



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2013년12월10일  
 (11) 등록번호 10-1340215  
 (24) 등록일자 2013년12월04일

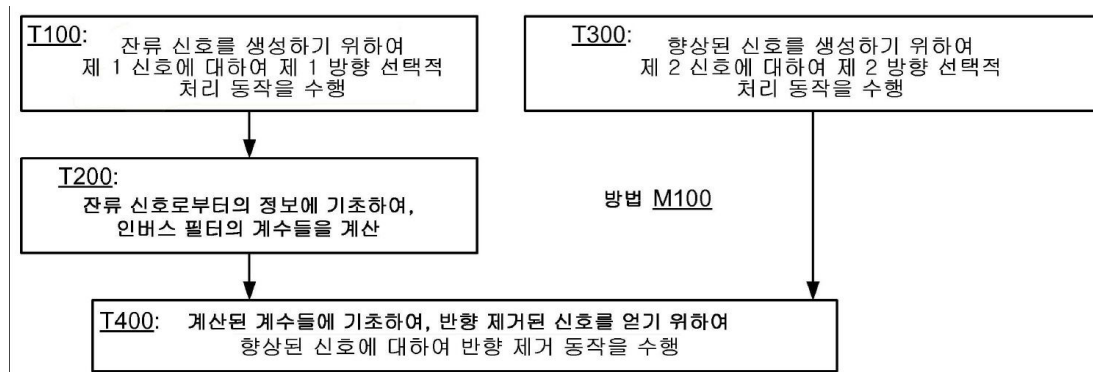
- |  |   |
|--|---|
| (51) 국제특허분류(Int. Cl.)<br>G10L 21/02 (2006.01)<br>(21) 출원번호 10-2012-7009000<br>(22) 출원일자(국제) 2012년09월07일<br>심사청구일자 2012년04월06일<br>(85) 번역문제출일자 2012년04월06일<br>(65) 공개번호 10-2012-0054087<br>(43) 공개일자 2012년05월29일<br>(86) 국제출원번호 PCT/US2010/048026<br>(87) 국제공개번호 WO 2011/029103<br>국제공개일자 2011년03월10일<br>(30) 우선권주장<br>12/876,163 2010년09월05일 미국(US)<br>61/240,301 2009년09월07일 미국(US)<br>(56) 선행기술조사문헌<br>US20040170284 A1<br>US20080059157 A1<br>전체 청구항 수 : 총 31 항 | (73) 특허권자<br><b>퀄컴 인코포레이티드</b><br>미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775<br>(72) 발명자<br><b>비셰르 에릭</b><br>미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775<br>(74) 대리인<br><b>특허법인코리아나</b> |
|--|---|
- 심사관 : 권영학

(54) 발명의 명칭 **멀티채널 신호의 반향 제거를 위한 시스템, 방법, 장치 및 컴퓨터 판독가능 매체**

**(57) 요약**

방향 선택적 처리 동작 (예를 들어, 빔형성) 의 사용을 비상관화 동작 (예를 들어, 블라인드 소스 분리 동작) 을 이용하여 얻어진 분리된 반향 추정에 트레이닝된 인버스 필터와 조합하는, 멀티마이크로폰 신호의 반향 제거를 위한 시스템들, 방법들, 장치들, 및 컴퓨터 판독가능 매체.

**대표도**



**특허청구의 범위**

**청구항 1**

방향성 컴포넌트를 포함하는 멀티채널 신호를 처리하는 방법으로서,  
 잔류 (residual) 신호를 생성하기 위하여 제 1 신호에 대하여 제 1 방향 선택적 처리 동작을 수행하는 단계;  
 향상된 신호를 생성하기 위하여 제 2 신호에 대하여 제 2 방향 선택적 처리 동작을 수행하는 단계;  
 상기 생성된 잔류 신호로부터의 정보에 기초하여, 룸-응답 (room-response) 전달 함수의 인버스 (inverse) 의 복수의 필터 계수들을 계산하는 단계; 및  
 반향 제거된 (dereverberated) 신호를 생성하기 위하여 상기 향상된 신호에 대하여 반향 제거 동작을 수행하는 단계를 포함하고,  
 상기 반향 제거 동작은 상기 계산된 복수의 필터 계수들에 기초하고,  
 상기 제 1 신호는 상기 멀티채널 신호의 적어도 두 채널들을 포함하고,  
 상기 제 2 신호는 상기 멀티채널 신호의 적어도 두 채널들을 포함하며,  
 상기 제 1 신호에 대하여 제 1 방향 선택적 처리 동작을 수행하는 단계는 상기 제 1 신호의 총 에너지에 관하여 상기 제 1 신호 내의 상기 방향성 컴포넌트의 에너지를 감소시키는 단계를 포함하고,  
 상기 제 2 신호에 대하여 제 2 방향 선택적 처리 동작을 수행하는 단계는 상기 제 2 신호의 총 에너지에 관하여 상기 제 2 신호 내의 상기 방향성 컴포넌트의 에너지를 증가시키는 단계를 포함하는, 멀티채널 신호를 처리하는 방법.

**청구항 2**

제 1 항에 있어서,  
 상기 제 1 방향 선택적 처리 동작은 블라인드 소스 분리 (blind source separation) 동작인, 멀티채널 신호를 처리하는 방법.

**청구항 3**

제 1 항에 있어서,  
 상기 제 1 방향 선택적 처리 동작은 널 (null) 빔형성 동작인, 멀티채널 신호를 처리하는 방법.

**청구항 4**

제 1 항에 있어서,  
 상기 제 1 방향 선택적 처리 동작은:  
 상기 제 1 신호의 복수의 상이한 주파수 컴포넌트들 각각에 대하여, 상기 제 1 신호의 제 1 채널 내의 상기 주파수 컴포넌트의 위상과 상기 제 1 신호의 제 2 채널 내의 상기 주파수 컴포넌트의 위상 사이의 차를 계산하는 단계, 및  
 상기 제 1 신호 내의 상기 계산된 위상 차에 기초하여, 상기 제 1 신호의 상기 복수의 상이한 주파수 컴포넌트들 가운데 적어도 하나의 레벨을 상기 제 1 신호의 상기 복수의 상이한 주파수 컴포넌트들 가운데 다른 것의 레벨에 관하여 감쇠시키는 단계를 포함하고,  
 상기 감쇠시키는 단계는 상기 제 1 신호 내의 상기 방향성 컴포넌트로부터의 에너지를 마스킹 (masking) 하기 위해 상보적 마스크 (mask) 를 이용하는 단계를 포함하는, 멀티채널 신호를 처리하는 방법.

**청구항 5**

제 1 항에 있어서,

상기 제 1 방향 선택적 처리 동작은:

상기 제 1 신호의 복수의 상이한 주파수 컴포넌트들 각각에 대하여, 상기 제 1 신호의 제 1 채널 내의 상기 주파수 컴포넌트의 위상과 상기 제 1 신호의 제 2 채널 내의 상기 주파수 컴포넌트의 위상 사이의 차를 계산하는 단계; 및

특정 판정 기준을 만족시키는 상기 계산된 위상 차의 수에 기초하여, 상기 방향성 컴포넌트가 상기 제 1 신호 내에서 활성인지를 표시하는 단계를 포함하는, 멀티채널 신호를 처리하는 방법.

**청구항 6**

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

상기 반향 제거 동작을 수행하는 단계는,

상기 향상된 신호에 기초하는 신호로, 시간 도메인에서, 상기 계산된 복수의 필터 계수들을 컨볼빙(convolving) 하는 단계를 포함하는, 멀티채널 신호를 처리하는 방법.

**청구항 7**

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

상기 제 2 방향 선택적 처리 동작은 블라인드 소스 분리 동작인, 멀티채널 신호를 처리하는 방법.

**청구항 8**

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

상기 제 2 방향 선택적 처리 동작은 빔형성 동작인, 멀티채널 신호를 처리하는 방법.

**청구항 9**

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

상기 제 2 방향 선택적 처리 동작은:

상기 제 2 신호의 복수의 상이한 주파수 컴포넌트들 각각에 대하여, 상기 제 2 신호의 제 1 채널 내의 상기 주파수 컴포넌트의 위상과 상기 제 2 신호의 제 2 채널 내의 상기 주파수 컴포넌트의 위상 사이의 차를 계산하는 단계, 및

특정 판정 기준을 만족시키는 상기 계산된 위상 차의 수에 기초하여, 상기 방향성 컴포넌트가 상기 제 2 신호 내에서 활성인지를 표시하는 단계를 포함하는, 멀티채널 신호를 처리하는 방법.

**청구항 10**

제 1 항에 있어서,

상기 방법은, 상기 멀티채널 신호에 대해 블라인드 소스 분리 동작을 수행하는 단계를 포함하고,

상기 블라인드 소스 분리 동작은 상기 제 1 방향 선택적 처리 동작 및 상기 제 2 방향 선택적 처리 동작을 포함하고,

상기 제 1 신호는 상기 멀티채널 신호이고 상기 제 2 신호는 상기 멀티채널 신호인, 멀티채널 신호를 처리하는 방법.

**청구항 11**

제 1 항 내지 제 5 항 및 제 10 항 중 어느 한 항에 있어서,

상기 복수의 필터 계수들을 계산하는 단계는 상기 생성된 잔류 신호에 자기회귀 모델 (autoregressive model) 을 맞추는 단계를 포함하는, 멀티채널 신호를 처리하는 방법.

**청구항 12**

제 1 항 내지 제 5 항 및 제 10 항 중 어느 한 항에 있어서,

상기 복수의 필터 계수들을 계산하는 단계는 상기 복수의 필터 계수들을 상기 생성된 잔류 신호에 기초하는 자기회귀 모델의 파라미터들로서 계산하는 단계를 포함하는, 멀티채널 신호를 처리하는 방법.

**청구항 13**

제 1 항 내지 제 5 항 및 제 10 항 중 어느 한 항에 있어서,

2 kHz 와 3 kHz 사이의 상기 반향 제거 동작의 평균 이득 응답은 300 Hz 와 400 Hz 사이의 상기 반향 제거 동작의 평균 이득 응답보다 적어도 3 데시벨 더 큰, 멀티채널 신호를 처리하는 방법.

**청구항 14**

제 1 항 내지 제 5 항 및 제 10 항 중 어느 한 항에 있어서,

상기 제 1 방향 선택적 처리 동작 및 상기 제 2 방향 선택적 처리 동작 가운데 적어도 하나에 대하여, 2 kHz 에서 3 kHz 사이의 주파수 범위에서, 방향에 관한, 상기 동작의 최소 이득 응답과 상기 동작의 최대 이득 응답 사이의 차의 절대값은, 300 Hz 에서 400 Hz 사이의 주파수 범위에서, 방향에 관한, 상기 동작의 최소 이득 응답과 상기 동작의 최대 이득 응답 사이의 차의 절대값보다 더 큰, 멀티채널 신호를 처리하는 방법.

**청구항 15**

프로세서에 의하여 임혀졌을 때 상기 프로세서가 제 1 항 내지 제 5 항 및 제 10 항 중 어느 한 항에 기재된 방향성 컴포넌트를 포함하는 멀티채널 신호를 처리하는 방법을 수행하도록 하는 유형적 특징들을 포함하는, 컴퓨터 판독가능 저장 매체.

**청구항 16**

방향성 컴포넌트를 포함하는 멀티채널 신호를 처리하기 위한 장치로서,

잔류 신호를 생성하기 위하여 제 1 신호에 대하여 제 1 방향 선택적 처리 동작을 수행하기 위한 수단;

향상된 신호를 생성하기 위하여 제 2 신호에 대하여 제 2 방향 선택적 처리 동작을 수행하기 위한 수단;

상기 생성된 잔류 신호로부터의 정보에 기초하여, 룸-응답 전달 함수의 인버스의 복수의 필터 계수들을 계산하기 위한 수단; 및

반향 제거된 신호를 생성하기 위하여 상기 향상된 신호에 대하여 반향 제거 동작을 수행하기 위한 수단을 포함하고,

상기 반향 제거 동작은 상기 계산된 복수의 필터 계수들에 기초하고,

상기 제 1 신호는 상기 멀티채널 신호의 적어도 두 채널들을 포함하고,

상기 제 2 신호는 상기 멀티채널 신호의 적어도 두 채널들을 포함하고,

상기 제 1 신호에 대하여 제 1 방향 선택적 처리 동작을 수행하기 위한 수단은 상기 제 1 신호의 총 에너지에 관하여 상기 제 1 신호 내의 상기 방향성 컴포넌트의 에너지를 감소시키도록 구성되고,

상기 제 2 신호에 대하여 제 2 방향 선택적 처리 동작을 수행하기 위한 수단은 상기 제 2 신호의 총 에너지에 관하여 상기 제 2 신호 내의 상기 방향성 컴포넌트의 에너지를 증가시키도록 구성된, 멀티채널 신호를 처리하기 위한 장치.

**청구항 17**

제 16 항에 있어서,

상기 제 1 방향 선택적 처리 동작은 블라인드 소스 분리 동작인, 멀티채널 신호를 처리하기 위한 장치.

**청구항 18**

제 16 항에 있어서,

상기 제 1 방향 선택적 처리 동작은 널 빔형성 동작인, 멀티채널 신호를 처리하기 위한 장치.

**청구항 19**

제 16 항에 있어서,

상기 제 1 방향 선택적 처리 동작은:

상기 제 1 신호의 복수의 상이한 주파수 컴포넌트들 각각에 대하여, 상기 제 1 신호의 제 1 채널 내의 상기 주파수 컴포넌트의 위상과 상기 제 1 신호의 제 2 채널 내의 상기 주파수 컴포넌트의 위상 사이의 차를 계산하는 것, 및

상기 제 1 신호 내의 상기 계산된 위상 차에 기초하여, 상기 제 1 신호의 상기 복수의 상이한 주파수 컴포넌트들 가운데 적어도 하나의 레벨을 상기 제 1 신호의 상기 복수의 상이한 주파수 컴포넌트들 가운데 다른 것의 레벨에 관하여 감쇠시키는 것을 포함하고,

상기 감쇠시키는 것은 제 1 신호 내의 상기 방향성 컴포넌트로부터의 에너지를 마스킹하기 위해 상보적 마스크를 이용하는 것을 포함하는, 멀티채널 신호를 처리하기 위한 장치.

**청구항 20**

제 16 항에 있어서,

상기 제 1 방향 선택적 처리 동작은:

상기 제 1 신호의 복수의 상이한 주파수 컴포넌트 각각에 대하여, 상기 제 1 신호의 제 1 채널의 상기 주파수 컴포넌트의 위상과 상기 제 1 신호의 제 2 채널의 상기 주파수 컴포넌트의 위상 사이의 차를 계산하는 것을 포함하고,

특정 판정 기준을 만족시키는 상기 계산된 위상 차들의 수에 기초하여, 상기 방향성 컴포넌트가 상기 제 1 신호 내에서 활성화인지 여부를 표시하는 것을 포함하는, 멀티채널 신호를 처리하기 위한 장치.

**청구항 21**

제 16 항 내지 제 20 항 중 어느 한 항에 있어서,

상기 반향 제거 동작을 수행하는 것은, 상기 향상된 신호에 기초하는 신호로, 시간 도메인에서, 상기 계산된 복수의 필터 계수들을 컨볼빙 (convolving) 하는 것을 포함하는, 멀티채널 신호를 처리하기 위한 장치.

**청구항 22**

제 16 항 내지 제 20 항 중 어느 한 항에 있어서,

상기 반향 제거된 신호를 생성하기 위하여 상기 향상된 신호에 대하여 반향 제거 동작을 수행하기 위한 수단은, 상기 향상된 신호에 기초하는 신호로, 시간 도메인에서, 상기 계산된 복수의 필터 계수들을 컨볼빙하도록 구성된 필터인, 멀티채널 신호를 처리하기 위한 장치.

**청구항 23**

제 16 항 내지 제 20 항 중 어느 한 항에 있어서,

상기 제 2 방향 선택적 처리 동작은 블라인드 소스 분리 동작인, 멀티채널 신호를 처리하기 위한 장치.

**청구항 24**

제 16 항 내지 제 20 항 중 어느 한 항에 있어서,

상기 제 2 방향 선택적 처리 동작은 빔형성 동작인, 멀티채널 신호를 처리하기 위한 장치.

