, 채널 보상 필터(220)는 채널 보상 응답을 추정한다. 채널 보상 응답은 비용 함수에 기반하여 추

정된다. 비용 함수는 수학식 7과 같이 정의된다.

수학식 7

[0077] $$J_i = E\left[\left|D_R(k) - w_i(k)D_i(k)\right|^2\right]$$

[0078] 여기에서, J_i는 제 i 음성 신호($x_i(n)$)의 보상 응답을 판별하기 위한 비용 함수이고, $D_R(k)$는 기준 채널에 대응하는 잡음 없는 음성 신호의 주파수 응답을 나타내고, $D_i(k)$는 제 i 음성 신호($x_i(n)$)의 잡음 없는 주파수 응답을 나타낸다. 그리고, $w_i(k)$는 제 i 음성 신호($x_i(n)$)를 보상하는 주파수 이득을 나타낸다.

[0079] 기준 채널에 대응하는 신호의 크기 및 제 i 음성 신호($x_i(n)$)의 크기의 차이가 최소화되도록 주파수 이득($w_i(k)$)이 설정되면, 음성 수신기(210)의 마이크들(MC_1~MC_N)의 위치 차이로 인한 신호들의 차이는 보상될 것이다. 따라서, 제 i 음성 신호($x_i(n)$)를 보상하는 최적화된 주파수 이득은 수학식 8과 같이 정의된다.

수학식 8

[0080] $$w_i^{opt} = \frac{E\left[D_R(k)D_i^*(k)\right]}{E\left[\left|D_i(k)\right|\right]}$$

[0081] 제 i 음성 신호($x_i(n)$)는 잡음 없는 음성 신호 및 잡음 신호로 구성된다. 마찬가지로, 기준 채널에 대응하는 음성 신호($x_R(n)$)는 잡음 없는 음성 신호 및 잡음 신호로 구성된다. 제 i 음성 신호($x_i(n)$) 및 기준 채널에 대응하는 음성 신호($x_R(n)$)는 수학식 9와 같이 표현될 수 있다.

수학식 9

[0082] $$X_i(k) = D_i(k) + V_i(k)$$

[0083] $$X_R(k) = D_R(k) + V_R(k)$$

[0084] 수학식 8 및 수학식 9를 결합하면, 최적화된 주파수 이득($w_i^{opt}(k)$)은 수학식 10과 같이 정리될 수 있다.

수학식 10

[0085] $$w_i^{opt}(k) = \frac{E\left[X_R(k)X_i^*(k)\right] - E\left[V_i(k)V_R(k)\right]}{E\left[\left|X_i(k)\right|^2\right] - E\left[\left|V_i(k)\right|^2\right]} = \frac{V2 - N2}{V1 - N1}$$

[0086] 즉, S240 단계에서 산출된 음성 통계들(V1, V2) 및 S250 단계에서 산출된 잡음 통계들(N1, N2)에 기반하여 최적화된 주파수 이득($w_i^{opt}(k)$)이 산출된다.

[0087] S280 단계에서, 채널 보상 필터(220)는 산출된 최적 주파수 이득($w_i^{opt}(k)$)을 역 푸리에(inverse Fourier) 변환한다. S290 단계에서, 채널 보상 필터(220)는 역 푸리에 변환의 결과를 제 i 채널에 대응하는 필터링 계수로 획득한다. 이후에, 획득된 필터링 계수에 기반하여, 제 i 음성 신호($x_i(n)$)가 필터링될 것이다. 필터링된 결과는 보상된 제 i 음성 신호($x'_i(n)$)로 출력될 것이다.

[0088] 예시적으로, 채널 보상 필터(220)는 제 1 내지 제 N 음성 신호들($x_1(n)$~$x_N(n)$)에 대하여 도 4를 참조하여 설명된 필터링 계수 획득 방법을 수행할 것이다.

[0089] 상술한 바와 같이, 채널 보상 필터(220)는 음성 통계 및 잡음 통계를 산출하고, 산출된 음성 및 잡음 통계들에 기반하여 최적화된 주파수 이득($w_i^{opt}(k)$)을 획득한다. 따라서, 음성 수신기(210)의 마이크들(MC_1~MC_N)의 위치 차이로 인한 음성 신호들($x_1(n)$~$x_N(n)$)의 차이가 보상된다. 즉, 잡음 제거 장치(100)의 동작 성능이 향상된다.

