



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2012년08월07일
(11) 등록번호 10-1171494
(24) 등록일자 2012년07월30일

(51) 국제특허분류(Int. Cl.)
H04B 1/10 (2006.01)
(21) 출원번호 10-2010-7008480
(22) 출원일자(국제) 2008년10월01일
심사청구일자 2010년04월19일
(85) 번역문제출일자 2010년04월19일
(65) 공개번호 10-2010-0054873
(43) 공개일자 2010년05월25일
(86) 국제출원번호 PCT/US2008/078395
(87) 국제공개번호 WO 2009/051959
국제공개일자 2009년04월23일
(30) 우선권주장
11/874,263 2007년10월18일 미국(US)
(56) 선행기술조사문헌
WO2004053839 A1*
WO2004083884 A2*
US20070030982 A1*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
모토로라 모빌리티, 인크.
미국 60048 일리노이주 리버티빌 노쓰 유에스 하
이웨이 45 600
(72) 발명자
주레크, 로버트, 에이.
미국 60002 일리노이주 안티오츠 어텀 드라이브
1055
엑셀로드, 제프리, 엠.
미국 60025 일리노이주 글렌뷰 파인 스트리트
1238
(뒷면에 계속)
(74) 대리인
백만기, 정은진, 양영준

전체 청구항 수 : 총 10 항

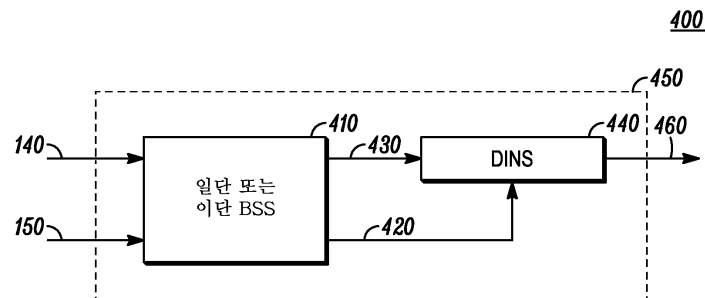
심사관 : 정성윤

(54) 발명의 명칭 **강인한 두 마이크로폰 잡음 억제 시스템**

(57) 요약

음성 신호를 잡음 음향 환경으로부터 분리하기 위한 시스템, 방법 및 장치가 개시된다. 분리 프로세스는 방향성 필터링, 블라인드 소스 분리(blind source separation), 및 듀얼 입력 스펙트럼 감산 잡음 억제기(dual input spectral subtraction noise suppressor)를 포함할 수 있다. 입력 채널은 두 개의 전방향성 마이크로폰들을 포함할 수 있고, 이들 마이크로폰들의 출력은 위상 지연 필터링을 이용하여 처리되어 음성 및 잡음 빔형상을 형성한다. 또한, 빔형상은 주파수 보정될 수 있다. 전방향성 마이크로폰들은 실질적으로 잡음만 있는 하나의 채널과, 잡음과 음성이 결합되어 있는 다른 채널을 발생한다. 블라인드 소스 분리 알고리즘은 통계적 기술을 통해 방향성 분리를 증강해준다. 그 다음, 잡음 신호 및 음성 신호는 잡음 성분을 효율적으로 줄이거나 제거하기 위하여 듀얼 입력 잡음 스펙트럼 감산 억제기(DINS)에서 프로세스 특성들(process characteristics)을 설정하는데 사용된다. 이러한 방식으로, 잡음은 결합 신호로부터 효과적으로 제거되어 양호한 품질의 음성 신호가 생성된다.

대표도 - 도4



(72) 발명자

클라크, 조엘, 에이.

미국 60517 일리노이주 우드릿지 햄프스테드 9017

프랜코즈, 할리, 엘.

영국 지유2 8디제이 길드포드 사우쓰 요크셔 사우
쓰웨이 163

이자벨레, 스코트, 케이.

미국 60087 일리노이주 위케간 빅토리아 레인 3255

피어스, 데이비드, 제이.

영국 알지24 8더블유이 베이징스토크 햄프셔 파요
츠 코프스 7

렉스, 제임스, 에이.

영국 에스오51 8에이치제이 램지 햄프셔 미들브릿
지 스트리트 49

특허청구의 범위

청구항 1

잡음 음향 환경으로부터 음성 신호를 분리하는 잡음 저감 시스템으로서,
 각기 하나 이상의 음향 신호를 수신하는 복수의 입력 채널들;
 상기 하나 이상의 음향 신호를 음성 빔 및 잡음 빔으로 분리하는 적어도 하나의 소스 필터 - 상기 소스 필터는 적어도 하나의 하이퍼카디오이드 방향성 필터를 포함함 -;
 상기 음성 빔 및 잡음 빔을 정제(refine)하도록 동작가능한 적어도 하나의 블라인드 소스 분리(blind source separation: BSS) 필터; 및
 상기 음성 빔으로부터 잡음을 제거하는 적어도 하나의 듀얼 입력 스펙트럼 감산 잡음 억제기(dual input spectral subtraction noise suppressor: DINS)
 를 포함하는 잡음 저감 시스템.

청구항 2

제1항에 있어서, 상기 소스 필터는 위상 지연 필터링을 이용하여 음성 빔 및 잡음 빔을 형성하고, 음성 및 잡음 빔은 상기 소스 필터에 의해 주파수 응답 보정되는 잡음 저감 시스템.