**청구항 25**

제 16 항 내지 제 20 항 중 어느 한 항에 있어서,

상기 제 2 방향 선택적 처리 동작은:

상기 제 2 신호의 상이한 복수의 주파수 컴포넌트들 각각에 대하여, 상기 제 2 신호의 제 1 채널 내의 상기 주파수 컴포넌트의 위상과 상기 제 2 신호의 제 2 채널 내의 상기 컴포넌트의 위상 사이의 위상 차를 계산하는 것, 및

특정 판정 기준을 만족시키는 상기 계산된 위상 차의 수에 기초하여, 상기 방향성 컴포넌트가 상기 제 2 신호 내에서 활성인지 여부를 표시하는 것을 포함하는, 멀티채널 신호를 처리하기 위한 장치.

**청구항 26**

제 16 항에 있어서,

상기 장치는, 상기 멀티채널 신호에 대하여 블라인드 소스 분리 동작을 수행하기 위한 수단을 포함하고,

상기 블라인드 소스 분리 동작을 수행하기 위한 수단은 상기 제 1 방향 선택적 처리 동작을 수행하기 위한 수단 및 상기 제 2 방향 선택적 처리 동작을 수행하기 위한 수단을 포함하며,

상기 제 1 신호는 상기 멀티채널 신호이고 상기 제 2 신호는 상기 멀티채널 신호인, 멀티채널 신호를 처리하기 위한 장치.

**청구항 27**

제 16 항 내지 제 20 항 및 제 26 항 중 어느 한 항에 있어서,

상기 복수의 필터 계수들을 계산하기 위한 수단은 상기 생성된 잔류 신호에 자기회귀 모델을 맞추도록 구성된, 멀티채널 신호를 처리하기 위한 장치.

**청구항 28**

제 16 항 내지 제 20 항 및 제 26 항 중 어느 한 항에 있어서,

상기 복수의 필터 계수들을 계산하기 위한 수단은 상기 복수의 필터 계수들을 상기 생성된 잔류 신호에 기초하는 자기회귀 모델의 파라미터들로서 계산하도록 구성된, 멀티채널 신호를 처리하기 위한 장치.

**청구항 29**

제 16 항 내지 제 20 항 및 제 26 항 중 어느 한 항에 있어서,

2 kHz 와 3 kHz 사이의 상기 반향 제거 동작의 평균 이득 응답은 300 Hz 와 400 Hz 사이의 상기 반향 제거 동작의 평균 이득 응답보다 적어도 3 데시벨 더 큰, 멀티채널 신호를 처리하기 위한 장치.

**청구항 30**

제 16 항 내지 제 20 항 및 제 26 항 중 어느 한 항에 있어서,

상기 반향 제거된 신호를 생성하기 위하여 상기 향상된 신호에 대하여 반향 제거 동작을 수행하기 위한 수단은 필터이고,

2 kHz 와 3 kHz 사이의 상기 필터의 평균 이득 응답은 300 Hz 와 400 Hz 사이의 상기 필터의 평균 이득 응답보다 적어도 3 데시벨 더 큰, 멀티채널 신호를 처리하기 위한 장치.

**청구항 31**

제 16 항 내지 제 20 항 및 제 26 항 중 어느 한 항에 있어서,

상기 제 1 방향 선택적 처리 동작 및 상기 제 2 방향 선택적 처리 동작 가운데 적어도 하나에 대하여,

2 kHz 에서 3 kHz 사이의 주파수 범위에서, 방향에 관한, 상기 동작의 최소 이득 응답과 상기 동작의 최대 이득 응답 사이의 차의 절대값은,

300 Hz 에서 400 Hz 사이의 주파수 범위에서, 방향에 관한, 상기 동작의 최소 이득 응답과 상기 동작의 최대 이득 응답 사이의 차의 절대값보다 더 큰, 멀티채널 신호를 처리하기 위한 장치.

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

**명세서**

**기술분야**

[0001] 35 U.S.C. § 119 하의 우선권 주장

[0002] 본 특허 출원은 양수인에게 양도되고, 발명의 명칭을 “SYSTEMS, METHODS, APPARATUS, AND COMPUTER-READABLE MEDIA FOR DEREVERBERATION OF MULTICHANNEL SIGNAL” 로 하는 2009년 9월 7일 출원된 미국 가출원 제 61/240,301 호를 우선권으로 주장한다.

[0003] 분야

[0004] 본 개시물은 신호 처리와 관련된 것이다.

**배경기술**

[0005] 반향 (reverberation) 은 특정 방향으로부터 발신하는 음향 신호 (예를 들어, 통신 디바이스의 사용자에게 의하여 발산된 음성 신호) 가 벽들 및/또는 다른 표면들로부터 반사될 때 생성된다. 마이크로폰으로 기록된 신호는 직접 경로 신호뿐만 아니라 그러한 복수의 반사들 (예를 들어, 오디오 신호의 지연된 인스턴스들) 을 포함할 수도 있다. 반향 음성은 먼대면 대화에서 들리는 음성보다 (예를 들어, 다양한 음성 경로를 상의 신호 인스턴스들의 상쇄 간섭으로 인해) 일반적으로 더 머플링 (muffling) 되거나/되고, 덜 선명하거나/하고, 덜 이해하기 쉽다. 이러한 효과들은 자동 음성 인식 (ASR) 애플리케이션들 (예를 들어, 계좌 잔고 또는 주식 시세 확인과 같은, 자동화된 비즈니스 거래; 자동화된 메뉴 안내; 자동화된 문의 처리) 에 있어서 특히 문제일 수도 있어, 정확성의 감소를 초래한다. 따라서, 음색에 변화들을 최소화하면서 기록된 신호에 반향 제거 (dereverberation) 동작을 수행하는 것이 바람직할 수도 있다.

**발명의 내용**

**과제의 해결 수단**

[0006] 일반적 구성에 따른, 방향성 컴포넌트를 포함하는 멀티채널 신호 처리의 방법은 잔류 (residual) 신호를 생성하기 위하여 제 1 신호에 대하여 제 1 방향 선택적 처리 동작을 수행하는 단계, 및 향상된 신호를 생성하기 위하여 제 2 신호에 대하여 제 2 방향 선택적 처리 동작을 수행하는 단계를 포함한다. 이 방법은, 생성된 잔류 신호에 기초하여, 인버스 (inverse) 필터의 복수의 필터 계수들을 계산하는 단계, 및 반향 제거된 신호를 생성하기 위하여 향상된 신호에 대하여 반향 제거 동작을 수행하는 단계를 포함한다. 반향 제거 동작은 계산된 복수의 필터 계수들에 기초한다. 제 1 신호는 멀티채널 신호의 적어도 두 개의 채널들을 포함하고, 제 2 신호는 멀티채널 신호의 적어도 두 개의 채널들을 포함한다. 이 방법에서, 제 1 신호에 대하여 제 1 방향 선택적 처리 동작을 수행하는 단계는 제 1 신호의 총 에너지에 상대적인 제 1 신호 내의 방향성 컴포넌트의 에너지를 감소시키는 단계를 포함하고, 제 2 신호에 대하여 제 2 방향 선택적 처리 동작을 수행하는 단계는 제 2 신호의 총 에너지에 상대적인 제 2 신호 내의 방향성 컴포넌트의 에너지를 증가시키는 단계를 포함한다. 이러한 방법을 수행하도록 구성되는 시스템들 및 장치들, 및 이러한 방법을 수행하기 위한 기계-실행 가능한 명령어들을 갖는 컴퓨터-판독가능 매체 또한 개시되어 있다.

[0007] 일반적 구성에 따른, 방향성 컴포넌트를 포함하는 멀티채널 신호 처리를 위한 장치는 잔류 신호를 생성하기 위하여 제 1 신호에 대하여 제 1 방향 선택적 처리 동작을 수행하도록 구성되는 제 1 필터, 및 향상된 신호를 생성하기 위하여 제 2 신호에 대하여 제 2 방향 선택적 처리 동작을 수행하도록 구성되는 제 2 필터를 포함한다. 이 장치는 생성된 잔류 신호로부터의 정보에 기초하여, 인버스 필터의 복수의 필터 계수들을 계산하도록 구성되는 계산기, 및 계산된 복수의 필터 계수들에 기초하여, 반향 제거된 신호를 생성하기 위하여 향상된 신호를 필터링하도록 구성된 제 3 필터를 포함한다. 제 1 신호는 멀티채널 신호의 적어도 두 개의 채널들을 포함하고, 제 2 신호는 멀티채널 신호의 적어도 두 개의 채널들을 포함한다. 이 장치에서, 제 1 방향 선택적 처리 동작은 제 1 신호의 총 에너지에 상대적인 제 1 신호 내의 방향성 컴포넌트의 에너지를 감소시키는 것을 포함하고, 제 2 방향 선택적 처리 동작은 제 2 신호의 총 에너지에 상대적인 제 2 신호 내의 방향성 컴포넌트의 에너지를 증가시키는 것을 포함한다.

[0008] 또 다른 일반적 구성에 따른, 방향성 컴포넌트를 포함하는 멀티채널 신호 처리를 위한 장치는 잔류 신호를 생성하기 위하여 제 1 신호에 대하여 제 1 방향 선택적 처리 동작을 수행하기 위한 수단, 및 향상된 신호를 생성하기 위하여 제 2 신호에 대하여 제 2 방향 선택적 처리 동작을 수행하기 위한 수단을 포함한다. 이 장치는 생성된 잔류 신호로부터의 정보에 기초하여, 인버스 필터의 복수의 필터 계수들을 계산하기 위한 수단, 및 반향 제거된 신호를 생성하기 위하여 향상된 신호에 대하여 반향 제거 동작을 수행하기 위한 수단을 포함한다. 이 장치에서, 반향 제거 동작은 계산된 복수의 필터 계수들에 기초한다. 제 1 신호는 멀티채널 신호의 적어도 두 개의 채널들을 포함하고, 제 2 신호는 멀티채널 신호의 적어도 두 개의 채널들을 포함한다. 이 장치에서, 제 1 신호에 대하여 제 1 방향 선택적 처리 동작을 수행하기 위한 수단은 제 1 신호의 총 에너지에 상대적인 제 1 신호 내의 방향성 컴포넌트의 에너지를 감소시키도록 구성되고, 제 2 신호에 대하여 제 2 방향 선택적 처리 동작을 수행하기 위한 수단은 제 2 신호의 총 에너지에 상대적인 제 2 신호 내의 방향성 컴포넌트의 에너지를 증가시키도록 구성된다.

**도면의 간단한 설명**

- [0009] 도 1a 및 도 1b 는 빔형성기 (beamformer) 응답 곡선들의 예들을 도시한다.
- 도 2a 는 일반적 구성에 따른 방법 (M100) 의 플로우차트를 도시한다.
- 도 2b 는 일반적 구성에 따른 장치 (A100) 의 플로우차트를 도시한다.
- 도 3a 및 도 3b 는 발생된 널 (null) 빔들의 예들을 도시한다.
- 도 4a 는 방법 (M100) 의 구현 (M102) 의 플로우차트를 도시한다.
- 도 4b 는 장치 (A100) 의 구현 (A104) 의 블록도를 도시한다.
- 도 5a 는 장치 (A100) 의 구현 (A106) 의 블록도를 도시한다.
- 도 5b 는 장치 (A100) 의 구현 (A108) 의 블록도를 도시한다.



- 도 6a 는 일반적 구성에 따른 장치 (MF100) 의 플로우차트를 도시한다.
- 도 6b 는 또 다른 구성에 따른 방법의 플로우차트를 도시한다.
- 도 7a 는 일반적 구성에 따른 디바이스 (D10) 의 블록도를 도시한다.
- 도 7b 는 디바이스 (D10) 의 구현 (D20) 의 블록도를 도시한다.
- 도 8a 내지 도 8d 는 멀티-마이크로폰 무선 헤드셋 (D100) 의 다양한 그림들을 도시한다.
- 도 9a 내지 도 9d 는 멀티-마이크로폰 무선 헤드셋 (D200) 의 다양한 그림들을 도시한다.
- 도 10a 는 멀티-마이크로폰 통신 핸드셋 (D300) 의 (중심 축에 대한) 횡단면도를 도시한다.
- 도 10b 는 디바이스 (D300) 의 구현 (D310) 의 횡단면도를 도시한다.
- 도 11a 는 멀티-마이크로폰 미디어 플레이어 (D400) 의 도면을 도시한다.
- 도 11b 및 도 11c 는, 각각, 디바이스 (D400) 의 구현들 (D410 및 D420) 의 도면을 도시한다.
- 도 12a 는 멀티-마이크로폰 핸즈프리 (hands-free) 자동차 키트 (D500) 의 도면을 도시한다.
- 도 12b 는 멀티-마이크로폰 필기 디바이스 (D600) 의 도면을 도시한다.
- 도 13a 및 도 13b 는, 각각, 디바이스 (D700) 의 정면도 및 평면도를 도시한다.
- 도 13c 및 도 13d 는, 각각, 디바이스 (D710) 의 정면도 및 측면도를 도시한다.
- 도 14a 및 도 14b 는, 각각, 핸드셋 (D300) 의 구현 (D320) 의 정면도 및 측면도를 도시한다.
- 도 14c 및 도 14d 는, 각각, 핸드셋 (D300) 의 구현 (D330) 의 정면도 및 평면도를 도시한다.
- 도 15 는 오디오 감지 디바이스 (D800) 의 디스플레이 도면을 도시한다.
- 도 16a 내지 도 16d 는 디바이스 (D10) 의 상이한 회의 구현들의 구성들을 도시한다.
- 도 17a 는 어레이 (R100) 의 구현 (R200) 의 블록도를 도시한다.
- 도 17b 는 어레이 (R200) 의 구현 (R210) 의 블록도를 도시한다.

**발명을 실시하기 위한 구체적인 내용**

- [0010] 본 개시물은, 블라인드 소스 분리 (Blind Source Separation; BSS) 를 이용하여 얻어지는 분리된 반향 추정치에 트레이닝된 인버스 필터들과 조합된 빔형성을 이용하는, 멀티 마이크로폰 신호의 반향 제거를 위한 시스템들, 방법들, 장치들, 및 컴퓨터-판독가능 매체의 설명을 포함한다.
- [0011] 그 문맥에 의하여 명백히 제한되지 않는다면, 본원에 사용되는 용어 “신호” 는, 전선, 버스 또는 다른 전송 매체 상에 보내지는 것과 같은 메모리 위치 (또는 메모리 위치들의 세트) 의 상태를 포함하는, 그 임의의 보통 의미를 나타내기 위하여 사용되었다. 그 문맥에 의하여 명백히 제한되지 않는다면, 본원에 사용되는 용어 “발생 (generating)” 은, 컴퓨팅 또는 그 외에 생성 (producing) 과 같은, 그 임의의 보통 의미를 나타내기 위하여 사용되었다. 그 문맥에 의하여 명백히 제한되지 않는다면, 본원에 사용되는 용어 “계산 (calculating)” 은, 컴퓨팅, 평가, 추정, 및/또는 복수의 값들로부터 선택과 같은, 그 임의의 보통 의미를 나타내기 위하여 사용되었다. 그 문맥에 의하여 명백히 제한되지 않는다면, 본원에 사용되는 용어 “획득 (obtaining)” 은 계산, 유도, (예를 들어, 외부 디바이스로부터의) 수신, 및/또는 (예를 들어, 저장 엘리먼트들의 어레이로부터의) 검색 과 같은, 그 임의의 보통 의미를 나타내기 위하여 사용되었다. 본원의 설명 및 청구항들에서 용어 “포함하는 (comprising)” 이 사용되는 것은, 다른 엘리먼트들 또는 동작들을 배제하지 않는다. 용어 “기초하는 (based on)” ( “A 는 B 에 기초하는” 에서와 같이) 은 (i) “~로부터 유도된” (예를 들어, “B 는 A 의 전구체이다”), (ii) “적어도 기초하는” (예를 들어, “A 는 적어도 B 에 기초한다”) 및, 특정 문맥에서 적절하다면, (iii) “~과 동일한” (예를 들어, “A 는 B 와 동일하다”) 의 경우들을 포함하는, 그 임의의 보통 의미를 나타내기 위하여 사용되었다. 유사하게, 용어 “~에 응답하여” 는, “적어도 ~에 응답하여” 를 포함하는, 그 임의의 보통 의미를 나타내기 위하여 사용되었다.
- [0012] 문맥에 의해 그렇지 않다고 나타나지 않는다면, 멀티-마이크로폰 오디오 감지 디바이스의 “위치 (location)” 에 대한 기준은 마이크로폰의 음성적 감지 면의 중심의 위치를 나타낸다. 용어 “채널” 은, 특정 문맥에

따라, 때로는 신호 경로를 나타내기 위해 사용되었고, 다른 때에는, 그러한 경로에 의해 운반되는 신호를 나타내기 위해 사용되었다. 달리 나타나지 아니하면, 용어 “시리즈 (series)” 는 둘 이상의 아이템들의 시퀀스를 나타내기 위하여 사용되었다. 용어 “주파수 컴포넌트” 는 주파수들의 세트 가운데 하나 또는, 신호의 주파수 도메인 표시의 샘플 (예를 들어, 고속 푸리에 변환에 의해 생성된 것과 같은) 또는 신호의 부대역 (subband) (예를 들어, 바크 스케일 (bark scale) 부대역) 과 같은, 신호의 주파수 대역들 가운데 하나를 나타내기 위하여 사용되었다.