[0090] 도 5는 도 1 내지 도 4를 참조하여 설명된 잡음 제거 장치(100)의 실험 환경(500)을 보여주는 블록도이다. 도 5를 참조하면, 실험은 가로 6m 그리고 세로 5m의 공간(510)에서 실행되었다. 음성 수신기(210)는 8채널을 갖는 것으로 설정되었으며, 음성 수신기(210)는 실험 공간(510)의 하단 벽으로부터 1.5m 이격된 위치에 설치되었다. 음성 소스(520)는 음성 수신기(210)로부터 2m 이격된 위치에 설치되었다. 그리고, 잡음 소스(530)는 음성 수신기(210)와 2m 이격되며, 음성 소스(520) 및 음성 수신기(540)를 연결하는 축선과 45도의 각도를 갖는 위치에 설치되었다.

[0091] 예시적으로, 데이터 표본화율은 16kHz로 설정되었고, 해상도는 16bit으로 설정되었다. 선형 예측 분석 차수(P)는 18로 설정되었고, 제 1 내지 제 N 필터들(311~31N) 각각의 길이는 127로 설정되었다. 제 1 내지 제 N 필터들(311~31N)의 학습 알고리즘은 정규화된 최소 평균 제곱근(normalized Least Mean Square) 알고리즘이 적용되었다. 제 1 내지 제 N 필터들(311~31N)의 학습률은 0.1로 설정되었다. 그리고, 잡음 구간을 판별하기 위한 로그 우도(LL_t)의 제 1 임계값은 잡음 구간의 평균 로그 우도에서 10dB를 감한 값으로 설정되었다.

[0092] 도 6은 도 5의 실험 환경(500)에서 도 1의 잡음 제거 장치(100)를 실험한 제 1 결과를 보여주는 테이블이다. 도 1, 도 5, 그리고 도 6을 참조하면, 입력 신호들($x_1(n)$~$x_N(n)$)의 신호대 잡음비(SNR, signal to noise ratio)를 -5dB, 0dB, 5dB, 10dB, 그리고 15dB로 변화하며 실험이 수행되었다.

[0093] 변화하는 SNR을 갖는 입력 신호들($x_1(n)$~$x_N(n)$)에 기반하여, 잡음 섞인 입력에 따른 잡음 구간의 인식률, 고정 빔포밍 결과의 에너지 궤적을 이용한 잡음 구간의 인식률, 고정 빔포밍 결과와 GSC(Generalized Sidelibe Canceller) 출력간의 상호 상관도를 이용하는 방식에 따른 잡음 구간의 인식률, 그리고 본 발명에 따른 잡음 구간의 인식률이 도 6에 기재되어 있다.

[0094] 도 6에 기재된 바와 같이, 기존의 방법들에 따른 잡음 구간의 인식률과 비교할 때, 본 발명에 따른 잡음 구간의 인식률이 더 높다. 특히, SNR이 낮은 열악한 환경에서, 본 발명에 따른 잡음 구간의 인식률이 기존의 방법들에 따른 잡음 구간의 인식률보다 월등한 것으로 나타난다.

[0095] 도 7은 도 5의 실험 환경(500)에서 도 1의 잡음 제거 장치(100)를 실험한 제 2 결과를 보여주는 테이블이다. 도 1, 도 5, 그리고 도 7을 참조하면, 입력 신호들($x_1(n)$~$x_N(n)$)의 신호대 잡음비(SNR, signal to noise ratio)를 -5dB, 0dB, 5dB, 10dB, 그리고 15dB로 변화하며 실험이 수행되었다.

[0096] 변화하는 SNR을 갖는 입력 신호들($x_1(n)$~$x_N(n)$)에 기반하여, 잡음 섞인 입력에 따른 출력 신호($y(n)$)의 SNR, 고정 빔포밍 결과의 에너지 궤적을 이용한 출력 신호($y(n)$)의 SNR, 고정 빔포밍 결과와 GSC(Generalized Sidelibe Canceller) 출력간의 상호 상관도를 이용하는 방식에 따른 출력 신호($y(n)$)의 SNR, 그리고 본 발명에 따른 출력 신호($y(n)$)의 SNR이 도 7에 기재되어 있다.