청구항 3

제1항에 있어서, 상기 블라인드 소스 분리(BSS) 필터로부터의 상기 정제된 음성 빔 및 잡음 빔은 듀얼 입력 스펙트럼 감산 잡음 억제기(DINS)로 공급되는 잡음 저감 시스템.

청구항 4

제1항에 있어서, 상기 블라인드 소스 분리(BSS) 필터로부터의 상기 정제된 잡음 빔, 및 상기 소스 필터로부터의 상기 음성 빔은 상기 듀얼 입력 스펙트럼 감산 잡음 억제기(DINS)로 공급되는 잡음 저감 시스템.

청구항 5

제1항에 있어서, 상기 시스템은, 두 개의 블라인드 소스 분리(BSS) 필터들의 캐스캐이딩을 더 포함하며,
 상기 캐스캐이딩된 두 개의 블라인드 소스 분리 필터들에는 상기 소스 필터로부터의 음성 빔 및 잡음 빔이 입력되며,
 상기 캐스캐이딩된 두 개의 블라인드 소스 분리 필터들의 출력은 상기 듀얼 입력 스펙트럼 감산 잡음 억제기(DINS)로 공급되는 잡음 저감 시스템.

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

잡음 저감 방법으로서,

복수의 입력 채널들로부터 하나 이상의 음향 신호를 수신하는 단계;

소스 필터로 상기 하나 이상의 음향 신호를 음성 빔 및 잡음 빔으로 분리하는 단계 - 상기 소스 필터는 적어도 하나의 하이퍼카디오이드 방향성 필터를 포함함 -;

적어도 하나의 블라인드 소스 분리(blind source separation: BSS) 필터를 이용하여 상기 음성 빔 및 잡음 빔을 정제하는 단계 - 상기 블라인드 소스 분리 필터는 상기 음성 및 잡음 빔을 정제하도록 동작함 -; 및

적어도 하나의 듀얼 입력 스펙트럼 감산 잡음 억제기(dual input spectral subtraction noise suppressor: DINS)를 통해 상기 음성 빔으로부터 잡음을 제거하는 단계

를 포함하는 잡음 저감 방법.

청구항 13

제12항에 있어서, 소스 필터에서 상기 분리하는 단계는 위상 지연 필터링을 통해 수행되고, 음성 빔 및 잡음 빔은 주파수 응답 보정되는 잡음 저감 방법.

청구항 14

제12항에 있어서, 상기 블라인드 소스 분리(BSS) 필터로부터의 상기 정제된 음성 빔 및 잡음 빔은 상기 듀얼 입력 스펙트럼 감산 잡음 억제기(DINS)로 공급되는 잡음 저감 방법.

청구항 15

제12항에 있어서, 상기 블라인드 소스 분리(BSS) 필터로부터의 상기 정제된 잡음 빔, 및 상기 소스 필터로부터의 상기 음성 빔은 상기 듀얼 입력 스펙트럼 감산 잡음 억제기(DINS)로 공급되는 잡음 저감 방법.

청구항 16

제12항에 있어서,

두 개의 블라인드 소스 분리(BSS) 필터들을 캐스케이딩하는 단계를 더 포함하며,

상기 캐스케이딩된 두 개의 블라인드 소스 분리 필터들에는 상기 소스 필터로부터의 상기 음성 빔 및 잡음 빔이 입력되며,

상기 캐스케이딩된 두 개의 블라인드 소스 분리 필터들의 출력은 상기 듀얼 입력 스펙트럼 감산 잡음 억제기(DINS)로 공급되는 잡음 저감 방법.

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

명세서

기술분야

[0001] 본 발명은 다중 음향 신호들을 처리하기 위한 시스템 및 방법에 관한 것으로, 특히, 음향 신호들을 필터링을 통해 분리하는 것에 관한 것이다.

배경기술

[0002] 잡음 환경에서 정보 신호를 검출하고 그에 반응하는 것은 대부분 어렵다. 흔히 사용자가 잡음 환경에서 이야기 하는 통화에서, 사용자의 음성 신호를 배경 잡음으로부터 분리하는 것이 바람직하다. 배경 잡음은 일반 환경에서 발생한 수많은 잡음 신호와, 배후의 다른 사람들의 대화 내용에 의해 발생한 신호뿐만 아니라 반향음(reflections), 및 각 신호들로부터 발생한 잔향음(reverberation)을 포함할 수 있다.

[0003] 잡음 환경에서, 업링크 통신은 심각한 문제가 될 수 있다. 이러한 잡음 문제에 대한 대부분의 해결책은 정적 잡음(stationary noise)과 같은 소정 형태의 잡음을 연구하거나, 또는 잡음 신호처럼 사용자를 성가시게 할 수 있는 상당한 오디오 아티팩트(significant audio artifacts)를 만드는 것뿐이다. 기존의 모든 해결책은 소스 및 잡음 위치, 그리고 억제하려는 잡음 형태에 관한 단점이 있다.

발명의 내용

해결하려는 과제

[0004] 본 발명의 목적은 모든 잡음 소스들을, 이들의 시간적 특성, 위치, 또는 이동과 무관하게 억제하는 수단을 제공하는 것이다.