[0013] 달리 나타나지 않는다면, 특정 특징을 갖는 장치의 동작의 임의의 개시 또한 유사한 특징 (그 역 또한 같다) 을 갖는 방법을 개시하도록 명백하게 의도된 것이며, 특정 구성에 따른 장치의 동작의 임의의 개시 또한 유사한 구성 (그 역 또한 같다) 에 따른 방법을 개시하도록 명백하게 의도된 것이다. 용어 “구성 (configuration)” 은 그 특정 문맥에 나타나는 대로의 방법, 장치들, 및/또는 시스템과 관련하여 사용될 수도 있다. 용어 “방법”, “프로세스”, “절차”, 및 “기술” 은 특정 문맥에 그와 달리 나타나지 않는다면 일반적으로 상호 교환 가능하게 사용된다. 용어 “장치” 및 “디바이스” 또한 특정 문맥에 그와 달리 나타나지 않는다면 일반적으로 상호 교환 가능하게 사용된다. 용어 “엘리먼트 (element)” 및 “모듈 (module)” 은 보통 더 큰 구성의 일부를 나타내도록 사용된다. 명백히 그 문맥에 의하여 제한되지 않는다면, 용어 “시스템” 은 본원에서, “일반 목적으로 쓰이도록 상호작용하는 엘리먼트의 그룹” 을 포함하는, 임의의 그 보통 의미를 나타내도록 사용된다. 삽입된 부분 내에 참조되는 임의의 그룹들뿐만 아니라, 문서의 일부분의 참조에 의한 임의의 삽입 또한 부분 내에 참조되는 용어들 또는 변수들의, 문서 내에 다른 곳에 나타나는, 정의를 포함하도록 이해되어야 한다.

[0014] 멀티 마이크로폰 신호의 반향 제거는, 빔 형성과 같은, 방향적 변별 (또는 “방향적 선택”) 필터링 기술을 이용하여 수행될 수도 있다. 그러한 기술은, 더 정확한 또는 덜 정확한 공간 레졸루션 (spatial resolution) 으로, 다른 방향들로부터 도착하는 사운드 컴포넌트들 (원하는 소리 컴포넌트의 반사된 인스턴스들을 포함) 로부터 특정 방향에서 도착하는 소리 컴포넌트들을 격리시키도록 사용될 수도 있다. 이러한 분리가 중간 내지 고 주파수들에 대해 일반적으로 잘 작용하는 반면, 저 주파수들에서의 결과는 일반적으로 실망스럽다.

[0015] 저 주파수들에서의 이 실패의 하나의 이유는 보통 소리-감지 소비자 디바이스 형태 인자들 (예를 들어, 무선 헤드셋들, 전화기 핸드셋들, 모바일 전화기들, 개인 휴대정보 단말기들 (PDA)) 상에 이용가능한 마이크로폰 간격이 상이한 방향들로부터 도착하는 저 주파수 컴포넌트들 사이의 양호한 분리를 보장하기에는 너무 작기 때문이다. 신뢰할만한 방향성 변별은 보통 파장에 유사한 어레이 구경 (aperture) 을 필요로 한다. 200 Hz 에서의 저-주파수 컴포넌트에 있어, 파장은 대략 170 cm 이다. 그러나, 보통의 소리-감지 소비자 디바이스에 있어서, 마이크로폰들 사이의 간격은 대략 10 cm 의 실질적 상한을 가질 수도 있다. 더하여, 백색 잡음 이득 제한의 바람직성 (desirability) 은 설계자가 저 주파수들에서 빔 (beam) 을 확장하는 것을 제약할 수도 있다. 센서 잡음 및 풍잡음 (wind noise) 과 같은, 마이크로폰 채널들 사이에 비상관화된 소음의 증폭을, 감소시키거나 피하기 위하여 백색 잡음 이득의 제한이 보통 부가된다.

[0016] 공간 앨리어싱 (spatial aliasing) 을 피하기 위하여, 마이크로폰들 사이의 거리는 최소 파장의 절반을 초과해서는 안 된다. 예를 들어, 8 kHz 샘플링 레이트는 0 내지 4 kHz 의 대역폭을 제공한다. 4 kHz 에서의 파장은 대략 8.5 cm 로, 따라서 이러한 경우에, 인접한 마이크로폰들 사이의 공간은 대략 4 cm 를 초과해서는 안 된다. 마이크로폰 채널들은 공간 앨리어싱을 증가시킬 수 있는 주파수들을 제거하기 위해 저역통과 필터링될 수도 있다. 공간 앨리어싱이 고 주파수들에서 공간적으로 선택적인 필터링의 효율성을 감소시킬 수도 있는 반면, 그러나, 반향 에너지는 보통 낮은 주파수들에 (예를 들어, 일반적 방의 기하학적 구조들 때문에) 집중된다. 방향 선택적 필터링 동작이 중간 및 고 주파수들에서의 충분한 반향의 제거를 수행할 수도 있으나, 그러나 저 주파수들에서의 반향제거 성능은 원하는 지각 (perceptual) 이득을 생성하기에 불충분할 수도 있다.

[0017] 도 1a 및 도 1b 는 4개의 인접 마이크로폰들 사이에 3.5 cm 의 간격을 갖는 4개의 마이크로폰 선형 어레이를 이용하여 기록된 멀티 마이크로폰 신호 상에서 얻어진 빔형성기 응답 곡선들을 도시한다. 도 1a 는 어레이 축에 대해 90 도의 스티어링 (steering) 방향에 대한 응답을 도시하고, 도 1b 는 어레이 축에 대해 0 도의 스티어링 방향에 대한 응답을 도시한다. 양 도면에서, 주파수 범위는 0 에서 4 kHz 이며, 낮은 이득에서 높은 이득으로는 어두운 것에서 밝은 것으로 명도에 의해 나타내어진다. 이해를 돕기 위해, 도 1a 에서 가장 높은 주파수에 경계선이 추가되었고, 메인 로브 (main lobe) 의 외곽선이 도 1b 에 추가되었다. 각각의 도면에서, 빔 패턴이 중간 및 고 주파수들에서 높은 방향성을 제공하는 것을 관찰할 수도 있으나, 저 주파수들에서는 빔 패턴이 퍼져나가는 것을 관찰할 수도 있다. 따라서, 반향 제거를 제공하기 위한 그러한 빔들의 애플리케이션은 중간 및 고 주파수들에서 효과적일 수도 있으나, 반향 에너지가 집중되는 경향이 있는, 저 주파수

대역에서는 덜 효과적이다.

- [0018] 다르게는, 멀티 마이크로폰 신호의 반향 제거는 반향 측정치들의 직접 인버스 필터링에 의해 수행될 수도 있다. 그러한 접근은,  $Y(t)$  는 관찰된 음성 신호를 의미하고,  $S(t)$  는 직접-경로 음성 신호를 의미하며,  $C(z^{-1})$  는 인버스 룸-응답 (room-response) 필터를 의미하는,  $C(z^{-1})Y(t) = S(t)$  와 같은 모델을 사용할 수도 있다.
- [0019] 보통의 직접 인버스 필터링 접근은, 의미 있는 솔루션으로 수립하기 위하여 각각의 수량의 분포함수 (예를 들어, 음성의 확률 분포함수 및 복원 에러의 확률 분포함수) 에 대한 적절한 가정을 이용하여, 직접-경로 음성 신호  $S(t)$  및 인버스 룸-응답 필터  $C(z^{-1})$  를 동시에 추정할 수도 있다. 그러나, 이 두 무관한 양들의 동시 추정은 문제가 될 수도 있다. 예를 들어, 그러한 접근은 반복적일 수 있고, 광범위한 계산으로 이어질 수도 있으며 보통 매우 정확하지 않은 결과로의 느린 수렴으로 이어질 수도 있다. 이러한 방법으로 기록된 신호에 직접적으로 인버스 필터링을 적용하는 것은 음성 포먼트 (formant) 구조를 백색화 (whitening) 하면서 룸 임펄스 응답 함수를 반전하는 경향이 있으며, 이는 음성이 부자연스럽게 들리게 한다. 이러한 백색화 아티팩트 (artifact) 를 막기 위해, 직접 인버스 필터링은 파라미터 튜닝에 지나치게 의존할 수도 있다.
- [0020] 블라인드 소스 분리 (BSS) 또는 다른 비상관화 (decorrelation) 기술을 이용하여 추정되는 반향 신호에 기초한 인버스 필터링을 수행하는 멀티-마이크로폰 반향제거를 위한 시스템들, 방법들, 장치들, 및 컴퓨터-판독가능 매체가 본원에 개시된다. 이러한 접근은 소스 쪽으로 안내되는 널 (null) 빔을 계산하기 위해 BSS 또는 다른 비상관화 기술을 이용함으로써 반향을 추정하는 것을 포함할 수도 있고, 그리고 인버스 룸-반향 필터를 추정하기 위해 잔류 신호 (예를 들어, 저 주파수 반향 잔류 신호) 결과로부터의 정보를 이용함으로써 반향을 추정하는 것을 포함할 수도 있다.
- [0021] 도 2a 는, 일반적 구성에 따른, 방향성 컴포넌트 (예를 들어, 사용자의 입에 의해 발산되는 음성신호와 같은, 원하는 신호의 직접-경로 인스턴스) 를 포함하는 멀티채널 신호 처리의 방법 (M100) 의 플로우차트를 도시한다. 방법 (M100) 은 태스크들 (task) (T100, T200, T300, 및 T400) 을 포함한다. 태스크 (T100) 는 잔류 신호를 생성하기 위하여 제 1 신호에 대하여 제 1 방향 선택적 처리 (DSP) 동작을 수행한다. 제 1 신호는 멀티채널 신호의 적어도 두 개의 채널들을 포함하며, 제 1 DSP 동작은 제 1 신호의 총 에너지에 상대적인 제 1 신호 내의 방향성 컴포넌트의 에너지를 감소시킴으로써 잔류 신호를 생성한다. 제 1 DSP 동작은, 예를 들어, 방향성 컴포넌트에 음 이득을 적용함으로써, 및/또는 신호의 하나 이상의 다른 컴포넌트들에 양 이득을 적용함으로써, 방향성 컴포넌트의 상대적인 에너지를 감소시키도록 구성될 수도 있다.
- [0022] 일반적으로, 제 1 DSP 동작은 신호의 총 에너지에 상대적인 방향성 컴포넌트의 에너지를 감소시키도록 구성된 임의의 비상관화 동작으로서 구현될 수도 있다. 예시들은 빔형성 동작 (널 빔 형성 동작으로써 구성되는), 방향성 컴포넌트를 분리해 내도록 구성되는 블라인드 소스 분리 동작, 및 방향성 컴포넌트의 주파수 컴포넌트를 감소하도록 구성된 위상-기반 동작을 포함한다. 그러한 동작은 시간 도메인 또는 변환 도메인 (예를 들어, FFT 또는 DCT 도메인 또는 다른 주파수 도메인) 에서 실행되도록 구성될 수도 있다.
- [0023] 한 예에서, 제 1 DSP 동작은 널 빔형성 동작을 포함한다. 이 경우에, 방향성 컴포넌트의 도착 방향 (예를 들어, 제 1 신호를 생성하는 마이크로폰 어레이에 상대적인 사용자의 입의 방향) 에서 널 빔을 계산함으로써 잔류물이 얻어진다. 널 빔형성 동작은 고정될 수도 있거나/있고 적응적일 수도 있다. 그러한 널 빔형성을 수행하기 위해 사용될 수도 있는 고정된 빔형성 동작들의 예들은 딜레이-앤드-섬 (delay-and-sum) 빔형성을 포함하고, 딜레이-앤드-섬 빔형성은 시간-도메인 딜레이-앤드-섬 빔형성 및 부대역 (subband) (예를 들어, 주파수-도메인) 위상 변이-앤드-섬 빔형성, 및 초지향성 빔형성을 포함한다. 그러한 널 빔형성 동작을 수행하기 위해 사용될 수도 있는 적응적 빔형성 동작들의 예들은 최소 분산 비왜곡 응답 (MVDR) 빔형성, 선형 제약 최소 분산 (LCMV) 빔형성, 및 범용 사이드로브 소거기 (generalized sidelobe canceller; GSC) 빔형성을 포함한다.
- [0024] 또 다른 예에서, 제 1 DSP 동작은 제 1 신호의 상이한 채널들 내의 주파수 컴포넌트의 위상 간의 차이에 기초하는 제 1 신호의 주파수 컴포넌트에 이득을 적용하는 것을 포함한다. 그러한 위상-차-기반 동작은, 제 1 신호의 복수의 상이한 주파수 컴포넌트들 각각에 대하여, 제 1 신호의 상이한 채널들 내의 주파수 컴포넌트의 대응 위상들 사이의 차를 계산하는 단계, 및 계산된 위상 차에 기초하여 주파수 컴포넌트들에 상이한 이득을 적용하는 단계를 포함할 수도 있다. 그러한 위상 차로부터 유도될 수도 있는 방향 표지들의 예들은 도착 방향 및 도착 시간차를 포함한다.
- [0025] 위상-차-기반 동작은 위상 차가 특정 관정 기준 (예를 들어, 특정 범위 내에 떨어지는 도착 대응 방향, 또는 특

정 범위 내에 떨어지는 도착의 대응 시간차, 또는 특정 범위 내에 떨어지는 주파수에 대한 위상차 비율) 을 만족시키는 주파수 컴포넌트들의 숫자에 따른 코히런스 (coherency) 측정을 계산하도록 구성될 수도 있다. 완벽하게 코히런트 (coherent) 한 신호에 대해, 주파수에 대한 위상 차의 비율은 상수이다. 이러한 코히런스 측정은 방향성 컴포넌트가 활성 (예를 들어, 목소리 활동 검출기) 일 때 인터벌들 (interval) 을 나타내기 위하여 사용될 수도 있다. 이러한 동작을, 특정 주파수 범위 (예를 들어, 대략 500, 600, 700, 또는 800 Hz 부터 대략 1700, 1800, 1900, 또는 2000 Hz 까지와 같은, 화자의 목소리의 대부분의 에너지를 포함할 것으로 기대될 수도 있는 범위) 의 주파수 컴포넌트들만의 위상 차들에 기초한 코히런스 측정을 계산하도록 구성하거나/하고, 원하는 화자의 목소리의 피치 (pitch) 주파수의 현재 추정치의 다중의 주파수 컴포넌트들만의 위상 차들에 기초한 코히런스 측정을 계산하도록 구성하는 것이 바람직할 수도 있다.

[0026] 추가적 예에서, 제 1 DSP 동작은 블라인드 소스 분리 (BSS) 동작을 포함한다. 블라인드 소스 분리는, 출력들 사이의 상호 정보가 최소화되는 정도로 분리된 출력들을 비상관화하는 분리 필터 솔루션을 계산하기 때문에, 특정 시나리오에서 반향을 추정하는 유용한 방법을 제공한다. 이러한 동작은 적응적이어서 발산하는 소스가 시간에 따라 이동함에 따라 방향성 컴포넌트의 에너지의 신뢰성있는 분리를 계속할 수도 있다.