[0097] 도 7에 기재된 바와 같이, 기존의 방법들에 따른 출력 신호($y(n)$)의 SNR과 비교할 때, 본 발명에 따른 출력 신호($y(n)$)의 SNR이 더 높다.

[0098] 도 8은 도 1 내지 도 7을 참조하여 설명된 잡음 제거 장치(100)를 포함하는 음성 인식 시스템(1000)을 보여주는 블록도이다. 도 8을 참조하면, 음성 인식 시스템(1000)은 잡음 제거 장치(100) 및 음성 인식 장치(1100)를 포함

한다.

[0099] 잡음 제거 장치(100)는 도 1 내지 도 7을 참조하여 설명된 바와 같이 구성된다. 예를 들면, 잡음 제거 장치(100)는 음성 신호가 잡음 구간일 확률에 기반하여 음성 신호의 잡음 구간을 판별할 것이다. 그리고, 잡음 제거 장치(100)는 수신 신호의 음성 구간 신호의 통계 및 비음성 구간 신호의 통계에 기반하여 수신 신호를 필터링할 것이다. 잡음 제거 장치(100)의 출력 신호(y(n))는 음성 인식 장치(1100)로 전달된다.

[0100] 음성 인식 장치(1100)는 잡음 제거 장치(100)로부터 출력 신호(y(n))를 수신한다. 음성 인식 장치(1100)는 수신된 출력 신호(y(n))에 기반하여 음성 인식 동작을 수행할 것이다.

[0101] 도 6 및 도 7을 참조하여 설명된 바와 같이, 본 발명의 실시 예에 따른 잡음 제거 장치(100)는 기존의 잡음 제거 방법들보다 향상된 잡음 제거 성능을 제공한다. 따라서, 음성 인식 시스템(1000)의 음성 인식률이 향상될 것이다.

[0102] 도 9는 도 8의 음성 인식 시스템(1000)을 포함하는 컴퓨팅 시스템(2000)을 보여주는 블록도이다. 도 9를 참조하면, 컴퓨팅 시스템(2000)은 시스템 버스(2100), 프로세서(2200), 메모리(2300), 대용량 저장장치(2400), 사용자 출력 인터페이스(2500), 그리고 음성 인식 시스템(1000)을 포함한다.

[0103] 시스템 버스(2100)는 컴퓨팅 시스템(2000)의 구성 요소들 사이에 통신 채널을 제공한다.

[0104] 프로세서(2200)는 컴퓨팅 시스템(2000)의 제반 동작을 제어하도록 구성된다. 예를 들면, 프로세서(2200)는 응용(Application) 및 운영체제(OS, Operating System)를 구동할 것이다.

[0105] 메모리(2300)는 컴퓨팅 시스템(2000)의 동작 메모리, 버퍼 메모리, 또는 캐시 메모리로 이용될 것이다. 예를 들면, 메모리(2300)는 SRAM, DRAM, SDRAM 등과 같은 고속 메모리를 포함할 것이다.

[0106] 대용량 저장장치(2400)는 컴퓨팅 시스템(2000)에 대용량 저장 수단을 제공한다. 예를 들면, 대용량 저장장치(2400)는 불휘발성 저장장치일 것이다. 예를 들면, 대용량 저장장치(2400)는 하드 디스크 드라이브(HDD, Hard Disk Drive), 반도체 드라이브(SSD, Solid State Drive), 또는 플래시 메모리, MRAM (magnetic RAM), FeRAM (Ferroelectric RAM), PRAM (Phase-change RAM), RRAM (Resistive RAM), PROM, EPROM, EEPROM 등과 같은 불휘발성 메모리를 포함할 것이다.

[0107] 사용자 출력 인터페이스(2500)는 사용자에게 데이터를 전달하는 다양한 수단을 포함할 것이다. 예를 들면, 사용자 출력 인터페이스(2500)는 모니터, 프린터, 스피커, LED, 모터 등과 같은 다양한 장치를 포함할 것이다.