과제의 해결 수단

[0005] 음성 신호를 잡음 음향 환경에서 분리하기 위한 시스템, 방법 및 장치가 제공된다. 분리 프로세스는 방향성 필터링(빔형성(beamforming)), 블라인드 소스 분리(blind source separation), 및 듀얼 입력 스펙트럼 감산 잡음 억제(dual input spectral subtraction noise suppression)일 수 있는 소스 필터링을 포함할 수 있다. 입력 채널들은 두 개의 전방향성 마이크로폰들을 포함할 수 있고, 이들 마이크로폰들의 출력은 위상 지연 필터링을 이용하여 처리되어, 음성 및 잡음 빔형상들(beamforms)을 형성한다. 또한, 빔형상들은 주파수 보정될(frequency corrected) 수 있다. 빔형성 동작은 실질적으로 잡음인 하나의 채널과, 잡음과 음성의 결합인 다른 채널을 발생한다. 블라인드 소스 분리 알고리즘은 통계적 기술을 통해 방향성 분리를 증강해준다. 그래서 잡음 신호 및 음성 신호는 잡음 성분을 효율적으로 줄이거나 제거하기 위하여 듀얼 입력 스펙트럼 감산 잡음 억제기(dual input spectral subtraction noise suppressor: DINS)에서 프로세스 특성들을 설정하는데 사용된다. 이러한 방식으로, 잡음은 결합 신호로부터 효과적으로 제거되어 양호한 품질의 음성 신호가 생성된다.

도면의 간단한 설명

[0006] 본 발명의 기술한 장점 및 특징 그리고 다른 장점 및 특징이 획득될 수 있는 방식을 기술하기 위하여, 앞에서 간략히 기술한 본 발명은 첨부 도면에 예시된 본 발명의 특정 실시예를 참조하여 더 상세히 설명될 것이다. 이들 도면들이 본 발명의 전형적인 실시예들만을 묘사하고 있으며 그러므로 본 발명의 범주가 제한되는 것으로 간주하지 않음은 당연하며, 본 발명은 첨부 도면을 사용하여 추가적인 특이성과 세부 사항으로 기술되고 설명될 것이다.

도 1은 두 개의 전방향성 마이크로폰들(omnidirectional microphones)로부터 잡음 및 음성 빔형상을 형성하는 전방 하이퍼카디오이드 방향성 필터(front hypercardioid directional filter)를 이용하는 빔형성기의 사시도이다.

도 2는 두 개의 전방향성 마이크로폰들로부터 잡음 및 음성 빔형상을 형성하는 전방 하이퍼카디오이드 방향성

필터 및 후방 카디오이드 방향성 필터를 이용하는 빔형성기의 사시도이다.

도 3은 본 발명의 실시가능한 실시예에 따른 강인한 듀얼 입력 스펙트럼 감산 잡음 억제기(robust dual input spectral subtraction noise suppressor: RDINS)의 블록도이다.

도 4는 본 발명의 실시가능한 실시예에 따른 블라인드 소스 분리(blind source separation: BSS) 필터 및 듀얼 입력 스펙트럼 감산 잡음 억제기(dual input spectral subtraction noise suppressor: DINS)의 블록도이다.

도 5는 본 발명의 실시가능한 실시예에 따른 BSS의 음성 출력을 바이패스하는 블라인드 소스 분리(BSS) 필터 및 듀얼 입력 스펙트럼 감산 잡음 억제기(DINS)의 블록도이다.

도 6은 본 발명의 실시가능한 실시예에 따른 정적 잡음 추정 방법의 플로우차트이다.

도 7은 본 발명의 실시가능한 실시예에 따른 연속 잡음 추정 방법의 플로우차트이다.

도 8은 본 발명의 실시가능한 실시예에 따른 강인한 듀얼 입력 스펙트럼 감산 잡음 억제기(RDINS)의 방법의 플로우차트이다.

발명을 실시하기 위한 구체적인 내용

[0007] 본 발명의 부가적인 특징 및 장점은 다음의 설명에서 기술될 것이며, 부분적으로 그 설명으로부터 자명해질 것이며, 또는 본 발명의 실시를 통해 알게 될 수 있다. 본 발명의 특징 및 장점은 첨부된 청구범위에서 특별히 지적인 장치들과 조합물들에 의해 실현되고 획득될 수 있다. 본 발명의 이러한 특징 및 다른 특징은 다음의 설명과 첨부된 청구범위로부터 더욱 자명해질 것이고, 또는 본 명세서에서 설명한 바와 같은 본 발명의 실시를 통해 알 수가 있다.

[0008] 본 발명의 여러 실시예들은 아래에서 상세히 설명된다. 특정한 구현예들이 기술되지만, 이것은 단지 예시 목적을 위한 것임을 이해하여야 한다. 관련 기술의 숙련자는 다른 컴포넌트 및 구성이 본 발명의 정신과 범주를 이탈함이 없이도 이용될 수 있음을 인식할 것이다.

[0009] 본 발명은 각종 실시예, 이를 테면, 본 발명의 기본 개념과 관련한 방법 및 장치 그리고 다른 실시예들을 포함한다.