[0027] 전통적인 빔형성 기술들대로 원하는 소스로 빔밍 (beaming) 하는 대신, BSS 동작이 다른 경쟁 방향들로 빔밍함으로써 원하는 소스 쪽으로 빔을 발생시키도록 설계될 수도 있다. 잔류 신호는 잡음으로부터 얻어질 수도 있고, 또는 방향성 컴포넌트의 에너지가 분리되는 (예를 들어, 잡음성 신호 출력과는 대조적으로, 방향성 컴포넌트의 에너지가 분리되어 들어가는), BSS 동작의 "잔류" 출력으로부터 얻어질 수도 있다.

[0028] 각각 개개의 주파수 빈 (frequency bin) 의 빔패턴들을 반복적으로 형성하기 위하여 제약된 BSS 접근을 사용하도록 제 1 DSP 동작을 구성하여, 비상관화된 잡음에 대하여 상관된 잡음을 트레이드 오프 (trade off) 하고 메인 빔에 대하여 사이드로브들을 트레이드 오프하는 것이 바람직할 수도 있다. 그러한 결과를 얻기 위하여, 모든 관찰각들 (look angle) 에 대해 노말라이제이션 (normalization) 절차를 이용하여 원하는 관찰 방향에 수렴된 빔들을 단위 이득으로 조정하는 것이 바람직할 수도 있다. 또한 각각의 널빔 방향 내 주파수 빈당 반복 프로세스 동안 강화된 널빔들의 깊이 및 빔폭 (beam width) 을 직접 제어하기 위하여 튜닝 매트릭스를 사용하는 것이 바람직할 수도 있다.

[0029] MVDR 설계와 같이, BSS 설계 자체로는 마이크로폰 어레이의 앞과 뒤 사이의 불충분한 변별을 제공할 수도 있다. 따라서, BSS 동작이 마이크로폰 어레이 전면의 소스들과 그 후면의 소스들 사이를 변별하기 위해 바람직한 애플리케이션들에 있어서, 다른 마이크로폰들과는 다른 방향을 향하는 적어도 하나의 마이크로폰을 포함하도록 어레이를 구현하는 것이 바람직할 수도 있고, 상기 적어도 하나의 마이크로폰은 후면으로부터의 소스들을 나타내기 위하여 사용될 수도 있다.

[0030] 수렴 시간을 감소시키기 위하여, BSS 동작은 방향성 컴포넌트의 추정된 방향을 나타내는 초기 조건들의 세트로 보통 초기화된다. 초기 조건들은 빔형성기(예를 들어, MVDR 빔형성기) 로부터, 및/또는 마이크로폰 어레이를 이용하여, 얻어지는 하나 이상의 방향성 소스들의 기록들에 디바이스를 트레이닝함으로써 얻어질 수도 있다. 예를 들어, 마이크로폰 어레이는 트레이닝 데이터 (training data) 를 얻기 위해 하나 이상의 확성기들의 어레이로부터 신호들을 기록하는데 사용될 수도 있다. 만약 특정 관찰 방향들로 빔들을 발생시키는 것이 바람직하다면, 어레이에 대한 그 각도에 배치될 수도 있다. 그 결과로 초래된 빔의 빔폭은 간섭하는 확성기들의 근접도에 의하여 결정될 수도 있으나, 제약된 BSS 규칙은 경쟁하는 소스들을 널 아웃 (null out) 하려 하고, 따라서 간섭하는 확성기들의 상대적 각 (angular) 거리에 의하여 결정되는 더 또는 덜 좁은 잔류 빔을 초래할 수도 있다.

[0031] 빔폭들은, 그 기하구조에 따라 공간으로 소리를 퍼뜨리는, 상이한 표면들 및 곡물에 확성기들을 이용함으로써 영향을 받을 수 있다. 이러한 응답들을 형성하기 위하여 마이크로폰들의 수와 같거나 더 적은 수의 소스 신호들이 사용될 수 있다. 확성기들에 의하여 재생되는 상이한 사운드 파일들이 상이한 주파수 콘텐츠를 생성하기 위해 사용될 수도 있다. 만약 확성기들이 상이한 주파수 콘텐츠를 포함한다면, 특정 대역들에서의 주파수 손실을 보상하기 위한 재생 전에 재생된 신호가 등화 (equalize) 될 수 있다.

[0032] BSS 동작은 방향적으로 제약될 수도 있어서, 특정 시간 인터벌 동안, 동작은 특정 방향으로부터 도착하는 에너지만을 분리한다. 다르게는, 특정 시간 인터벌 동안, BSS 동작이 상이한 주파수들에서 다소 상이한 방향들로부터 도착하는 에너지를 분리하도록 허용하기 위하여, 그러한 제약은 어느 정도 완화될 수도 있고, 이는 실세계 (real world) 조건들에서 더 양호한 분리 성능을 생성할 수도 있다.

[0033] 도 3a 및 3b 는 마이크로폰 어레이에 대한 소리 소스 (예를 들어, 사용자의 입) 의 상이한 공간적 구성들을 위한 BSS 를 이용하여 생성되는 널 빔들의 예를 도시한다. 도 3a 에 있어서, 원하는 소리 소스는 어레이 축에 대하여 30° 에 있고, 도 3b 에 있어서, 원하는 소리 소스는 어레이 축에 대하여 120° 에 있다. 이 예시들 모두에서, 주파수 범위는 0 내지 4 kHz 이고, 낮은 이득에서 높은 이득까지는 어두운 것에서 밝은 것으로의 명도에 의하여 표시되었다. 이해를 돕기 위하여 각각의 도면에서 가장 높은 주파수와 낮은 주파수에 윤곽선들이 추가되었다.

[0034] 태스크 (T100) 에서 수행되는 제 1 DSP 동작은 원하는 소스를 향해 충분히 날카로운 널 빔을 생성할 수도 있는 반면, 이 공간적 방향은 모든 주파수 대역들, 특히 저-주파수 대역 (예를 들어, 대역 반향 축적으로 인해) 에서, 매우 잘 정의되지 않을 수도 있다. 상기에 언급한 바와 같이, 방향 선택적 처리 동작들은, 특히 작은 형태 인자들을 가지고 있어 마이크로폰 어레이의 넓이가 저-주파수 컴포넌트들의 파장보다 훨씬 작은 디바이스들에 있어서, 저 주파수들에서 보통 덜 효과적이다. 따라서, 태스크 (T100) 에서 수행되는 제 1 DSP 동작은 제 1 신호의 중간 및 고 주파수 대역들로부터 방향성 컴포넌트의 반향을 제거하는데 효과적일 수도 있으나, 방향성 컴포넌트의 저-주파수 잔향 제거에는 덜 효율적일 수도 있다.

[0035] 태스크 (T100) 에 의하여 생성되는 잔류 신호는 원하는 음성 신호의 구조를 덜 포함하기 때문에, 이 잔류 신호에 트레이닝된 인버스 필터는 음성 포먼트 (formant) 구조를 덜 반전할 것이다. 따라서, 트레이닝된 인버스 필터를 기록된 신호들 또는 향상된 신호들에 적용하는 것은 인공적 음성 효과들을 생성하지 않고 고-품질 반향 제거를 생성하도록 기대될 수도 있다. 또한 잔류 신호로부터 방향성 컴포넌트를 억제하는 것은 방향성 컴포넌트의 동시 추정 없이 인버스 룸 임펄스 응답의 추정을 가능하게 하고, 이는 전통적 인버스 필터링 접근들에 비하여 인버스 필터 응답 함수의 더욱 효과적인 계산을 가능하게 할 수도 있다.

[0036] 태스크 (T200) 는 룸-응답 전달 함수 ("룸 임펄스 응답 함수" 라 불리기도 한다)  $F(z)$  의 역을 계산하기 위하여 태스크 (T100) 에서 얻어진 잔류 신호로부터의 정보를 이용한다. 기록된 신호  $Y(z)$  (예를 들어, 멀티채널 신호) 가 원하는 방향성 신호  $S(z)$  의 직접-경로 인스턴스와 방향성 신호  $S(z)$  의 반향된 인스턴스의 합으로써 모델링된다 가정하면:

$$Y(z) = S(z) + S(z)F(z) = S(z)(1 + F(z)).$$

[0037]

[0038] 이 모델은 기록된 신호  $Y(z)$  에 관하여 방향성 신호  $S(z)$  를 표현하도록 재배치될 수도 있다:

$$S(z) = \frac{1}{F(z) + 1} Y(z).$$

[0039]

[0040] 또한 룸-응답 전달 함수  $F(z)$  가 전극 (all-pole) 필터  $1/C(z)$  로써 모델링될 수 있다 가정하면, 인버스 필터  $C(z)$  는 유한-임펄스-응답 (FIR) 필터이다:

$$C(z) = 1 + \sum_{i=1}^q c_i z^{-1}.$$

[0041]

[0042] 이 두 개의 모델들이 원하는 신호  $S(z)$  에 대한 다음 수식을 얻기 위해 조합되면:

$$S(z) = \frac{C(z)}{C(z) + 1} Y(z).$$

[0043]

[0044] 임의의 반향의 부재시 (예를 들어, 모든 필터 계수들  $c_i$  가 0과 같을 때), 함수들  $C(z)$  및  $F(z)$  는 1과 같다.

상기의 수식에서, 이 조건은  $S(z) = Y(z)/2$  의 결과를 낳는다. 따라서, 다음과 같이, 기록된 신호  $Y(z)$  및 인버스 필터  $C(z)$  에 관하여, 음성 신호  $S(z)$  의 모델을 획득하기 위하여 2의 노말라이제이션 인자를 포함하는 것이 바람직할 수도 있다:

$$S(z) = \frac{2C(z)}{C(z) + 1} Y(z).$$

[0045]

[0046]

일 예에서, 태스크 (T200) 는 자기회귀 모델 (autoregressive model) 을 계산된 잔류물에 맞춤으로써 인버스 필터  $C(z)$  의 필터 계수들  $c_i$  를 계산하도록 구성될 수 있다. 예를 들어, 그러한 모델은,  $r(t)$  는 계산된 잔류 신호를 의미하고  $e(t)$  는 백색 잡음 시퀀스를 의미하는,  $C(z)r(t) = e(t)$  로 표현될 수도 있다. 또한 이 모델은 다음과 같이 표현될 수도 있는데

[0047]

$$r[t] - \sum_{i=1}^q c_i r[t-i] = e[t],$$

[0048]

기호 "a[b]" 는 시간 b 에서의 시간-도메인 시퀀스 a의 값을 나타내고 필터 계수들  $c_i$  는 모델의 파라미터들이다. 모델의 차수 q 는 고정될 수도 있고 적응성일 수도 있다.

[0049]

태스크 (T200) 는 임의의 적합한 방법을 사용하여 그러한 자기회귀 모델의 파라미터들  $c_i$  를 계산하도록 구성될 수도 있다. 일 예에서, 태스크 (T200) 는 모델에 대하여 (예를 들어, 에러  $e(t)$  의 에너지를 최소화하기 위하여) 최소-자승 최소화 동작을 수행할 수도 있다. 모델 파라미터들  $c_i$  를 계산하기 위하여 사용될 수도 있는 다른 방법들은 전후진 (forward backward) 접근, 율-워커 방법 (the Yule-Walker method), 및 Burg 방법을 포함한다.

[0050]

0이 아닌  $C(z)$  를 얻기 위하여, 태스크 (T200) 는 에러  $e(t)$  에 대한 분포 함수를 가정하도록 구성될 수도 있다. 예를 들어,  $e(t)$  는 최대 우도 (maximum likelihood) 함수에 따라 분포되는 것으로 가정될 수도 있다. 태스크 (T200) 가  $e(t)$  를 성긴 임펄스 트레인 (sparse impulse train) (예를 들어, 가능한 한 적은 임펄스들, 또는 가능한 한 많은 0들을 포함하는 일련의 델타 함수들) 으로 제약하도록 구성하는 것이 바람직할 수도 있다.

[0051]

모델 파라미터들  $c_i$  는 잔류물에서 배워지는 백색화 필터를 정의하기 위해 고려될 수도 있고, 에러  $e(t)$  는 잔류물  $r(t)$  을 상승시키는 가상 여기 (hypothetical excitation) 신호로 간주될 수도 있다. 이러한 문맥에서, 필터  $C(z)$  계산의 프로세스는 LPC 음성 포먼트 구조 모델링에서 여기 (excitation) 벡터를 찾는 프로세스와 유사하다. 따라서, LPC 분석에 사용되었던 하드웨어 또는 펌웨어를 이용하여 필터 계수들  $c_i$  를 푸는 것이 가능할 수도 있다. 잔류 신호는 음성 신호의 직접-경로 인스턴스를 제거함으로써 계산되었기 때문에, 모델 파라미터 추정 동작은 음성 포먼트 구조를 반전하려 하지 않고 림 전달 함수  $F(z)$  의 극점들을 추정할 수 있을 것이라 기대될 수도 있다.

[0052]

태스크 (T100) 에 의하여 생성된 잔류 신호의 저-주파수 컴포넌트들은 방향성 컴포넌트의 반향 에너지의 대부분을 포함하는 경향이 있다. 방법 (M100) 의 구현을 잔류 신호의 중간 및/또는 고 주파수 에너지의 양을 더 감소시키도록 구성하는 것이 바람직할 수도 있다. 도 4a 는 태스크 (T150) 를 포함하는 방법 (M100) 의 구현 (M102) 의 예를 도시한다. 태스크 (T150) 는 태스크 (T200) 의 잔류 신호 업스트림 (upstream) 에 대해 저역 통과 필터링을 수행하여, 태스크 (T200) 에서 계산된 필터 계수들이 이 필터링된 잔류에 기초하도록 한다. 방법 (M100) 의 관련된 대체적 구현에서, 태스크 (T100) 에서의 제 1 방향 선택적 처리 동작은 저역 통과 필터링 동작을 포함한다. 양쪽의 경우에, 저역 통과 필터링 동작이, 예를 들어, 500, 600, 700, 800, 900 또는 1000 Hz 의 차단 주파수를 갖도록 하는 것이 바람직할 수도 있다.

[0053]

태스크 (T300) 는, 향상된 신호 생성을 위해, 제 2 신호에 대해, 제 2 방향 선택적 처리 동작을 수행한다. 제 2 신호는 멀티채널 신호의 적어도 두 개의 채널들을 포함하고, 제 2 DSP 동작은 제 2 신호의 총 에너지에 상대적인 제 2 신호 내의 방향성 컴포넌트의 에너지를 증가시킴으로써 향상된 신호를 생성한다. 제 2 DSP 동작은 방향성 컴포넌트에 양의 이득을 적용시킴으로써, 및/또는 제 2 신호의 하나 이상의 다른 컴포넌트들에 음의 이득을 적용시킴으로써 방향성 컴포넌트의 상대적 에너지를 증가시키도록 구성될 수도 있다. 제 2 DSP 동작은 시간 도메인 또는 변환 도메인 (예를 들어, FFT 또는 DCT 도메인 또는 또 다른 주파수 도메인) 에서 실행되도록 구성될 수도 있다.

[0054] 일 예에서, 제 2 DSP 동작은 빔형성 동작을 포함한다. 이 경우에, 향상된 신호는 방향성 컴포넌트의 도착 방향 (예를 들어, 제 2 신호를 생성하는 마이크로폰 어레이에 상대적인 화자의 입의 방향) 에서 빔을 계산함으로써 얻어진다. 고정될 수도 있거나/있고 적응적일 수도 있는, 빔형성 동작은 태스크 (T100) 와 관련하여 상기에 언급된 임의의 빔형성 예들을 이용하여 구현될 수도 있다. 또한 태스크 (T300) 는 상이한 특정 방향들 (예를 들어, 현재 가장 높은 에너지 또는 가장 높은 SNR 을 생성하는 빔에 따라) 로 안내되는 복수의 빔들 가운데서 빔을 선택하도록 구성될 수도 있다. 또 다른 예에서, 태스크 (T300) 는, 다중 신호 분류 (Multiple Signal Classification; MUSIC) 알고리즘과 같은, 소스 국지화 (source localization) 방법을 이용하여 빔 방향을 선택하도록 구성된다.