[0108] 음성 인식 시스템(1000)은 컴퓨팅 시스템(2000)의 사용자 입력 인터페이스로 동작할 것이다. 예를 들면, 컴퓨팅 시스템(2000)은 음성 인식 시스템(1000)을 통해 사용자로부터 데이터를 수신할 것이다.

[0109] 도 6 및 도 7을 참조하여 설명된 바와 같이, 본 발명의 실시 예에 따른 잡음 제거 장치(100)는 기존의 잡음 제거 방법들보다 향상된 잡음 제거 성능을 제공한다. 따라서, 컴퓨팅 시스템(2000)이 사용자로부터 데이터를 정상적으로 수신할 확률이 증가한다. 즉, 컴퓨팅 시스템(2000)의 신뢰도가 증가할 것이다.

[0110] 예시적으로, 컴퓨팅 시스템(2000)은 음성 인식 시스템(1000) 이외의 다양한 사용자 입력 인터페이스(미도시)를 추가적으로 포함할 수 있다. 예를 들면, 컴퓨팅 시스템(2000)은 키보드, 마우스, 터치 패널, 카드 리더, 센서, 카메라 등과 같은 다양한 사용자 입력 인터페이스를 더 포함할 수 있다.

[0111] 예시적으로, 컴퓨팅 시스템(2000)은 다른 컴퓨팅 시스템(미도시)과 통신할 수 있는 통신 모뎀을 추가적으로 포함할 수 있다. 예를 들면, 컴퓨팅 시스템(2000)은 ethernet, WiFi, WCDMA, CDMA 2000, Wibro, WiMAX 등과 같은 다양한 프로토콜을 제공하는 모뎀을 더 포함할 수 있다.

[0112] 예시적으로, 컴퓨팅 시스템(2000)은 컴퓨터, UMPC (Ultra Mobile PC), 워크스테이션, 넷북(net-book), PDA (Personal Digital Assistants), 포터블(portable) 컴퓨터, 웹 타블렛(web tablet), 무선 전화기(wireless phone), 모바일 폰(mobile phone), 스마트폰(smart phone), e-북(e-book), PMP (portable multimedia player), 휴대용 게임기, 네비게이션(navigation) 장치, 블랙박스(black box), 디지털 카메라(digital camera), DMB (Digital Multimedia Broadcasting) 재생기, 디지털 음성 녹음기(digital audio recorder), 디지털 음성 재생기(digital audio player), 디지털 영상 녹화기(digital picture recorder), 디지털 영상 재생기(digital picture player), 디지털 동영상 녹화기(digital video recorder), 디지털 동영상 재생기(digital video player), 정보를 무선 환경에서 송수신할 수 있는 장치, 홈 네트워크를 구성하는 다양한 전자 장치들 중 하나, 컴퓨터 네트워크를 구성하는 다양한 전자 장치들 중 하나, 또는 텔레매틱스 네트워크를 구성하는 다양한

전자 장치들 중 하나로 제공될 수 있다.

[0113] 본 발명의 상세한 설명에서는 구체적인 실시 예에 관하여 설명하였으나, 본 발명의 범위와 기술적 사상에서 벗어나지 않는 한도 내에서 여러 가지 변형이 가능하다. 그러므로 본 발명의 범위는 상술한 실시 예에 국한되어 정해져서는 안되며 후술하는 특허청구범위뿐만 아니라 이 발명의 특허청구범위와 균등한 것들에 의해 정해져야 한다.