[0010] 도 1은 본 발명의 실시가능한 실시예에 따라서 두 개의 전방향성 마이크로폰으로부터 잡음 및 음성 빔형상(speech beamforms)을 형성하기 위한 빔형성기(beamformer)(100)의 예시적인 도면을 예시한다. 두 개의 마이크로폰들(110)은 서로 이격되어 있다. 각 마이크로폰은 입력 신호를 직접 또는 간접적으로 수신하고, 신호를 출력할 수 있다. 두 개의 마이크로폰들(110)은 전방향성이어서, 이들 마이크로폰에 대해 모든 방향으로부터 사운드를 거의 동일하게 수신한다. 마이크로폰들(110)은 음성 및 잡음의 혼합된 사운드를 나타내는 음향 신호 또는 에너지를 수신할 수 있으며, 이 입력들은 음성이 우세한 제1 신호(140)와 음성 및 잡음을 갖는 제2 신호(150)로 변환될 수 있다. 도시되지 않았지만, 마이크로폰들은 내부 또는 외부의 아날로그-디지털 변환기를 포함할 수 있다. 마이크로폰들(110)로부터의 신호들은 하나 이상의 변환 함수를 이용하여, 시간 도메인과 주파수 도메인 사이에서 스케일링 또는 변환될 수 있다. 빔형성은 마이크로폰들(110)에 의해 수신된 상이한 신호들의 상이한 전파 시간들을 보상할 수 있다. 도 1에 도시된 바와 같이, 마이크로폰들의 출력들은, 마이크로폰들(110)로부터의 신호들을 주파수 응답 보정하도록, 소스 필터링 또는 방향성 필터링(120)을 이용하여 처리된다. 빔형성기(100)는 전방 하이퍼카디오이드 방향성 필터(front hypercardioid directional filter)(130)를 이용하여 마이크로폰들(110)로부터의 신호들을 추가로 필터링한다. 일 실시예에서, 방향성 필터는 주파수에 따라 가변하여 모든 주파수 범위에 걸쳐 이상적인 빔형상을 형성하는 진폭 및 위상 지연 값들을 갖는다. 이들 값들은 자유 공간에 놓인 마이크로폰들이 필요로하는 이상적인 값들과 상이할 수 있다. 상이한 것은 마이크로폰들이 배치된 물리적 하우징의 기하구성을 고려한 것일 것이다. 이 방법에서, 마이크로폰들(110)의 공간적 차이로 인한 신호들 간의 시간 차이는 신호를 강화시키는데 사용된다. 보다 상세히 말해서, 마이크로폰들(110) 중 하나는 음성 소스(화자)의 근방에 더 가까이 있을 것이며, 반면 다른 마이크로폰은 비교적 감쇄된 신호를 발생할 수 있다. 도 2는 본 발명의 실시가능한 실시예에 따른, 두 개의 전방향성 마이크로폰들로부터 잡음(250) 및 음성(240) 빔형상을 형성하기 위한 빔형성기(200)의 예시적인 도면을 예시한다. 빔형성기(200)는 마이크로폰들(110)로부터 신호를 추가로 필터링하기 위하여 후방 카디오이드(cardioid) 방향성 필터(260)를 더 포함한다.

[0011] 전방향성 마이크로폰들(110)은 사운드 신호를 마이크로폰 주변의 어떤 방향에서도 대략 동일하게 수신한다. 감지 패턴(도시되지 않음)은 마이크로폰 주변의 모든 방향으로부터 대략 동일한 진폭의 수신 신호 전력을 보여준

다. 따라서, 마이크로폰으로부터의 전기적 출력은 사운드가 어느 방향으로부터 마이크로폰에 도달하는가에 무관하게 동일하다.

[0012] 전방 하이퍼카디오이드(230) 감지 패턴은 카디오이드 패턴과 비교하여 협소한 각도의 기본 감도(primary sensitivity)를 제공한다. 더욱이, 하이퍼카디오이드 패턴은 전방으로부터 대략 ± 140 각도에 위치한, 두 지점의 최저 감도의 지점을 갖는다. 그와 같이, 하이퍼카디오이드 패턴은 마이크로폰의 측면 그리고 마이크로폰의 후면 양측 모두에서 수신한 사운드를 억제한다. 그러므로, 하이퍼카디오이드 패턴들은 악기와 보컬리스트를 실내 환경으로부터 그리고 서로로부터 격리하는데 가장 적합하다.

[0013] 후 방면 카디오이드 또는 후방 카디오이드(260) 감지 패턴(도시되지 않음)은 방향성이고, 사운드 소스가 마이크로폰 쌍의 후방에 있을 때 완전한 감도를 제공한다. 마이크로폰 쌍의 측면에서 수신된 사운드는 출력의 대략 절반을 가지며, 마이크로폰 쌍의 전방에서 나오는 사운드는 실질적으로 감쇄된다. 이러한 후방 카디오이드 패턴은 무형의 가상 마이크로폰이 원하는 음성 소스(화자)를 겨냥하도록 만들어진다.

[0014] 모든 경우에서, 하나의 전방향성 마이크로폰을, 위상 지연 필터(그 출력은 다른 전방향성 마이크로폰 신호와 합쳐져서 널 위치들(null locations)을 설정함)와, 그리고 나서 결과 신호의 주파수 응답을 보정하는 보정 필터를 이용하여 필터링함으로써 빔이 형성된다. 적절한 주파수-중속 지연을 포함하는 분리 필터는 카디오이드(260) 및 하이퍼카디오이드(230) 응답을 생성하는데 사용된다. 대안으로, 빔은 전술한 프로세스를 이용하여 전 방면 및 후 방면 카디오이드 빔을 먼저 생성하고, 그 카디오이드 신호를 합하여 가상의 전방향성 신호를 생성하고, 신호들의 차를 취하여 양방향성 필터 또는 다이폴 필터를 생성함으로써 생성될 수 있다. 가상의 전방향성 및 다이폴 신호는 수학적 1을 이용하여 결합되어 하이퍼카디오이드 응답을 생성한다.