[0055] 일반적으로, 딜레이-앤드-섬 또는 MVDR 빔형성기와 같은 전통적 접근은, 빔형성기 출력 에너지가 단위 (unity) 와 함께 제약된 관찰 방향 에너지로 최소화된, 자유장 (free-field) 모델에 기초하는 하나 이상의 빔패턴들을 설계하는데 이용될 수도 있다. 예를 들어, 폐쇄형 MVDR 기술들은, 주어진 관찰 방향, 상호-마이크로폰 거리, 및 잡음 상호-상관 (cross-correlation) 매트릭스에 기초하여 빔패턴들을 설계하는데 사용될 수도 있다. 보통, 그 결과의 설계들은, 잡음 상호-상관 매트릭스의 주파수-의존 대각 로딩 (diagonal loading) 에 의하여 메인 빔에 대해 트레이드 오프될 수도 있는, 원치않는 사이드로브들을 강조한다. 메인 빔폭과 사이드로브 크기 (magnitude) 사이의 트레이드 오프에 대해 더 양호한 제어를 제공할 수도 있는, 선형 프로그래밍 기술들에 의하여 해결되는 특수 제약된 MVDR 비용 함수를 이용하는 것이 바람직할 수도 있다. 마이크로폰 어레이 전면의 소스들과 그 후면의 소스들 사이를 변별하기 위한 제 1 또는 제 2 DSP 동작에 바람직한 애플리케이션들에 있어서, MVDR 설계 자체만으로는 마이크로폰 어레이의 전면과 후면 사이 불충분한 변별을 제공할 수도 있어, 후면으로부터의 소스들을 나타내도록 사용될 수도 있는, 다른 마이크로폰들과는 다른 방향을 향하는 적어도 하나의 마이크로폰을 포함하도록 어레이를 구현하는 것이 바람직할 수도 있다.

[0056] 또 다른 예에서, 제 2 DSP 동작은 제 2 신호의 상이한 채널들의 주파수 컴포넌트의 위상들 사이의 차이에 기초하는 제 2 신호의 주파수 컴포넌트에 이득을 적용시키는 것을 포함한다. 태스크 (T100) 와 관련해 상기에 언급된 임의의 위상-차-기반 예시들을 이용하여 구현될 수도 있는, 그러한 동작은, 제 2 신호의 복수의 상이한 주파수 컴포넌트들 각각에 대하여, 제 2 신호의 상이한 채널들의 주파수 컴포넌트의 상응 (corresponding) 위상들 사이의 차를 계산하는 단계, 및 계산된 위상 차이에 기초하여 주파수 컴포넌트들에 상이한 이득을 적용시키는 단계를 포함할 수도 있다. 제 1 및/또는 제 2 DSP 동작들 (예를 들어, 제 1 필터 (F110) 및/또는 제 2 필터 (F120)) 을 구현하는데 사용될 수도 있는 위상-차-기반 방법들 및 구조들에 관한 정보는, 예를 들어, 미국 특허 출원 번호 제 12/605,158 호 ("SYSTEMS, METHODS, APPARATUS, AND COMPUTER-READABLE MEDIA FOR COHERENCE DETECTION" 이라는 제목의, 2009년 10월 23일자 출원) 및 미국 특허 출원 번호 제 12/796,566 호 ("SYSTEMS, METHODS, APPARATUS, AND COMPUTER-READABLE MEDIA FOR PHASE-BASED PROCESSING OF MULTICHANNEL SIGNAL" 이라는 제목의, 2010년 6월 8일자 출원) 에서 찾아볼 수 있다. 예를 들어, 그러한 방법들은, 위상차, 상이한 어레이 축들을 따르는 마이크로폰들로부터의 신호에 기초한 전-후 (front-to-back) 변별, 및 방향적 소스로부터의 에너지를 (예를 들어, 잔류 신호 계산을 위해) 마스킹 (masking) 하기 위한 상보적 마스킹에 기초한 부대역 이득 제어를 포함한다.

[0057] 제 3 예로, 제 2 DSP 동작은, 태스크 (T100) 와 관련하여 상기에 언급된 임의의 BSS 예들을 이용하여 구현되거나/되고, 초기화되거나/되고, 제약될 수도 있는, 블라인드 소스 분리 (BSS) 동작을 포함한다. 제 1 및/또는 제 2 DSP 동작들을 구현하는데 사용될 수도 있는 BSS 기술들 및 구조들에 관한 추가적 정보는, 예를 들어, 미국 특허 출원 공개 번호 제 2009/0022336 호 (발명자 비썬 등, "SYSTEMS, METHODS, AND APPARATUS FOR SIGNAL SEPARATION" 라는 제목의, 2009년 1월 22일자 공개) 및 미국 특허 출원 공개 번호 제 2009/0164212 호 (발명자 쉐넬 등, "SYSTEMS, METHODS, AND APPARATUS FOR MULTI-MICROPHONE BASED SPEECH ENHANCEMENT" 라는 제목의, 2009년 6월 25일자 공개) 에서 찾아볼 수 있다.

[0058] 제 4 예로, BSS 동작은 태스크 (T100 및 T300) 들 양자를 구현하도록 사용된다. 이 경우에, 잔류 신호는 BSS 동작의 한 출력에서 생성되고 향상된 신호는 BSS 동작은 다른 출력에서 생성된다.

[0059] 또한 제 1 및 제 2 DSP 동작들 양자는 동작으로의 입력 신호의 각각의 채널의 신호 레벨들 사이의 관계 (예를 들어, 제 1 또는 제 2 신호의 채널들의 선형레벨들의 비율, 또는 제 1 또는 제 2 신호의 채널들의 대수적 레벨들의 차) 에 기초하여 신호 방향을 구별하도록 구현될 수도 있다. 그러한 레벨-기반 (예를 들어, 이득-기반 또는 에너지-기반) 동작은, 신호의 복수의 부대역 각각의, 또는 신호의 복수의 주파수 컴포넌트 각각의 신호의 현재 방향을 나타내도록 구성될 수도 있다. 이 경우에, 마이크로폰 채널들의 이득 응답들 (특히, 마이크로

폰들의 이득 응답들) 이 서로에 관하여 잘 조정되는 것이 바람직할 수도 있다.

[0060] 상기에 언급된 바와 같이, 방향적으로 선택적인 처리 동작들은 보통 저 주파수들에서 덜 효과적이다. 따라서, 태스크 (T300) 에서 수행되는 제 2 DSP 동작이 원하는 신호의 중간 및 고 주파수들에서 효과적으로 반향 제거를 할 수도 있는 반면, 이 동작은 반향 에너지의 대부분을 포함한다 기대될 수도 있는 저 주파수들에서 덜 효과적일 수 있다.

[0061] 빔형성, BSS 또는 마스킹 동작의 방향성의 손실은 보통 주파수가 감소함에 따라 이득 응답의 메인로브의 폭이 증가함으로써 명시된다. 메인로브의 폭은, 예를 들어, 이득 응답이 최대치로부터 3 데시벨 떨어지는 지점들 간의 각도으로써 취해질 수도 있다. 제 1 및/또는 제 2 DSP 동작의 방향성의 손실을, 주파수 감소에 따라, 특정 주파수에서의 동작의 최소 이득 응답들과 최대 이득 응답들 사이의 차의 절대값에서의 감소로써 설명하는 것이 바람직할 수도 있다. 예를 들어, 차의 절대값은 저-주파수 범위 (예를 들어, 300 Hz 내지 400 Hz) 에 대해서 보다 중간 및/또는 고-주파수 범위 (예를 들어, 2 내지 3 kHz) 에서 더 크도록 기대될 수도 있다.

[0062] 다르게는, 제 1 및/또는 제 2 DSP 동작의 방향성의 상실을, 주파수가 감소함에 따라, 방향에 관하여, 동작의 최소 이득 응답들과 최대 이득 응답들 사이의 차의 절대값의 감소로써 설명하는 것이 바람직할 수도 있다. 예를 들어, 이 차의 절대값은 저-주파수 범위 (예를 들어, 300 Hz 내지 400 Hz) 에 대해서 보다 중간 및/또는 고-주파수 범위 (예를 들어, 2 내지 3 kHz) 에서 더 크도록 기대될 수도 있다. 다르게는, 중간 및/또는 고-주파수 범위 (예를 들어, 2 내지 3 kHz) 에 대한, 범위 내의 각각의 주파수 컴포넌트에서의 이 차의 절대값의 평균이, 저-주파수 범위 (예를 들어, 300 내지 400 Hz) 에 대한, 범위 내의 각각의 주파수 컴포넌트에서의 이 차의 절대값의 평균보다, 더 크도록 기대될 수도 있다.

[0063] 태스크 (T400) 는 반향 제거된 신호를 생성하기 위해 상향된 신호에 반향 제거 동작을 수행한다. 반향 제거 동작은 계산된 필터 계수들  $c_i$  에 기초하고 태스크 (T400) 는 시간 도메인 또는 변환 도메인 (예를 들어, FFT 또는 DCT 도메인 또는 다른 주파수 도메인) 에서 반향 제거 동작을 수행하도록 구성될 수도 있다. 한 예에서, 태스크 (T400) 는 다음과 같은 표현에 따라 반향 제거 동작을 수행하도록 구성되고

$$D(z) = \frac{2C(z)}{C(z) + 1} G(z),$$

[0064]

[0065]  $G(z)$  는 향상된 신호 (S40) 를 나타내며  $D(z)$  는 반향 제거된 신호 (S50) 를 나타낸다. 또한 그러한 동작은 시간-도메인 차 공식으로써 표현될 수도 있는데

$$d[t] = g[t] + \sum_{i=1}^q c_i(g[t - i] - 0.5d[t - i]),$$

[0066]

[0067] 여기서  $d$  와  $g$  는 각각, 시간 도메인에서의 반향 제거된 신호 (S50) 와 향상된 신호 (S40) 를 나타낸다.

[0068] 상기에 언급된 바와 같이, 태스크 (T100) 에서의 제 1 DSP 동작은 제 1 신호의 중간 및 고-주파수 대역들로부터 방향성 컴포넌트의 반향을 제거하는데 효과적일 수도 있다. 따라서, 태스크 (T200) 에서 수행되는 인버스 필터 계산은 주로 저-주파수 에너지에 기초할 수도 있어, 태스크 (T400) 에서 수행되는 반향 제거 동작이 중간 또는 고 주파수들보다 향상된 신호의 저 주파수들을 더 감쇠하도록 한다. 예를 들어, 태스크 (T400) 에서 수행되는 반향 제거 동작의 이득 응답은 저-주파수 영역 (예를 들어, 300 내지 400 Hz 사이) 에 대한 반향 제거 동작의 평균 이득 응답보다 (예를 들어, 적어도 3,6,9,12, 또는 20 데시벨 만큼) 더 큰 중간 및/또는 고-주파수 영역 (예를 들어, 2 내지 3 kHz 사이) 에 대한 평균 이득 응답을 가질 수도 있다.

[0069] 방법 (M100) 은 일련의 세그먼트들 (segment) 로써 멀티채널 신호를 처리하도록 구성될 수도 있다. 보통의 세그먼트 길이는 범위가 대략 5 또는 10 mSec 에서 대략 40 또는 50 mSec 이고, 세그먼트들은 오버래핑 (overlapping) (예를 들어, 인접 세그먼트들과 25% 또는 50% 오버래핑) 하거나 논오버래핑 (nonoverlapping) 할 수도 있다. 일 특정 예에서, 멀티채널 신호는 일련의 논오버래핑 세그먼트들 또는 각각 10 mSec 의 길이를 갖는 "프레임들" 로 나뉜다. 방법 (M100) 에 의해 처리되는 세그먼트 또한 상이한 동작에 의해 처리되는 더 큰 세그먼트의 세그먼트 (예를 들어, "서브 프레임 (subframe)") 일 수도 있거나, 그 반대일 수도 있다.

[0070] 제 1 방향 선택적 처리 동작의 적응성 구현 (예를 들어, 적응성 빔형성기 또는 적응성 BSS 동작) 은, 각각의 프



레이에서, 또는 덜 빈번한 인터벌로 (예를 들어, 매 5 또는 10 프레임들마다 한번), 또는 특정 이벤트에 대응하여 (예를 들어, 도착 방향의 검출된 변화) 적응 (adaptatioin) 을 수행하도록 구성될 수도 있다. 그러한 동작은, 예를 들어, 필터 계수들의 하나 이상의 상용 세트들을 갱신함으로써, 적응을 수행하도록 구성될 수도 있다. 제 2 방향 선택적 처리 동작의 적응성 구현 (예를 들어, 적응성 빔형성기 또는 적응성 BSS 동작) 은 유사하게 구성될 수도 있다.

[0071] 태스크 (T200) 는 잔류 신호  $r(t)$  의 프레임 동안 또는 다중의 연속되는 프레임들의 윈도우 (window) 동안 필터 계수들  $c_i$  를 계산하도록 구성될 수도 있다. 태스크 (T200) 는 음성 활동 검출 (VAD) 동작 (예를 들어, 에너지-기반 VAD 동작, 또는 상기에 설명된 위상-기반 코히런스 측정) 에 따라 필터 계수들을 계산하는데 사용되는 잔류 신호의 프레임들을 선택하도록 구성될 수도 있어서, 필터 계수들이 반향 에너지를 포함하는 잔류 신호의 세그먼트들에 기초하도록 할 수도 있다. 태스크 (T200) 는, 각각의 프레임에, 또는 각각의 활성 프레임에; 또는 덜 빈번한 인터벌에 (예를 들어, 매 5 또는 10 프레임들마다 한번, 또는 매 5 또는 10 활성 프레임들마다 한번); 또는 몇몇 이벤트 (예를 들어, 방향성 컴포넌트의 도착 방향의 검출된 변화) 에 대응하여, 필터 계수들을 갱신 (예를 들어, 재계산) 하도록 구성될 수도 있다.

[0072] 태스크 (T200) 에서의 필터 계수들의 갱신은 필터 계수들을 획득하기 위해 시간 동안 계산된 값들을 평활화 (smoothing) 하는 것을 포함할 수도 있다. 그러한 임시적 평활화 동작은 다음과 같은 표현에 따라 수행될 수도 있고:

$$c_i[n] = \alpha c_i[n - 1] + (1 - \alpha)c_{in},$$

[0073]  $c_{in}$  은 필터 계수  $c_i$  의 계산된 값을 의미하고,  $c_i[n-1]$  은 필터 계수  $c_i$  의 이전 값을 의미하고,  $c_i[n]$  은 필터 계수  $c_i$  의 갱신된 값을 의미하고, 그리고  $\alpha$  는 0 (즉, 평활화 없음) 에서 1 (즉, 갱신 없음) 까지의 범위의 값을 갖는 평활화 인자를 의미한다. 평활화 인자  $\alpha$  의 일반적 값들은 0.5, 0.6, 0.7, 0.8 및 0.9 를 포함한다.

[0075] 도 2b 는, 방향성 컴포넌트를 포함하는 멀티채널 신호를 처리하기 위한 일반적 구성에 따른, 장치 (A100) 의 블록도를 도시한다. 장치 (A100) 는 잔류 신호 (S30) 를 생성하기 위해 제 1 신호에 (S10) 대해 제 1 방향 선택적 처리 동작 (예를 들어, 본원에 태스크 (T100) 와 관련하여 설명되는 바와 같이) 을 수행하도록 구성되는 제 1 필터 (F110) 를 포함한다. 장치 (A100) 는 상향된 신호 (S40) 를 생성하기 위하여 제 2 신호 (S20) 에 대하여 제 2 방향 선택적 처리 동작 (예를 들어, 본원에 태스크 (T300) 와 관련하여 설명되는 바와 같이) 을 수행하도록 구성되는 제 2 필터 (F120) 또한 포함한다. 제 1 신호 (S10) 는 멀티채널 신호의 적어도 두 개의 채널들을 포함하고, 제 2 신호 (S20) 는 멀티채널 신호의 적어도 두 개의 채널들을 포함한다.