부호의 설명

[0114] 100 : 잡음 제거 장치

200 : 공간 필터링부

220 : 채널 보상 필터

300 : 시간 필터링부

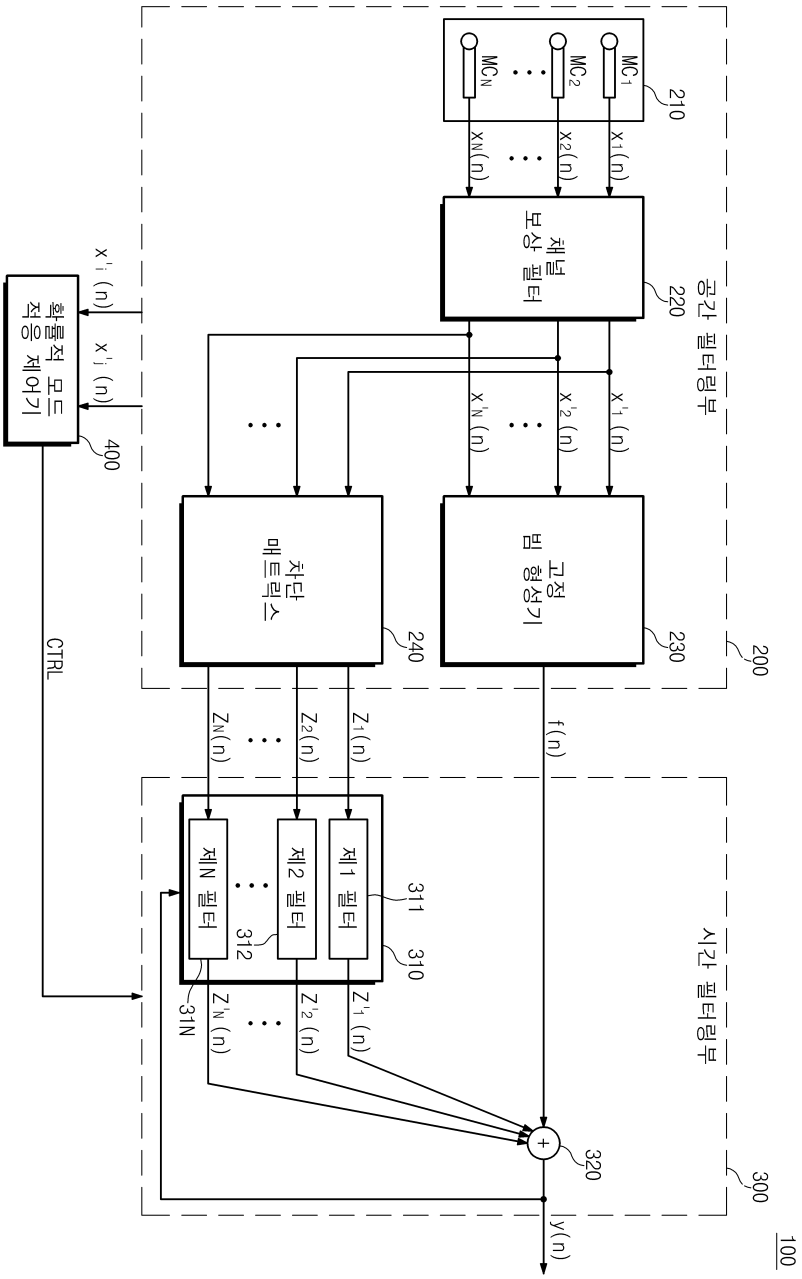
400 : 확률적 모드 적응 제어기

1000 : 음성 인식 시스템

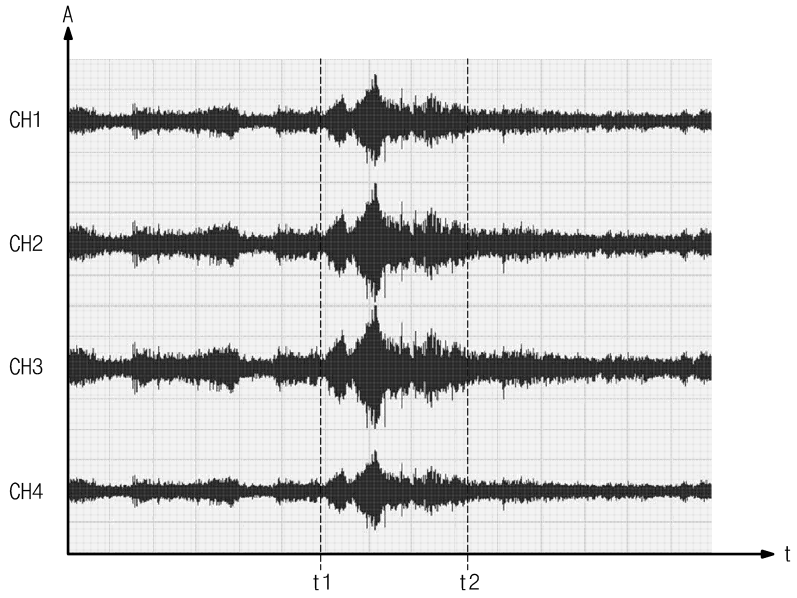
2000 : 컴퓨팅 시스템

도면

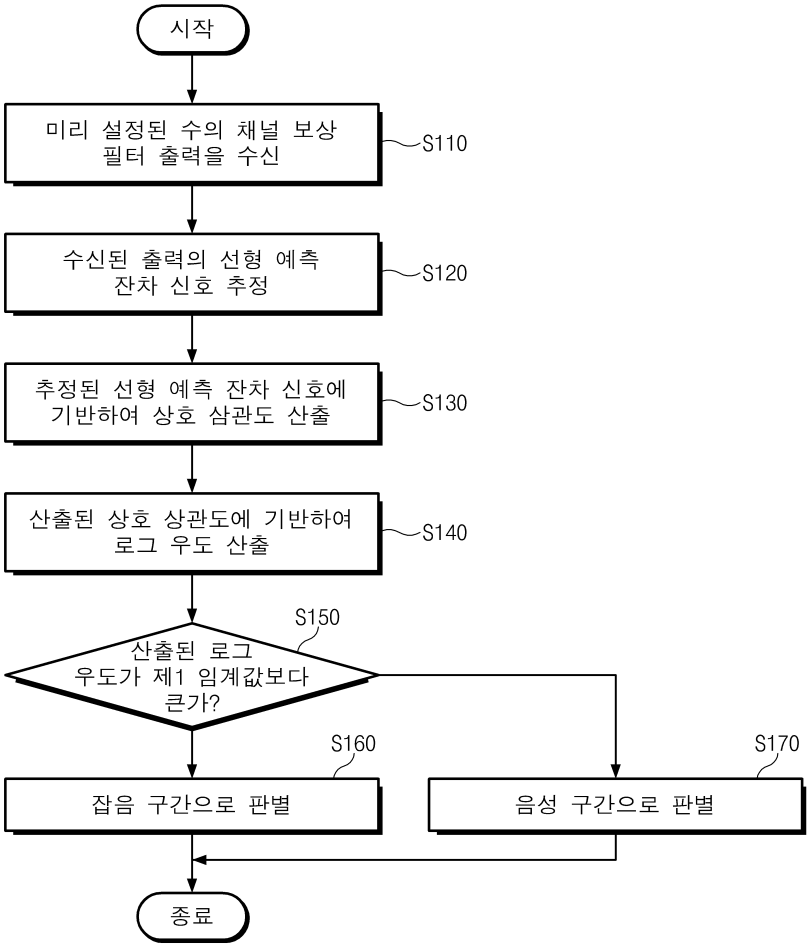
도면1


MC1
MC2
MCN
210
x1(n)
x2(n)
xN(n)
채널 보상 필터
220
공간 필터링부
x'1(n)
x'2(n)
x'N(n)
고정 빔 형성기
230
차단 매트릭스
240
200
f(n)
Z1(n)
Z2(n)
ZN(n)
제1 필터
제2 필터
제N 필터
311
310
312
31N
Z'1(n)
Z'2(n)
Z'N(n)
320
시간 필터링부
300
y(n)
100
CTRL
확률적 적응 제어기
400
x'i(n)
x'j(n)