[0015] [수학식 1]

[0016] 하이퍼카디오이드(Hypercardioid) = 0.25*(omni +3*dipole)

[0017] 대안의 실시예는 고정된 지향성 단일 요소의 하이퍼카디오이드 및 카디오이드 마이크로폰 캡슐을 사용하는 것이다. 이것은 신호 처리에서 빔형성 단계를 필요없게 해 주겠지만, 장치에서 하나의 사용 모드로부터 다른 사용 모드로 빔형상의 변경이 더 어려우며, 장치에서 다른 처리 동안 진짜 전방향성 신호가 이용불가능할 것이라는 점에서, 시스템의 적응성을 제한할 것이다. 이 실시예에서, 소스 필터는 고역 통과 필터, 저역 통과 안티앨리어싱(antialiasing) 필터, 또는 대역 통과 필터와 같이 대역 잡음을 줄여주는 통과 대역을 갖는 단순 필터 또는 주파수 보정 필터 둘 중 하나일 수 있다.

[0018] 도 3은 본 발명의 실시가능한 실시예에 따라서 강인한 듀얼 입력 스펙트럼 감산 잡음 억제기(robust dual input spectral subtraction noise suppressor (RDINS))의 예시적인 도면을 예시한다. 음성 추정 신호(240) 및 잡음 추정 신호(250)는 RDINS(305)로의 입력으로서 공급되어 음성과 잡음의 스펙트럼 특성의 차이를 활용하여 음성 신호(140) 중 잡음 성분을 억제한다. RDINS(305)의 알고리즘은 방법(600 내지 800)을 참조하여 더 잘 설명된다.

[0019] 도 4는 블라인드 소스 분리(BSS) 필터 및 듀얼 입력 스펙트럼 감산 잡음 억제기(DINS)를 이용하여 음성(140) 및 잡음(150) 빔형상들을 처리하는 잡음 억제 시스템(400)의 예시적인 도면을 예시한다. 잡음 및 음성 빔형상은 주파수 응답 보정되었다. 블라인드 소스 분리(BSS) 필터(410)는 잡음 신호로부터 잔류 음성 신호를 제거한다. BSS 필터(410)는 정제된 잡음 신호(420) 만 또는 정제된 잡음 및 음성 신호(420, 430)를 생성할 수 있다. BSS 는 두 입력(음성과 잡음) 및 원하는 개수의 출력을 갖는 일단(single stage)의 BSS 필터일 수 있다. 두 단의 BSS 필터는 원하는 개수의 출력들과 함께 캐스케이드 연결한 또는 연결한 두 단의 BSS를 가질 것이다. 블라인드 소스 분리 필터는 통계적으로 서로 독립적이라 추정되는 혼합된 소스 신호들을 분리한다. 블라인드 소스 분리 필터(410)는 가중치들의 비혼합(un-mixing) 매트릭스에 혼합된 신호들을 곱셈하여 분리 신호를 생성함으로써 가중치들의 비혼합 매트릭스를 혼합된 신호들에 가중한다. 매트릭스에서 가중치들은 할당된 초기값들이며 정보 중복(information redundancy)을 최소화하기 위하여 조정된다. 이러한 조정은 출력 신호(420, 430)의 정보 중복이 최소한으로 줄어들 때까지 반복된다. 이러한 기술은 각 신호의 소스에 관한 정보를 필요로하지 않기 때문에, 이를 블라인드 소스 분리라고 지칭한다. BSS 필터(410)는 통계적으로 잡음으로부터 음성을 제거하여 저감된-음성 잡음 신호(reduced-speech noise signal)(420)를 생성한다. DINS 유닛(440)은 실질적으로 잡음이 없는 음성 신호(460)를 생성하기 위하여 저감된-음성 잡음 신호(420)를 이용하여 음성(430)으로부터 잡음을 제거한다. DINS 유닛(440) 및 BSS 필터(410)는 단일 유닛(450)으로 통합될 수 있거나, 또는 개별의 컴포넌트들로 분리될 수 있다.