[0076] 장치 (A100) 는, 잔류 신호 (S30) 로부터의 정보에 기초하여, 인버스 필터의 복수의 필터 계수들을 계산 (예를 들어, 본원에 태스크 (T200) 와 관련하여 설명되는 바와 같이) 하도록 구성되는 계산기 (CA100) 또한 포함한다. 장치 (A100) 는, 계산된 복수의 필터 계수들에 기초하여, 반향 제거된 신호 (S50) 를 생성하기 위하여 향상된 신호 (S40) 를 필터링 (예를 들어, 본원에 태스크 (T400) 과 관련하여 설명되는 바와 같이) 하도록 구성되는 제 3 필터 (F130) 또한 포함한다.

[0077] 상기에 언급된 바와 같이, 제 1 및 제 2 DSP 동작들 각각은 시간 도메인 또는 변환 도메인 (예를 들어, FFT 또는 DCT 도메인 또는 다른 주파수 도메인) 에서 실행되도록 구성될 수도 있다. 도 4b 는 필터들 (F110 및 F120) 의 FFT 도메인 업스트림으로의 (변환 모듈들 (TM10a 및 TM20b) 을 통해) 제 1 및 제 2 신호들 (S10 및 S20) 의 변환을 명백히 도시하고, 필터 (F110 및 F120) 의 시간 도메인 다운스트림으로 (인버스 변환 모듈들 (TM20a 및 TM20b) 을 통해) 잔류 신호 (S30) 및 상향된 신호 (S40) 의 후속 변환을 명백히 도시하는, 장치 (A100) 의 구현의 예 (A104) 의 블록도를 도시한다. 또한 방법 (M100) 및 장치 (A100) 는 제 1 방향 선택적 처리 동작 및 제 2 방향 선택적 처리 동작 모듈들, 시간 도메인에서 수행되도록, 또는 제 1 방향 선택적 처리 동작은 시간 도메인에서 수행되고 제 2 방향 선택적 처리 동작은 변환 도메인에서 수행 되도록 (또는 그 반대도 같다), 구현될 수도 있음을 알아야 한다. 추가적 예들은 제 1 및 제 2 방향 선택적 처리 동작들의 하나 또는 양자 모두 내의 변환을 포함하여 동작의 입력 및 출력이 상이한 도메인들 (예를 들어, FFT 도메인에서 시간 도메인으로 변화) 에 있도록 한다.

[0078] 도 5a 는 장치 (A100) 의 구현 (A106) 의 블록도를 도시한다. 장치 (A106) 는 멀티채널의 4-채널 구현 (MCS4) 의 모든 4 개의 채널들을 제 2 신호 (S20) 로써 수신하도록 구성된 제 2 필터 (F120) 의 구현 (F122)

을 포함한다. 한 예에서, 장치 (A106) 는 제 1 필터 (F110) 가 BSS 동작을 수행하고 제 2 필터 (F122) 가 빔형성 동작을 수행하도록 구현된다.

[0079] 도 5b 는 장치 (A100) 의 구현 (A108) 의 블록도를 도시한다. 장치 (A108) 는 제 1 필터 (F110) 및 제 2 필터 (F120) 모두를 포함하도록 구성되는 비상관기 (DC10) 를 포함한다. 예를 들어, 비상관기 (DC10) 는, 일 출력 (예를 들어, 잡음 출력) 에서 잔류 신호를 생성하고 다른 출력 (예를 들어, 분리된 신호 출력) 에서 향상된 신호를 생성하기 위해, 멀티채널 신호의 2-채널 구현 (MCS2) 에 BSS 동작 (예를 들어, 본원에 설명된 임의의 BSS 예들에 따라) 을 수행하도록 구성될 수도 있다.

[0080] 도 6a 는, 일반적 구성에 따른, 방향성 컴포넌트를 포함하는 멀티채널 신호 처리를 위한 장치 (MF100) 의 블록도를 도시한다. 장치 (MF100) 는 잔류 신호를 생성하기 위하여 제 1 신호에 대해 제 1 방향 선택적 처리 동작 (예를 들어, 본원에 태스크 (T100) 와 관련하여 설명된 대로) 을 수행하기 위한 수단들 (F100) 을 포함한다. 또한 장치 (MF100) 는 향상된 신호를 생성하기 위하여 제 2 신호에 대해 제 2 방향 선택적 처리 동작 (예를 들어, 본원에 태스크 (T300) 와 관련하여 설명되는 대로) 을 수행하기 위한 수단들 (F300) 을 포함한다. 제 1 신호는 멀티채널 신호의 적어도 두 개의 채널들을 포함하고, 제 2 신호는 멀티채널 신호의 적어도 두 개의 채널들을 포함한다. 장치 (MF100) 는, 생성된 잔류 신호로부터의 정보에 기초하여, 인버스 필터의 복수의 필터 계수들을 (예를 들어, 본원에 태스크 (T200) 와 관련하여 설명된 대로) 계산하기 위한 수단들 (F200) 또한 포함한다. 장치 (MF100) 는, 계산된 복수의 필터 계수들에 기초하여, 반향 제거된 신호를 생성하기 위하여 향상된 신호에 대해 (예를 들어, 본원에 태스크 (T400) 와 관련하여 설명된 대로) 반향 제거 동작을 수행하기 위한 수단들 (F400) 또한 포함한다.

[0081] 태스크 (T300) 에서 수행되는 (다르게는, 제 2 필터 (F120) 에 의해 수행되는) 멀티채널 방향 선택적 처리 동작은 두 출력들: 방향성 컴포넌트의 에너지가 집중된 잡음성 신호 출력, 및 제 2 신호의 다른 컴포넌트들 (예를 들어, 다른 방향성 컴포넌트들 및/또는 분산된 잡음 컴포넌트) 의 에너지를 포함하는 잡음 출력을 생성하도록 구현될 수도 있다. 빔형성 및 BSS 동작들은, 예를 들어, 보통 그러한 출력들 (예를 들어, 도 5b 에 도시된 대로) 을 생성하도록 구현된다. 그러한 태스크 (T300) 또는 필터 (F120) 의 구현은 향상된 신호로써 잡음성 신호 출력을 생성하도록 구성될 수도 있다.

[0082] 다르게는, 그러한 경우에, 태스크 (T300) 에서 수행되는 (다르게는, 제 2 필터 (F120) 또는 비상관기 (DC10) 에 의해 수행되는) 제 2 방향 선택적 처리 동작을, 잡음성 신호를 출력의 잡음을 더 감소시키기 위하여 잡음 출력을 이용함으로써 향상된 신호를 생성하는 후-처리 (post-processing) 동작을 포함하도록 구현하는 것이 바람직할 수도 있다. 그러한 후-처리 동작 (또한 "잡음 감소 동작" 이라 불리기도 한다) 은, 예를 들어, 잡음 출력의 스펙트럼에 기초하여, 잡음성 신호 출력에 대한 위너 (Wiener) 필터링 동작으로써 구성될 수도 있다. 다르게는, 그러한 잡음 감소 동작은, 향상된 신호를 생성하기 위하여, 잡음성 신호 출력으로부터 잡음 출력에 기초하는 추정된 잡음 스펙트럼을 감산하는 스펙트럼 감산 동작으로써 구성될 수도 있다. 또한 그러한 잡음 감산 동작은, 스펙트럼 감산 또는 신호-대-잡음-비 (SNR) 기반 이득 규칙에 기초하는 부대역 이득 제어 동작으로써 구성될 수도 있다. 그러나, 적극적 (aggressive) 세팅들에서, 그러한 부대역 이득 제어 동작은 음성 왜곡으로 이어질 수도 있다.

[0083] 특정 설계 선택에 따라서, 태스크 (T300) (다르게는, 제 2 필터 (F120)) 은 향상된 신호를 싱글-채널 신호 (예를 들어, 본원에 설명되고 도시된 대로) 또는 멀티채널 신호로써 생성하도록 구현될 수도 있다. 향상된 신호가 멀티채널 신호인 경우에, 태스크 (T400) 는 각각의 채널에 대해 반향 제거 동작의 상응 인스턴스를 수행하도록 구성될 수도 있다. 그러한 경우, 하나 이상의 결과 채널들에 대해, 다른 하나 이상의 결과 채널들에 기초하여, 상기에 설명된 대로 잡음 감소 동작을 수행할 수 있다.

[0084] 멀티채널 신호 처리 (또는 상응 장치) 의 방법을 도 6b 의 플로우차트에 도시된 대로 구현할 수 있고, 상기 플로우 차트에서, 태스크 (T500) 는, 태스크 (T400) 와 관련하여 본원에 설명된 대로, 태스크 (T300) 에 의하여 생성되는 향상된 신호에 대해서보다, 멀티채널 신호의 하나 이상의 채널들에 대해서 반향 제거 동작을 수행한다. 이 경우에, 태스크 (T300) (또는 필터 (F120)) 는 생략되거나 또는 우회될 수도 있다. 태스크 (T300) 의 멀티채널 DSP 동작이 인버스 롬-반응 필터에 기초하는 반향 제거보다 중간 및 고 주파수들에서 방향성 컴포넌트의 더 양호한 반향 제거를 수행할 것으로 기대될 수도 있기에, 방법 (M100) 은 그런 방법 (또는 상응 장치) 보다 더 양호한 결과를 생성하도록 기대될 수도 있다.

[0085] 태스크 (T100) (다르게는, 제 1 필터 (F110)) 에 의해 수행되는 제 1 DSP 동작 및/또는 태스크 (T300) (다르게는, 제 2 필터 (F120)) 에 의해 수행되는 제 2 DSP 동작을 구현하는데 사용될 수도 있는 블라인드 소스 분리

(BSS) 알고리즘들의 범위는, 필터 계수 값들이 주파수 도메인에서 직접 계산되는, 주파수-도메인 ICA 또는 복합 ICA 라 불리는 접근법을 포함한다. 피드포워드 (feedforward) 필터 구조를 이용하여 구현될 수도 있는, 그러한 접근법은 입력 채널들에 대해 FFT 또는 다른 변환을 수행하는 것을 포함할 수도 있다. 이 ICA 기술은 각각의 주파수 빈  $\omega$  에 대해  $M \times M$  언믹싱 (unmixing) 매트릭스  $W(\omega)$  를 계산하도록 설계되어 분리된 (demixed) 출력 벡터들  $Y(\omega, l) = W(\omega)X(\omega, l)$  이 상호 독립적이도록 하고, 여기서  $X(\omega, l)$  은 주파수 빈  $\omega$  및 윈도우  $l$  에 대하여 관찰된 신호를 의미한다. 언믹싱 매트릭스들  $W(\omega)$  은 다음과 같이 표현되는 규칙에 따라 갱신될 수 있으며:

[0086] 
$$W_{l+r}(\omega) = W_l(\omega) + \mu[I - \langle \Phi(Y(\omega, l))Y(\omega, l)^H \rangle]W_l(\omega) \quad (1)$$

[0087]  $W_l(\omega)$  는 주파수 빈  $\omega$  및 윈도우  $l$  에 대한 언믹싱 매트릭스를 의미하고,  $Y(\omega, l)$  은 주파수 빈  $\omega$  및 윈도우  $l$  에 대한 필터 출력을 의미하고,  $W_{l+r}(\omega)$  는 주파수 빈  $\omega$  및 윈도우  $l+r$  에 대한 언믹싱 매트릭스를 의미하고,  $r$  은 1보다 작지 않은 정수 값을 갖는 갱신을 파라미터이고,  $\mu$  는 학습률 (learning rate) 파라미터이고,  $I$  는 항등행렬이고,  $\Phi$  는 활성화 함수를 의미하고, 위첨자  $H$  는 쾨주게이트 (conjugate) 전치 (transpose) 동작을 의미하고, 괄호  $\langle \rangle$  는 시간  $l = 1, \dots, L$  에서의 평균 동작을 의미한다. 한 예에서, 활성화 함수  $\Phi(Y_j(\omega, l))$  는  $Y_j(\omega, l)/|Y_j(\omega, l)|$  과 같다. 잘 알려진 ICA 구현들의 예들은 Infomax, FastICA (온라인 [www-dot-cis-dot-hut-dot-fi/projects/ica/fastica](http://www-dot-cis-dot-hut-dot-fi/projects/ica/fastica) 에서 이용가능), 및 JADE (Joint Approximate Diagonalization of Eigenmatrices) 를 포함한다.

[0088] 그러한 합성된 빔형성기의 각각의 출력 채널  $j$  에 대한 빔 패턴은 표현식

[0089] 
$$W_{j1}(i \times \omega)D(\omega)_{1j} + W_{j2}(i \times \omega)D(\omega)_{2j} + \dots + W_{jM}(i \times \omega)D(\omega)_{Mj}$$

[0090] 의 크기 플롯을 계산함으로써 주파수-도메인 변환 함수  $W_{jm}(i \times \omega)$  ( $m$  은 입력 채널을 의미하고,  $1 \leq m \leq M$ ) 로부터 얻어질 수도 있다. 이 표현식에서,  $D(\omega)$  는 주파수  $\omega$  에 대한 지향성 매트릭스를 나타내어

[0091] 
$$D(\omega)_{ij} = \exp(-i \times \cos(\theta_j) \times pos(i) \times \omega / c), \quad (2)$$

[0092] 와 같이 되도록 하고, 여기서  $pos(i)$  는  $M$  마이크로 폰들의 어레이의  $i$ -번째 마이크로폰의 공간적 좌표들을 의미하고,  $c$  는 매체 내의 소리의 전파 속도 (예를 들어, 공기 중에서 340 m/s) 이고,  $\theta_j$  는 마이크로폰 어레이의 축에 대한  $j$ -번째 소스의 도착 입사각을 나타낸다.

[0093] 복합 ICA 해결책들은 보통, 관찰 방향이 변함에 따라 빔패턴 이득 및/또는 응답 색깔에 변화를 초래할 수도 있는, 스케일링 모호성 (scaling ambiguity) 의 문제가 있다. 만약 소스들이 정지되어 있고 모든 주파수 빈들에 대해 소스들의 변동들이 알려져 있다면, 스케일링 문제는 알려진 값들로 변동들을 조정함으로써 해결될 수도 있다. 그러나, 자연 신호 소스들은 동적이며, 보통 비-정지상태이고, 알려지지 않은 변동들을 갖는다.

[0094] 소스 변동들을 조정하는 대신, 스케일링 문제는 학습된 분리 필터 매트릭스를 조정함으로써 해결될 수도 있다. 최소 왜곡 원칙에 의하여 얻어지는, 한 잘 알려진 해결책은, 다음과 같은 표현식에 따라 학습된 언믹싱 매트릭스를 스케일링한다.

[0095] 
$$W_{l+r}(\omega) \leftarrow \text{diag}(W_{l+r}^{-1}(\omega))W_{l+r}(\omega)$$

[0096] 원하는 관찰 방향에 단위 이득을 생성함으로써 스케일링 문제를 다루는 것이 바람직할 수도 있고, 이는 원하는 화자의 목소리의 주파수 착색 (coloration) 을 감소 또는 피하는데 도움을 줄 수도 있다. 하나의 그러한 접근법은, 모든 각에 대한 필터 응답 크기의 최대값에 의하여, 매트릭스  $W$  의 각 행을 노멀라이징한다:

[0097] 
$$\max_{\theta_j \in [-\pi, \pi]} |W_{j1}(i \times \omega)D(\omega)_{1j} + W_{j2}(i \times \omega)D(\omega)_{2j} + \dots + W_{jM}(i \times \omega)D(\omega)_{Mj}|$$

[0098] 몇몇 복합 ICA 구현들의 또 다른 문제는 동일 소스와 관련된 주파수 빈들 간의 코히런스 (coherence) 의 손실이다. 이 손실은, 정보 소스로부터의 에너지를 주로 포함하는 주파수 빈들이, 간접 출력 채널로 및/또는 그 반대로 잘못 할당되는, 주파수 치환 문제를 일으킬 수도 있다.