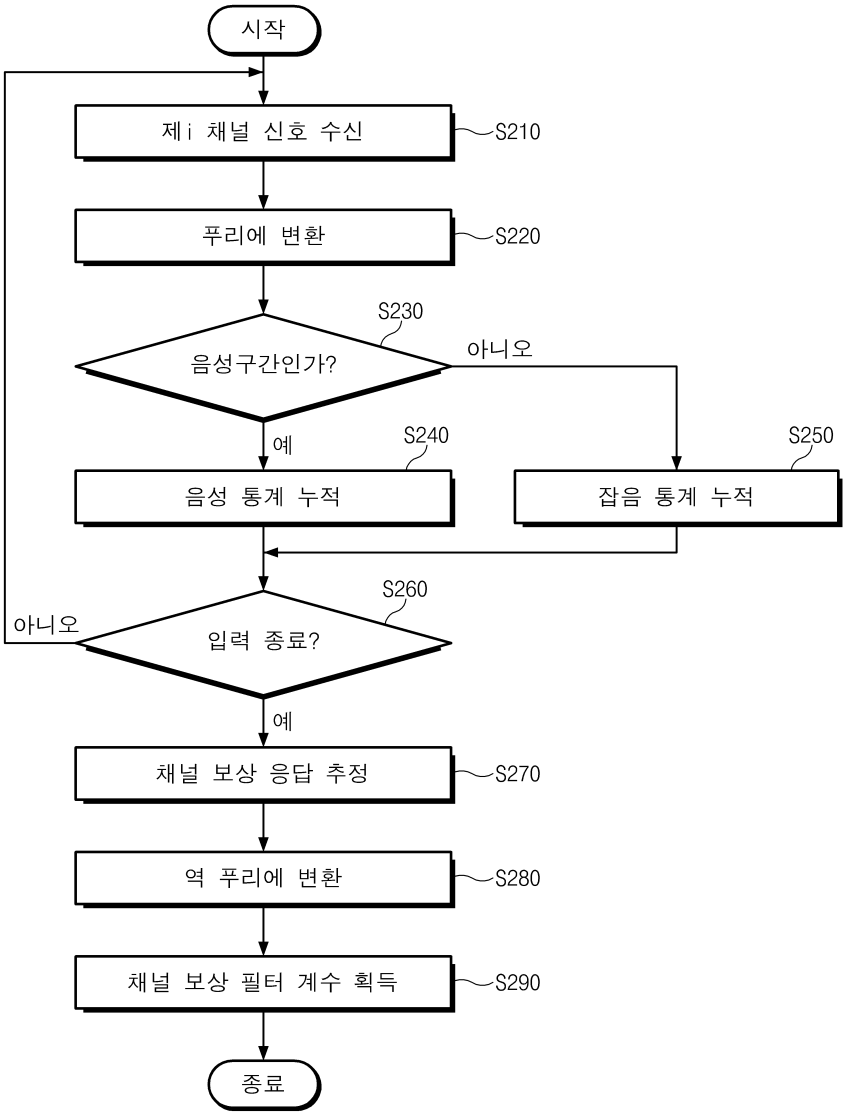
도면2


A
CH1
CH2
CH3
CH4
t
t1
t2

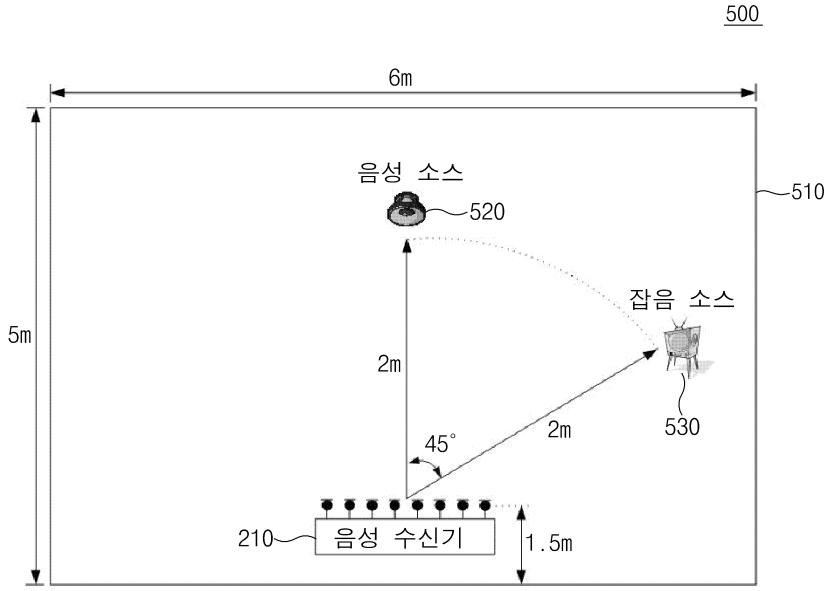
도면3


시작
미리 설정된 수의 채널 보상 필터 출력을 수신
S110
수신된 출력의 선형 예측 잔차 신호 추정
S120
추정된 선형 예측 잔차 신호에 기반하여 상호 삼관도 산출
S130
산출된 상호 상관도에 기반하여 로그 우도 산출
S140
S150
산출된 로그 우도가 제1 임계값보다 큰가?
S160
잡음 구간으로 판별
S170
음성 구간으로 판별
종료