- [0020] 마이크로폰(110)으로부터 처리된 신호들에 의해 제공된 음성 신호(140)는, 입력으로서 블라인드 소스 분리 필터(410)에 전달되며, 블라인드 소스 분리 필터에서, 처리된 음성 신호(430) 및 잡음 신호(420)는 DINS(440)로 출력되며, 여기서 처리된 음성 신호(430)는 BSS 필터(410)에서 수행된 블라인드 소스 분리 알고리즘의 동작에 의해 주변 사운드(잡음)로부터 분리된 사용자의 음성으로 완전히 또는 적어도 본질적으로 이루어진다. 그러한 BSS 신호 처리는 환경을 향하여 지향된 마이크로폰 및 화자를 향하여 지향된 마이크로폰에 의해 픽업된 혼합음들이 주변 사운드와 사용자의 음성의 상이한 혼합음으로 이루어진다는 사실을 활용하며, 이들은 이 두 신호들의 기여자들 또는 소스들의 진폭 비율에 관해 그리고 혼합음 중 이들 두 신호의 기여자들의 위상 차에 관해서는 서로 상이하다.
- [0021] DINS 유닛(440)은 처리된 음성 신호(430) 및 잡음 신호(420)를 더 강화시키며, 잡음 신호(420)는 DINS 유닛(440)의 잡음 추정(noise estimate)으로서 사용된다. 결과적인 잡음 추정(420)에는 음성 신호가 대폭 줄어 있어야 하는데, 이것은 남아있는 원하는 음성(460) 신호가 음성 강화 절차에 유리하지 않을 것이며 그래서 출력의 품질을 더 낮게 할 것이기 때문이다.
- [0022] 도 5는 블라인드 소스 분리(BSS) 필터 및 듀얼 입력 스펙트럼 감산 잡음 억제기(DINS)를 이용하여 음성(140) 및 잡음(150) 빔형상을 처리하는 잡음 억제 시스템(500)의 예시적인 도면을 예시한다. DINS 유닛(440)의 잡음 추정은 여전히 BSS 필터(410)로부터의 처리된 잡음 신호이다. 그러나, 음성 신호(430)는 BSS 필터(410)에 의해 처리되지 않는다.
- [0023] 도 6 내지 도 8은 본 발명의 실시가능한 실시예에 따른 강인한 듀얼 입력 스펙트럼 감산 잡음 억제기(RDINS)에 필요한 정적 잡음 추정(static noise estimates)을 결정하기 위한 기본 단계의 일부를 예시하는 예시적인 플로우차트이다.
- [0024] BSS가 사용되지 않을 때, 방향성 필터링의 출력(240, 250)은 듀얼 입력 채널 잡음 억제기(DINS)로 직접 인가될 수 있고, 유감스럽게도 후 방면 카디오이드 패턴(260)은 원하는 화자에게 부분적인 널(partial null)을 배치할 뿐이며, 이로써 잡음 추정에서 원하는 화자의 3dB 내지 6dB 만 억제되는 결과를 가져온다. DINS 유닛(440) 자체에서, 이러한 음성 누설 양은 음성이 처리된 이후 음성에 허용할 수 없는 왜곡을 일으킨다. RDINS는 잡음 추정 신호(250)에서 이러한 음성 누설에 더욱 강인해지도록 설계된 DINS의 한가지 버전이다. 이러한 견고성(robustness)은 두 가지 별개 잡음 추정을 이용함으로써 성취되는데, 즉, 하나는 방향성 필터링으로부터의 연속 잡음 추정(continuous noise estimate)이며 다른 것은 단일 채널 잡음 억제기에서도 사용될 수 있는 정적 잡음 추정(static noise estimate)이다.
- [0025] 방법(600)은 음성 빔(240)을 이용한다. 연속 음성 추정(continuous speech estimate)은 음성 빔(240)으로부터 구하며, 이 추정은 음성 구간과 무음성(speech free) 구간 동안 구한다. 음성 추정의 에너지 레벨이 단계(610)에서 계산된다. 단계(620)에서, 음성 액티비티 검출기(voice activity detector)가 사용되어 매 프레임마다 음성 추정 시에 무음성 구간들을 찾는다. 단계(630)에서, 음성 추정의 무음성 구간에서 평탄한 정적 잡음 추정이 형성된다. 이러한 정적 잡음 추정은 원하는 입력 음성의 지속기간 동안 금지(frozen)되므로 어떤 음성도 포함하지 않을 것인데, 그러나, 이것은 비정적 잡음 동안 잡음 추정이 변동을 포착하지 못함을 의미한다. 단계(640)에서, 정적 잡음 추정의 에너지가 계산된다. 단계(650)에서, 연속 음성 신호(615)의 에너지 및 정적 잡음 추정의 에너지로부터 정적 신호대 잡음비가 계산된다. 단계(620) 내지 단계(650)는 각 서브밴드마다 반복된다.
- [0026] 방법(700)은 연속 잡음 추정(250)을 이용한다. 단계(710)에서, 잡음 빔(250)으로부터 연속 잡음 추정을 구하며, 이 추정은 음성 구간 및 무음성 구간 동안 구한다. 이러한 연속 잡음 추정(250)은 불완전한 널(imperfect null)로 인하여 원하는 화자로부터 음성 누설이 있을 것이다. 단계(720)에서, 서브밴드 동안 잡음 추정의 에너지가 계산된다. 단계(730)에서, 서브밴드의 연속 신호대 잡음비가 계산된다.
- [0027] 방법(800)은, 연속 잡음 추정의 계산된 신호대 잡음비 및 정적 잡음 추정의 계산된 신호대 잡음비를 이용하여, 잡음을 억제할지를 결정한다. 단계(810)에서, 만일 연속 SNR이 제1 임계값보다 크면, 본 방법은 억제값을 연속 SNR과 함께 설정하는 단계(820)로 진행한다. 단계(810)에서, 연속 SNR이 제1 임계값보다 크지 않으면, 본 방법은 단계(830)로 진행한다. 단계(830)에서, 만일 연속 SNR이 제2 임계값보다 작으면, 본 방법은 억제값을 정적 SNR로 설정하는 단계(840)로 진행한다. 만일 연속 SNR이 제2 임계값보다 작지 않으면, 본 방법은 가중 평균 잡음 억제기를 사용하는 단계(850)로 진행한다. 가중 평균은 정적 및 연속 SNR들의 평균이다. 낮은 SNR 서브-밴드(잡음에 비해 음성이 없는/약한)의 경우, 연속 잡음 추정을 사용하여 비정적인 잡음 동안 효과적인 억제량을 결정한다. 높은 SNR 서브-밴드(잡음에 비해 음성이 강한)의 경우, 연속 잡음 추정에서 누설(leakage)이 우세할 때, 정적 잡음 추정을 사용하여 음성을 과하게 억제하고 왜곡을 일으키는 음성 누설을 방지하는 억제량을 결정

한다. 중간 SNR 서브-밴드 동안, 두 가지 추정을 결합하여 전술한 두 경우들이 부드럽게 전환되게 한다. 단계 (860)에서, 채널 이득이 계산된다. 단계(870)에서, 채널 이득은 음성 추정에 적용된다. 이들 단계들은 각 서브밴드에 대해 반복된다. 그런 다음 채널 이득은 DINS에서와 동일한 방식대로 SNR이 높은 채널들을 통과시키면서 SNR이 낮은 채널들을 감쇄시키도록 적용된다. 이 구현예에서, 음성 파형은 윈도우드 역 FFT(windowed Inverse FFT)의 오버랩-애드(overlap add)에 의해 재구성된다.