[0099] 치환 문제에 사용될 수도 있는 하나의 응답은, 소스를 이용하여 먼저 주파수 빈들 간의 추정된 의존성을 모델링

하는 복합 ICA 의 변형인, 독립적 벡터 분석 (IVA) 이다. 이 방법에서, 활성 함수  $\Phi$  는 다음과 같은 다중 변동성 활성 함수이고:

$$\Phi(Y_j(\omega, l)) = \frac{Y_j(\omega, l)}{\left( \sum_{\omega} |Y_j(\omega, l)|^p \right)^{1/p}}$$

[0100]

[0101] 여기서 p 는 1 과 같거나 더 큰 정수 값 (예를 들어, 1, 2, 또는 3) 을 갖는다. 이 함수에서, 분모의 항은 모든 주파수 빈들에 대한 분리된 소스 스펙트럼들에 관련된 것이다.

[0102] BSS 알고리즘은 자연스레, 원하는 관찰 방향의 에너지만을 남기고, 간섭 소스들을 빔 아웃 (beam out) 하려 시도할 수도 있다. 모든 주파수 빈들에 대한 노말라이제이션 후에, 그러한 동작은 원하는 소스 방향에 단위 이득을 초래할 수도 있다. BSS 알고리즘은 특정 방향에 완벽히 정렬된 빔을 산출하지 못할 수도 있다. 만약 특정 공간적 픽업 (pickup) 패턴으로 빔형성기들을 생성하는 것이 바람직하다면, 그 깊이 및 폭이 각각의 주파수 빈 및 각각의 널 빔 방향에 대한 특정 튜닝 요소들에 의해 보장될 수 있는, 널빔들을 특정 관찰 방향들로 강화시킴으로써, 사이드로브들은 최소화될 수 있고 빔폭들이 형성된다.

[0103] 사이드로브 최소화 및/또는 특정 관찰 방향들의 빔 패턴 정규화 (regularizing) 를 선택적으로 강제함으로써, BSS 알고리즘에 의해 제공되는 미가공 (raw) 빔패턴들을 미세 조정 (fine-tune) 하는 것이 바람직할 수도 있다. 원하는 관찰 방향은, 예를 들어, 어레이 관찰 방향들에 대한 필터 공간적 응답의 최대값을 계산하고 이 최대 관찰 방향 주위에 제약을 강제함으로써 얻어질 수 있다.

[0104] 지향성 매트릭스  $D(\omega)$  (상기 표현식 (2) 에서와 같이) 에 기초하여 정규화 항  $J(\omega)$  를 추가함으로써 빔들 및/또는 널 빔들을 강화하는 것이 바람직할 수도 있고:

$$J(\omega) = S(\omega) \|W(\omega)D(\omega) - C(\omega)\|^2 \quad (3)$$

[0105]

[0106] 여기서  $S(\omega)$  는 주파수  $\omega$  및 각각의 널 빔 방향에 대한 튜닝 매트릭스 이고,  $C(\omega)$  는 원하는 빔 패턴의 선택을 설정하고 각각의 출력 채널 j 에 대한 간섭 방향들에 널들 (null) 을 위치시키는  $\text{diag}(W(\omega) * D(\omega))$  와 동일한  $M \times M$  대각 매트릭스이다. 그러한 정규화는 사이드로브들을 제어하는데 도움이 될 수도 있다. 예를 들어, 매트릭스  $S(\omega)$  는, 각각의 주파수 빈에서 각각의 널 방향의 강화의 양을 제어함으로써, 특정 방향  $\theta_j$  의 각각의 널 빔의 깊이를 형성하도록 사용될 수도 있다. 그러한 제어는, 좁거나 넓은 널 빔들에 대하여 사이드로브들의 발생을 트레이드 오프하기 위해 중요할 수도 있다.

[0107] 정규화 항 (3) 은 언믹싱 매트릭스 갱신 공식에 대한 제약으로써 다음과 같은 표현식으로 표현될 수도 있다:

$$\text{constr}(\omega) = (dJ/dW)(\omega) = \mu * S(\omega) * 2 * (W(\omega) * D(\omega) - C(\omega)) D(\omega)^H$$

[0108]

[0109] 그러한 제약은, 다음 표현식과 같이, 필터 학습 규칙 (예를 들어, 표현식 (1)) 에 그러한 항을 더함으로써 구현될 수도 있다:

$$W_{constr.l+p}(\omega) = W_l(\omega) + \mu [I - \langle \Phi(Y(\omega, l))Y(\omega, l)^H \rangle] W_l(\omega) + 2S(\omega)(W_l(\omega)D(\omega) - C(\omega))D(\omega)^H$$

[0110]

[0111] 소스 도착 방향 (DOA) 값들  $\theta_j$  는 사이드로브들을 제거하기 위하여 수립된 BSS 빔패턴들에 기초하여 결정될 수도 있다. 원하는 애플리케이션에 있어 매우 클 수도 있는 이 사이드로브들을 제거하기 위하여, 선택적 널 빔들을 강화하는 것이 바람직할 수도 있다. 좁혀진 빔은, 각각의 주파수 빈에 특정 매트릭스  $S(\omega)$  를 통해 강화되는 추가적 널 빔을 적용시킴으로써, 얻어질 수도 있다.

[0112] 음향 신호를 수신하도록 구성된 둘 이상의 마이크로폰들의 어레이 (R100) 및 장치 (A100) 의 구현을 갖는 휴대용 오디오 감지 디바이스를 생성하는 것이 바람직할 수도 있다. 그러한 어레이를 포함하도록 구현될 수도 있고 오디오 기록 및/또는 음성 통신 애플리케이션들에 사용될 수도 있는, 휴대용 오디오 감지 디바이스의 예들은, 전화기 헤드셋 (예를 들어, 셀룰러 전화기 헤드셋); 유선 또는 무선 헤드셋 (예를 들어, 블루투스 헤드셋); 휴대용 오디오 및/또는 비디오 기록장치; 오디오 및/또는 비디오 콘텐츠를 기록하도록 구성된 개인용 미디어 플레이어; 개인 정보 단말기 (PDA) 또는 다른 휴대용 컴퓨팅 디바이스; 및 노트북 컴퓨터, 랩탑 컴퓨터, 넷북 컴

퓨터, 테이블 컴퓨터 또는 다른 휴대용 컴퓨팅 디바이스를 포함한다. 장치 (A100) 및 어레이 (R100) 의 인스턴스들을 포함하도록 구축될 수도 있고 오디오 기록 및/또는 음성 통신 애플리케이션들에 사용될 수도 있는, 오디오 감지 디바이스들의 다른 예들은, 셋톱박스들 및 오디오 및/또는 비디오-회의 디바이스들을 포함한다.

[0113] 도 7a 는 일반적 구성에 따른 멀티 마이크로폰 오디오 감지 디바이스 (D10) 의 블록도를 도시한다. 디바이스 (D10) 는 본원에 개시된 마이크로폰 어레이 (R100) 의 임의의 구현을 포함하고, 본원에 개시된 임의의 오디오 감지 디바이스들은 디바이스 (D10) 의 인스턴스로서 구현될 수도 있다. 디바이스 (D10) 는 본원에 개시된 대로 장치 (A100) 의 구현 (예를 들어, 장치 (A100, A104, A106, A108 및/또는 MF100)) 인 장치 (A200) 또한 포함하거나/하고 디바이스 (D10) 는 본원에 개시된 대로 방법 (M100) 의 구현 (예를 들어, 방법 (M100 또는 M102)) 을 수행함으로써 멀티채널 오디오 신호 MCS 를 처리하도록 구성된다. 장치 (A200) 는 하드웨어 및/또는 소프트웨어 (예를 들어, 펌웨어) 로 구현될 수도 있다. 예를 들어, 장치 (A200) 는 디바이스 (D10) 의 프로세서에서 실행되도록 구현될 수도 있다.

[0114] 도 7b 는 디바이스 (D10) 의 구현인 통신 디바이스 (D20) 의 블록도를 도시한다. 디바이스 (D20) 는 장치 (A200) 를 포함하는 칩 또는 칩셋 (CS10) (예를 들어, 모바일 스테이션 모뎀 (MSM) 칩셋) 을 포함한다. 칩/칩셋 (CS10) 은, 장치 (A200) (예를 들어, 명령어들) 의 전부 또는 일부를 실행하도록 구성되는, 하나 이상의 프로세서들을 포함할 수도 있다. 칩/칩셋 (CS10) 은 어레이 (R100) 의 처리 엘리먼트들 (예를 들어, 이하에 설명되는 바와 같이 오디오 전처리 단계 (AP10) 의 엘리먼트들) 또한 포함할 수도 있다. 칩/칩셋 (CS10) 은, 무선 주파수 (RF) 통신 신호를 수신하고 RF 신호 내에 인코딩 (encoding) 된 오디오 신호를 디코딩 (decoding) 및 재생하도록 구성된 수신기, 및 장치 (A200) 에 의해 생성되는 처리된 신호에 기초하는 오디오 신호를 인코딩하고 인코딩된 오디오 신호를 설명하는 RF 통신 신호를 송신하도록 구성된 송신기를 포함한다. 예를 들어, 칩/칩셋 (CS10) 의 하나 이상의 프로세서들은 멀티채널 신호의 하나 이상의 채널에 대해 상기에 설명된 대로 잡음 감소 동작을 수행하도록 구성될 수도 있어서 인코딩된 신호가 잡음-감소된 신호에 기초하도록 한다.

[0115] 어레이 (R100) 의 각각의 마이크로폰은 전방향적 (omnidirectional) 응답, 양방향성 (bidirectional) 응답, 또는 단일 방향성 (unidirectional) (예를 들어, 카디오이드 (cardioid)) 응답을 가질 수도 있다. 어레이 (R100) 에서 사용될 수도 있는 마이크로폰들의 다양한 형태들은 (제한되는 것이 아니라) 압전기 (piezoelectric) 마이크로폰들, 동적 마이크로폰들, 및 일렉트릿 (electret) 마이크로폰들을 포함한다. 핸드셋 또는 헤드셋과 같은, 휴대용 음성 통신을 위한 디바이스에서, 어레이 (R100) 의 인접 마이크로폰들 사이의 중심 간격은, 비록 핸드셋 또는 스마트폰 같은 디바이스에서는 더 큰 간격 (예를 들어, 10 또는 15 cm 까지) 이 가능하고, 테이블 컴퓨터와 같은 디바이스에서는 심지어 더 큰 간격들 (예를 들어, 20, 25, 또는 30 cm 또는 그 이상) 이 가능하기는 하지만, 보통 대략 1.5 cm 내지 대략 4.5 cm 의 범위에 있다. 어레이 (R100) 의 마이크로폰들은 선을 따라서 (균일 또는 비균일 마이크로폰 간격으로) 또는, 다르게는, 어레이 (R100) 의 마이크로폰들의 중심들이 2-차원 (예를 들어, 삼각형의) 또는 3-차원 형태의 꼭짓점들에 놓이도록, 정렬될 수도 있다.

[0116] 마이크로폰들이 소리보다 방사 또는 방출에 민감한 변환기들로서 더욱 일반적으로 구현될 수도 있음을 명백히 알 수 있다. 그러한 한 예에서, 마이크로폰 쌍은 한 쌍의 초음파 변환기들 (예를 들어, 15, 20, 25, 30, 40, 또는 50 kHz 또는 그 이상보다 더 큰 음향 주파수들에 민감한 변환기) 로써 구현된다.

[0117] 도 8a 및 도 8b 는 멀티 마이크로폰 오디오 감지 디바이스 (D10) 의 휴대용 구현 (D100) 의 다양한 모습들을 도시한다. 디바이스 (D100) 는, 어레이 (R100) 의 2-마이크로폰 구현 및 하우징으로부터 연장되는 이어폰 (Z20) 을 지니는 하우징 (Z10) 을 포함하는, 무선 헤드셋이다. 그러한 디바이스는 셀룰러 전화기 핸드셋과 같은 전화기 디바이스와 (예를 들어, the Bluetooth Special Interest Group, Inc., 워싱턴, 벨뷰, 에 의해 배포된 블루투스™ 프로토콜의 버전을 이용하여) 통신을 통해 하프-듀플렉스 또는 풀-듀플렉스 전화를 지원하도록 구성될 수도 있다. 일반적으로, 헤드셋의 하우징은, 직사각형이거나 그렇지 않으면 도 8a, 8b, 및 8d 도시된 바와 같이 (예를 들어, 미니붐 (miniboom) 같은 형상의) 길게 늘어질 수도 있으며, 또는 더 둥글거나 심지어 원형일 수도 있다. 하우징은 배터리 및 프로세서 및/또는 다른 처리 회로망 (예를 들어, 인쇄회로기판 및 그 위에 장착된 컴포넌트들) 또한 포함할 수도 있고, 하우징은 전기 포트 (예를 들어, 미니-범용 직렬 버스 (USB) 또는 배터리 충전을 위한 다른 포트) 및 하나 이상의 버튼 스위치들 및/또는 LED 등과 같은 사용자 인터페이스 기능들을 포함할 수도 있다. 보통 그 주축에 따른 하우징의 길이는 1 내지 3 인치의 범위 내에 있다.

[0118] 보통, 어레이 (R100) 의 각각의 마이크로폰은 디바이스 내의 음향 포트로서 쓰이는 하우징의 하나 이상의 작은

구멍들 뒤에 장착된다. 도 8b 내지 8d 는, 디바이스 (D100) 의 어레이의 1차 마이크로폰을 위한 음향 포트 (Z40) 의 위치, 및 디바이스 (D100) 의 어레이의 2차 마이크로폰을 위한 음향 포트 (Z50) 의 위치를 도시한다.

[0119] 헤드셋은, 보통 헤드셋으로부터 분리가능한 이어 후크 (ear hook) 와 같은, 고정 디바이스 또한 포함할 수도 있다. 외부 이어 후크는, 예를 들어, 사용자가 헤드셋을 양쪽 귀에 사용하기 위해 구성하는 것을 허용하기 위하여, 가역적일 수도 있다. 다르게는, 헤드셋의 이어폰은, 상이한 사용자들이 특정 사용자의 이도 (ear canal) 의 바깥쪽 부분에 더 잘 맞기 위해 상이한 크기 (예를 들어, 지름) 의 이어피스 (earpiece) 를 사용하는 것을 허용하도록 탈착 가능한 이어피스를 포함할 수도 있는, 내부 고정 디바이스 (예를 들어, 귀마개) 로써 설계될 수도 있다.

[0120] 도 9a 내지 도 9d 는 무선 헤드셋의 또 다른 예인 멀티 마이크로폰 오디오 감지 디바이스 (D10) 의 휴대용 구현 (D200) 의 다양한 시점들을 도시한다. 디바이스 (D200) 는 둥근, 타원형 하우징 (Z12) 및 귀마개로써 구성될 수도 있는 이어폰 (Z22) 을 포함한다. 도 9a 내지 도 9d 는 디바이스 (D200) 의 어레이의 마이크로폰의, 1차 마이크로폰을 위한 음향 포트 (Z42) 의 위치 및 2차 마이크로폰을 위한 음향 포트 (Z52) 의 위치 또한 도시한다. 2차 마이크로폰 포트 (Z52) 는 적어도 부분적으로 가려질 수도 (예를 들어, 사용자 인터페이스 버튼에 의하여) 있다.

[0121] 도 10a 는 통신 핸드셋인, 멀티 마이크로폰 오디오 감지 디바이스 (D10) 의, 휴대용 구현 (D300) 의 횡단면도 (중심 축을 따른) 를 도시한다. 디바이스 (D300) 는, 1차 마이크로폰 (MC10) 및 2차 마이크로폰 (MC20) 을 갖는 어레이 (R100) 의 구현을 포함한다. 이 예에서, 디바이스 (D300) 는 1차 확성기 (SP10) 및 2차 확성기 (SP20) 또한 포함한다. 그러한 디바이스는 하나 이상의 인코딩 및 디코딩 방식 (또한 "코덱" 이라 불리는) 을 통해 무선으로 음성 통신 데이터를 송신 및 수신하도록 구성될 수도 있다. 그러한 코덱들의 예는, 2007년 2월의 "Enhanced Variable Rate Codec, Speech Service Option 3, 68, and 70 for Wideband Spread Spectrum Digital Systems" (온라인 [www-dot-3gpp-dot-org](http://www-dot-3gpp-dot-org) 에서 이용가능) 라는 제목의, 제 3 세대 파트너십 프로젝트 2 (3GPP2) 문서 C.S0014-C, v1.0 에서 설명되는, 향상된 가변 레이트 코덱을 포함하고; 2004년 1월의 "Selectable Mode Vocoder (SMV) Service Option for Wideband Spread Spectrum Communication Systems" (온라인 [www-dot-3gpp-dot-org](http://www-dot-3gpp-dot-org) 에서 이용가능) 라는 제목의, 3GPP2 문서 C.S0030-0, v3.0 에서 설명되는, 선택가능 모드 보코더 음성 (Selectable mode vocoder speech) 코덱을 포함하고; 문서 ETSI TS 126 092 V6.0.0 (2004년 12월, 유럽 전기 통신 표준 협회 (ETSI), 소피아 안티폴리스 세덱스 (Sophia Antipolis Cedex), FR) 에서 설명되는, 적응성 멀티 레이트 (AMR) 음성 코덱을 포함하고; 그리고 문서 ETSI TS 126 192 V6.0.0 (2004년 12월, ETSI) 에서 설명되는, AMR 광대역 음성 코덱을 포함한다.