도면4


시작
제i 채널 신호 수신
S210
푸리에 변환
S220
S230
음성구간인가?
아니오
예
S240
음성 통계 누적
S250
잡음 통계 누적
S260
입력 종료?
아니오
예
채널 보상 응답 추정
S270
역 푸리에 변환
S280
채널 보상 필터 계수 획득
S290
종료

도면5



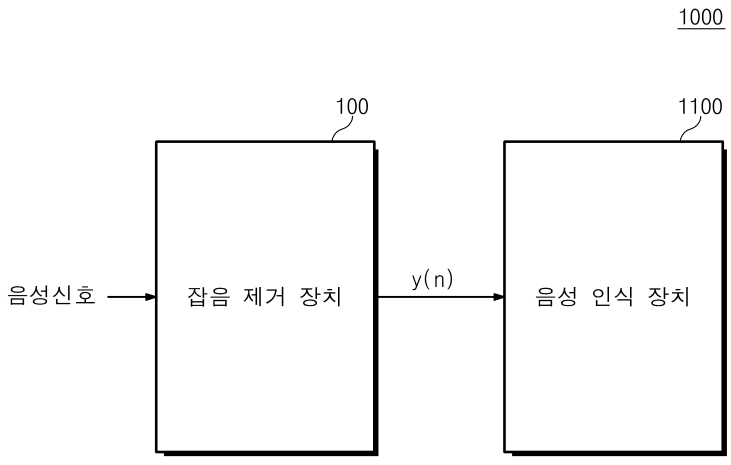
도면6

입력 신호대 잡음비(dB)	잡음 섞인 입력에 따른 인식률(%)	고정빔포밍 결과의 에너지 궤적을 이용한 인식률(%)	고정빔포밍 결과와 GSC 출력간의 상호 상관도를 이용하는 방식에 따른 인식률(%)	본 발명에 따른 인식률(%)
-5	0.67	27.1	35.7	50.66
0	3.37	64.4	72.35	78.1
5	25.17	87.0	88.05	88.72
10	70.13	89.4	90.27	91.37
15	88.72	93.8	92.26	92.25

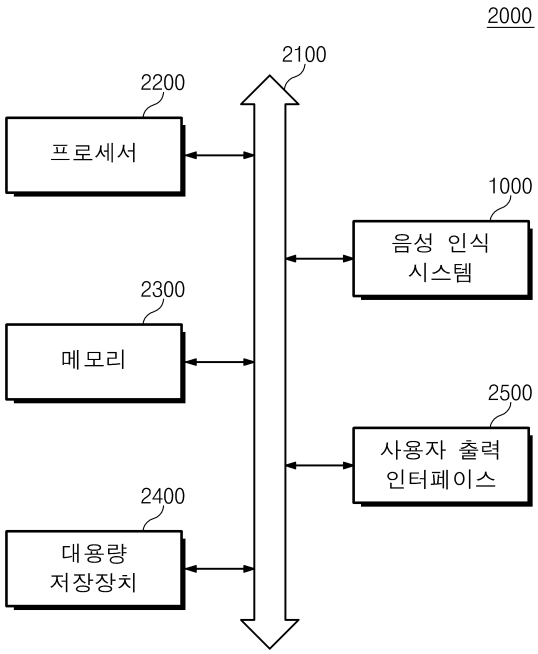
도면7

입력 신호대 잡음비(dB)	잡음 섞인 입력에 따른 SNR(dB)	고정빔포밍 결과의 에너지 궤적을 이용한 SNR(dB)	고정빔포밍 결과와 GSC 출력간의 상호 상관도를 이용하는 방식에 따른 SNR(dB)	본 발명에 따른 SNR(dB)
-5	-5.00	1.43	2.54	4.41
0	0.00	7.42	8.68	11.42
5	5.00	11.66	12.76	15.17
10	10.00	17.53	18.42	20.99
15	15.00	22.13	23.02	25.50

도면8



도면9


2000
2100
2200
프로세서
1000
음성 인식
시스템
2300
메모리
2500
사용자 출력
인터페이스
2400
대용량
저장장치