[0028] 실제로, 쌍방향 통신 장치는 사용 모드에 따라서 전환되는 본 발명의 여러 실시예들을 포함할 수 있다. 예를 들어, 도 1에서 기술된 빔형성 동작은 근접-대화(close-talking) 또는 개인 모드에서 사용하기 위해 도 4에서 기술된 BSS 단 및 DINS와 결합될 수 있지만, 핸드프리 또는 스피커폰(speakerphone) 모드에서, 도 2의 빔형성기는 도 3의 RDINS와 결합될 수 있다. 이들 동작 모드들 간의 전환은 본 기술 분야에서 공지된 많은 구현예들 중 하나에 의해 이루어질 수 있다. 예를 들어, 제한하지는 않지만, 전환 방법은 접근 기반 판단 로직(logic decision based on proximity), 자기 또는 전기 스위치, 또는 본 명세서에서 기술되지 않은 어떤 대등한 방법을 통하여 가능할 수 있다.

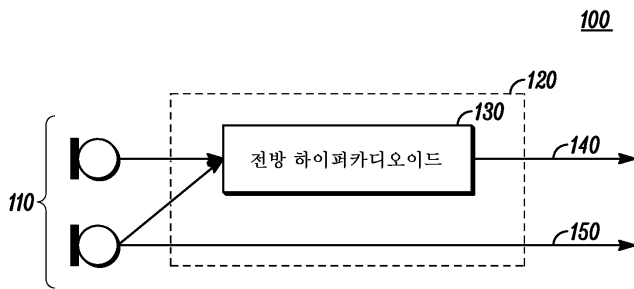
[0029] 본 발명의 범주 내에 속하는 실시예들은 저장된 컴퓨터 실행가능 명령어들 또는 데이터 구조들을 담거나 가지고 있는 컴퓨터 판독가능 매체를 또한 포함할 수 있다. 그러한 컴퓨터 판독가능한 매체는 범용 또는 전용 컴퓨터에 의해 액세스될 수 있는 모든 이용가능한 매체일 수 있다. 예를 들어, 제한하지 않지만, 그러한 컴퓨터 판독가능한 매체는 원하는 프로그램 코드 수단을 컴퓨터 실행가능한 명령어들 또는 데이터 구조들의 형태로 담거나 저장하는데 사용될 수 있는 RAM, ROM, EEPROM, CD-ROM 또는 다른 광 디스크 스토리지, 자기 디스크 스토리지 또는 다른 자기 스토리지 장치, 또는 다른 모든 매체를 포함할 수 있다. 정보가 네트워크 또는 다른 (유선, 무선, 또는 이들의 조합) 통신 연결을 통해 컴퓨터로 전달 또는 제공될 때, 컴퓨터는 당연히 그 연결을 컴퓨터 판독가능한 매체로 간주한다. 따라서, 그러한 어떤 연결이라도 당연히 컴퓨터 판독가능한 매체라고 불리운다. 전술한 결합은 또한 컴퓨터 판독가능한 매체의 범주 내에 속하여야 할 것이다.

[0030] 컴퓨터 실행가능한 명령어들은, 예를 들어, 범용 컴퓨터, 전용 컴퓨터, 또는 전용 처리 장치가 어떤 기능이나 기능들의 그룹을 수행하게 하는 명령어들 및 데이터를 포함한다. 컴퓨터 실행가능한 명령어들은 또한 독립형 또는 네트워크 환경에서 컴퓨터들에 의해 실행되는 프로그램 모듈들을 포함한다. 일반적으로, 프로그램 모듈들은 특정한 작업을 실행하거나 특정한 추상적 데이터 형태를 구현하는 루틴, 프로그램, 객체, 컴포넌트, 및 데이터 구조 등을 포함한다. 컴퓨터 실행가능 명령어들, 연관된 데이터 구조들, 및 프로그램 모듈들은 본 명세서에서 개시된 방법들의 단계들을 실행하기 위한 프로그램 코드 수단의 예를 나타낸다. 그러한 실행가능한 명령어들 또는 연관된 데이터 구조들의 특정한 순서는 그러한 단계들에서 설명된 기능들을 구현하기 위한 해당 행위의 예를 나타낸다.

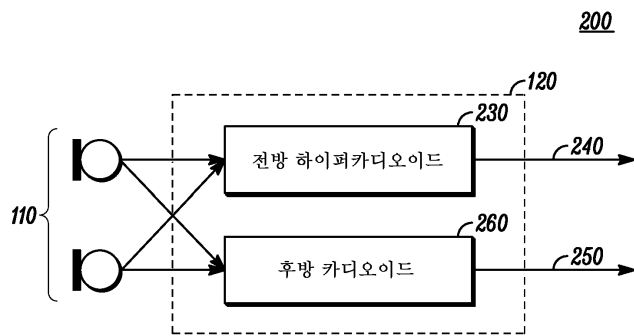
[0031] 비록 전술한 설명이 구체적인 세부 내용을 포함하고 있을지라도, 이 세부 내용들은 어떤 방식으로도 청구범위를 제한하는 것으로 해석되지 않아야 한다. 본 발명의 기술된 실시예들의 다른 구성들은 본 발명의 범주의 일부이다. 예를 들어, 본 발명의 원리는 각 개별 사용자에게 적용될 수 있으며, 각 사용자는 그러한 시스템을 개별적으로 효율적으로 사용할 수 있다. 이로써 각 사용자는 다수의 가능한 응용예들 중 어느 것이 본 명세서에서 기술된 기능을 필요로 하지 않을지라도 본 발명의 이익을 활용하는 것이 가능하다. 다시 말해서, 도 1 내지 도 8에서 방법 및 장치들의 많은 사례들이 있을 수 있으며 각기 여러 가능한 방식으로 콘텐츠를 처리할 수 있다. 최종 사용자들이 모두 반드시 한가지 시스템을 사용할 필요는 없다. 따라서, 첨부된 청구범위 및 그의 법적인 등가물은 주어진 어떤 특정한 예들이라기보다 본 발명을 규정할 뿐이어야 한다.

도면

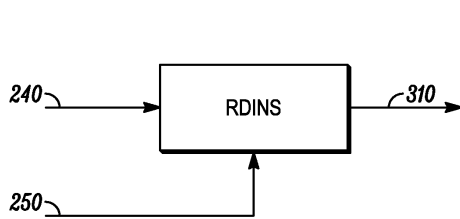
도면1



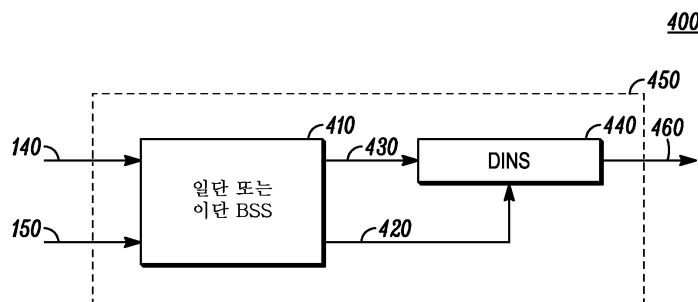
도면2



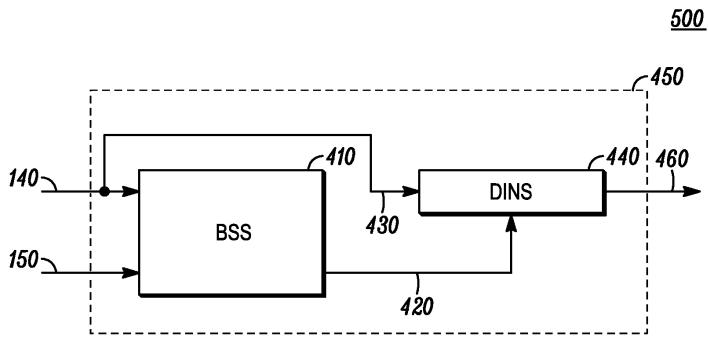
도면3



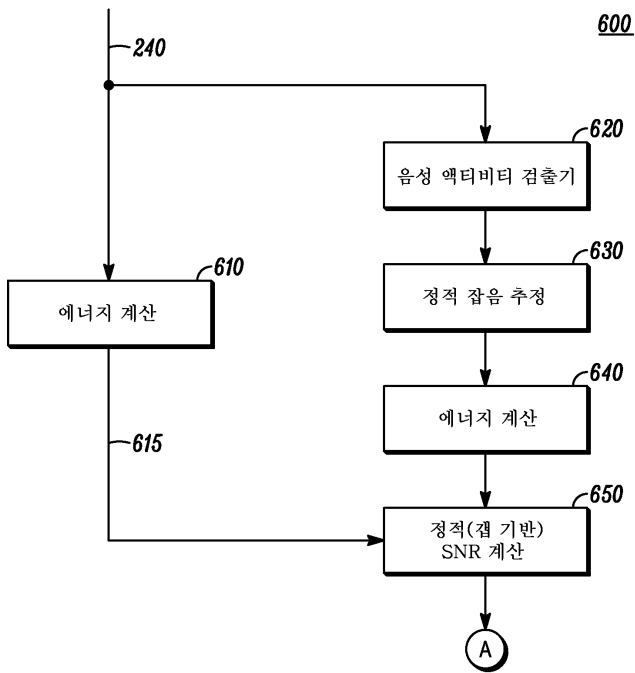
도면4



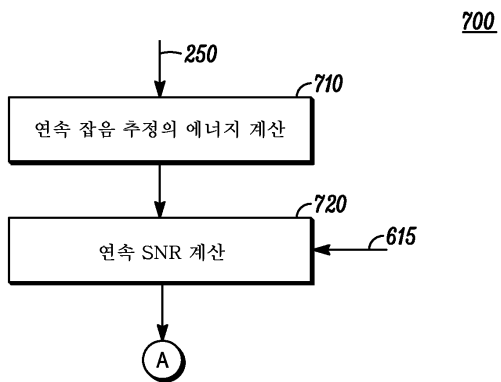
도면5



도면6



도면7



도면8