[0122] 도 10a 의 예에서, 핸드셋 (D300) 은 폴더형 셀룰러 전화 핸드셋 ("플립" 핸드셋이라고도 불리는) 이다. 그러한 멀티 마이크로폰 통신 핸드셋의 다른 구성들은 막대 (bar) 형, 슬라이드형, 및 터치스크린 전화 핸드셋들을 포함하고, 디바이스 (D10) 는 임의의 이러한 형식들에 따라 구현될 수도 있다. 도 10b 는, 제 3 마이크로폰 (MC30) 을 포함하는 어레이 (R100) 의 3-마이크로폰 구현을 포함하는, 디바이스 (D300) 의 구현 (D310) 의 횡단면도를 도시한다.

[0123] 도 11a 는 미디어 플레이어인 멀티 마이크로폰 오디오 감지 디바이스 (D10) 의 휴대용 구현 (D400) 의 도면을 도시한다. 그러한 디바이스는, 표준 압축 포맷 (예를 들어, 동영상을 전문으로 연구하는 전문가 그룹 (Moving Picture Experts Group) (MPEG)-1 오디오 레이어 3 (MP3), MPEG-4 Part 14 (MP4), 윈도우즈 미디어 오디오/비디오 (WMA/WMV) 의 버전 (워싱턴, 레드몬드, 마이크로소프트 사), 고급 오디오 코딩 (AAC), 국제 전기통신 연합 (ITU) -T H.264, 또는 그와 유사한 것) 에 따라 인코딩된 파일 또는 스트림과 같은, 압축된 오디오 또는 시청각적 정보의 재생을 위해 구성될 수도 있다. 디바이스 (D400) 는, 디바이스의 전면면에 배치되는 디스플레이 스크린 (SC10) 및 확성기 (SP10) 를 포함하고, 어레이 (R100) 의 마이크로폰들 (MC10 및 MC20) 은 디바이스의 동일한 면 (예를 들어, 이 예에서와 같이 상측 면의 반대 측에, 또는 전면의 반대 측에) 에 배치된다. 도 11b 는, 마이크로폰들 (MC10 및 MC20) 이 디바이스의 반대 면들에 배치되는, 디바이스 (D400) 의 또 다른 구현 (D410) 을 도시하고, 도 11c 는, 마이크로폰들 (MC10 및 MC20) 이 디바이스의 근접한 면들에 배치되는, 디바이스 (D400) 의 추가적 구현 (D420) 을 도시한다. 또한 미디어 플레이어는 의도된 사용시에 더 긴 축이 가로축이 되도록 설계될 수도 있다.

[0124] 도 12a 는, 핸즈-프리 자동차 키트인 멀티 마이크로폰 오디오 감지 디바이스 (D10) 의 구현 (D500) 의 도면을 도시한다. 그러한 디바이스는, 계기판, 윈드실드 (windshield), 백미러, 차양 (visor), 또는 다른 자동차의 내부 표면에 설치되거나 탈착 가능하게 고정되도록 구성될 수도 있다. 예를 들어, 그러한 디바이스를 앞좌

석 탑승자의 앞 및 드라이버의 차양과 승객의 차양 사이에 (예를 들어, 백미러 내 또는 백미러 상에) 위치시키는 것이 바람직할 수도 있다. 디바이스 (D500) 는 확성기 (85) 및 어레이 (R100) 의 구현을 포함한다. 이 특정 예에서, 디바이스 (D500) 는 어레이 (R100) 의 4-마이크로폰 구현 (R102) 을 포함한다. 그러한 디바이스는, 상기에 열거된 예들과 같은, 하나 이상의 코텍들을 통해 음성 통신 데이터를 무선으로 송신 및 수신하도록 구성될 수도 있다. 다르게는 또는 추가적으로, 그러한 디바이스는 셀룰러 전화기 핸드셋과 같은 전화기 디바이스와의 통신을 통해 (예를 들어, 상기에 설명된 바와 같은 블루투스™ 프로토콜의 버전을 이용하여) 하프-듀플렉스 또는 풀-듀플렉스 전화를 지원하도록 구성될 수도 있다.

[0125] 도 12b 는, 스타일러스 (stylus) 또는 필기 디바이스 (예를 들어, 펜 또는 연필) 인 멀티 마이크로폰 오디오 감지 디바이스 (D10) 의 휴대용 구현 (D600) 의 도면을 도시한다. 디바이스 (D600) 는 어레이 (R100) 의 구현을 포함한다. 그러한 디바이스는, 상기에 열거된 예들과 같은, 하나 이상의 코텍들을 통해 무선으로 음성 통신 데이터를 송신 및 수신하도록 구성될 수도 있다. 다르게는 또는 추가적으로, 그러한 디바이스는 셀룰러 전화기 핸드셋 및/또는 무선 헤드셋과 같은 디바이스와의 통신을 통해 (예를 들어, 상기에 설명된 바와 같은 블루투스™ 프로토콜의 버전을 이용하여) 하프-듀플렉스 또는 풀-듀플렉스 전화를 지원하도록 구성될 수도 있다. 디바이스 (D600) 는, 어레이 (R100) 에 의해 생성되는 신호의, 드로잉 (drawing) 표면 (81) (예를 들어, 종이 한 장) 에 걸친 디바이스 (D600) 의 팁 (tip) 의 이동으로 인하여 초래될 수 있는, 스크래칭 (scratching) 잡음 (82) 의 레벨을 감소시키기 위하여 공간적으로 선택적 처리 동작을 수행하도록 구성되는 하나 이상의 프로세서들을 포함할 수도 있다.

[0126] 어레이 (R100) 의 비선형 4-마이크로폰 구현의 한 예는, 중앙 마이크로폰과 각각의 외곽 마이크로폰들 사이의 공간이 5 cm 인, 하나의 선상의 3 개의 마이크로폰들을 포함하고, 상기 선 4 cm 위쪽에 위치되고 외곽 마이크로폰 양쪽보다 중앙 마이크로폰에 더 가깝게 위치되는 또 다른 마이크로폰을 포함한다. 그러한 어레이를 위한 애플리케이션의 일 예시는 헬즈-프리 자동차 키트 (D500) 의 대체적 구현이다.

[0127] 휴대용 컴퓨팅 디바이스들의 클래스는, 랩탑 컴퓨터들, 노트북 컴퓨터들, 울트라-휴대용 컴퓨터들, 태블릿 컴퓨터들, 모바일 인터넷 디바이스들, 및 스마트폰들과 같은 이름들을 갖는 디바이스들을 현재 포함한다. 그러한 디바이스는, 디스플레이 스크린을 포함하는 상측 패널 및 키보드를 포함할 수도 있는 하측 패널을 가질 수도 있고, 상기 두 개의 패널들은 폴더 또는 힌지 (hinge) 관계로 연결될 수도 있다.

[0128] 도 13a 는 디바이스 (D10) 의 그러한 휴대용 컴퓨팅 구현 (D700) 의 예의 정면도를 도시한다. 디바이스 (D700) 는, 상측 패널 (PL10) 상의 선형 어레이에 정렬된, 디스플레이 스크린 (SC10) 위쪽의 4 개의 마이크로폰들 (MC10, MC20, MC30, MC40) 을 갖는 어레이 (R100) 의 구현을 포함한다. 도 13b 는 또 다른 차원의 4 개의 마이크로폰들의 위치들을 도시하는 상측 패널 (PL10) 의 평면도를 도시한다. 도 13c 는, 4 개의 마이크로폰들 (MC10, MC20, MC30, MC40) 이 디스플레이 스크린 (SC10) 위쪽에 비선형적 방법으로 상측 패널 (PL12) 상에 정렬된, 어레이 (R100) 의 구현을 포함하는 그러한 휴대용 컴퓨팅 디바이스 (D710) 의 또 다른 예의 정면도를 도시한다. 도 13d 는, 마이크로폰들 (MC10, MC20, 및 MC30) 이 패널의 전면에 배치되고 마이크로폰 (MC40) 이 패널의 후면에 배치된, 또 다른 차원에서의 4 개의 마이크로폰들의 위치들을 도시하는, 상측 패널 (PL12) 의 평면도를 도시한다.

[0129] 사용자가 사용중에, 그러한 디바이스 (D700 또는 D710) 의 앞에서 좌우로, 디바이스 쪽으로 또는 반대로, 및/또는 심지어 디바이스 주변으로 (예를 들어, 디바이스의 앞에서 디바이스의 뒤로) 움직일 수도 있다는 것이 예상될 수도 있다. 근접장 (near-field) 음성의 보존과 원거리장 (far-field) 간섭의 감쇄 사이에 적절한 트레이드 오프를 제공하거나/하고 원치 않는 방향들의 비선형적 신호 감쇄를 제공하기 위하여 그러한 디바이스 내에 디바이스 (D10) 를 구현하는 것이 바람직할 수도 있다. 최소 음성 왜곡을 위해 선형 마이크로폰 구성을 선택하거나, 또는 더 양호한 잡음 제거를 위해 비선형 마이크로폰 구성을 선택하는 것이 바람직할 수도 있다.

[0130] 어레이 (R100) 의 4-마이크로폰 인스턴스의 또 다른 예에서, 마이크로폰들은 대략 4 면의 구성으로 정렬되어, 하나의 마이크로폰이, 그 모서리들이, 대략 3 cm 간격의, 다른 3 개의 마이크로폰들의 위치에 의해 정의되는 삼각형 뒤에 (예를 들어, 대략 1 cm 뒤에) 위치되도록 한다. 그러한 어레이를 위한 잠재적 애플리케이션들은, 화자의 입과 어레이 사이의 예상 거리가 대략 20 내지 30 cm 인, 스피커폰 모드에서 동작하는 핸드셋을 포함한다. 도 14a 는, 4 개의 마이크로폰들 (MC10, MC20, MC30, MC40) 이 대략 삼면 구성으로 정렬된, 어레이 (R100) 의 그러한 구현을 포함하는 핸드셋 (D300) 의 구현 (D320) 의 정면도를 도시한다. 도 14b 는 핸드셋 내의 마이크로폰들 (MC10, MC20, MC30, 및 MC40) 의 위치를 도시하는 핸드셋 (D320) 의 측면도를 도시한다.

- [0131] 핸드셋 애플리케이션을 위한 어레이 (R100) 의 4-마이크로폰 인스턴스의 또 다른 예는 핸드셋의 전면 (예를 들어, 키패드의 1, 7, 및 9 위치 근처) 3 개의 마이크로폰들을 포함하고 후면에 (예를 들어, 키패드의 7 또는 9 위치 뒤에) 하나의 마이크로폰을 포함한다. 도 14c 는, 4 개의 마이크로폰들 (MC10, MC20, MC30, MC40) 이 "별" 구성으로 정렬된, 어레이 (R100) 의 그러한 구현을 포함하는 핸드셋 (D300) 의 구현 (D330) 의 정면도를 도시한다. 도 14d 는, 핸드셋 내의 마이크로폰들 (MC10, MC20, MC30, 및 MC40) 의 위치를 도시하는, 핸드셋 (D330) 의 측면도를 도시한다. 디바이스 (D10) 의 다른 예들은, 터치스크린의 주변에 마이크로폰들이 유사한 방법으로 정렬된, 핸드셋 (D320 및 D330) 의 터치스크린 구현들 (예를 들어, iPhone (캘리포니아, 쿠파티노, 애플 사), HD2 (중화민국, 대만, HTC), 또는 CLIQ (일리노이, 샤움버그, 모토롤라) 와 같은, 편평한, 접히지 않는 평판) 을 포함한다.
- [0132] 도 15 는, 핸드헬드 애플리케이션들을 위한 멀티 마이크로폰 오디오 감지 디바이스 (D10) 의 휴대용 구현 (D800) 의 도면을 도시한다. 디바이스 (D800) 는 터치스크린 디스플레이, 사용자 인터페이스 선택 제어 (좌측), 사용자 인터페이스 내비게이션 제어 (우측), 두 개의 확성기들, 및 3 개의 전면 마이크로폰들 및 일 후면 마이크로폰을 포함하는 어레이 (R100) 의 구현을 포함한다. 사용자 인터페이스 제어의 각각은, 하나 이상의 푸시버튼들 (pushbutton), 트랙볼들, 클릭-휠들 (click-wheel), 터치패드들, 조이스틱들, 및/또는 다른 포인팅 디바이스들, 등을 이용하여 구현될 수도 있다. 브라우저-토크 (browse-talk) 모드 또는 게임 플레이 모드에서 사용될 수도 있는, 디바이스 (D800) 의 일반 사이즈는 대략 15 cm x 20 cm 이다. 디바이스 (D10) 는, 상측 면의 여백 및/또는 태블릿 컴퓨터의 하나 이상의 측 표면 내에 배치된 어레이 (R100) 의 마이크로폰들을 구비한, 상측 면에 터치스크린 디스플레이를 포함하는 태블릿 컴퓨터 (예를 들어, iPad (애플 사), Slate (캘리포니아, 파울로 알토, 휴렛-팩커드 사, 또는 Streak (텍사스, 라운드 락, 텔 사) 와 같은, "슬레이트 (slate)") 로써 유사하게 구현될 수도 있다.
- [0133] 멀티채널 기록된 신호 내의 반향 에너지는 원하는 소스와 어레이 (R100) 사이의 길이가 증가함에 따라 증가하는 경향이 있다. 방법 (M100) 을 실행시키기에 바람직할 수도 있는 또 다른 애플리케이션은 오디오 및/또는 비디오 회의이다. 도 16a 내지 16d 는 디바이스 (D10) 의 회의 구현들의 몇몇 예들의 평면도를 도시한다. 도 16a 는 어레이 (R100) 의 3-마이크로폰 (마이크로폰들 (MC10, MC20, 및 MC30)) 구현을 포함한다. 도 16b 는 어레이 (R100) 의 4-마이크로폰 (마이크로폰들 (MC10, MC20, MC30, 및 MC40)) 구현을 포함한다. 도 16c 는 어레이 (R100) 의 5-마이크로폰 (마이크로폰들 (MC10, MC20, MC30, MC40, 및 MC50)) 구현을 포함한다. 도 16d 는 어레이 (R100) 의 6-마이크로폰 (마이크로폰들 (MC10, MC20, MC30, MC40, MC50 및 MC60)) 구현을 포함한다. 어레이 (R100) 의 각각의 마이크로폰들을 정다각형의 상응 꼭지점에 위치시키는 것이 바람직할 수도 있다. 원단 (far-end) 오디오 신호의 재생을 위한 확성기 (SP10) 는 디바이스 내에 (예를 들어, 도 16a 에 도시된 바와 같이) 포함될 수도 있거나/있고 그러한 확성기는 디바이스로부터 분리되어 (예를 들어, 음향 피드백을 감소시키기 위하여) 위치될 수도 있다.
- [0134] 디바이스 (D10) 의 회의 구현이, 각각의 마이크로폰 쌍에 대하여 또는, 적어도 각각의 활성 마이크로폰 쌍 (예를 들어, 하나 이상의 근접단 (near-end) 스피커의 각각의 음성을 분리하여 반향 제거하기 위해) 에 대해 방법 (M100) 의 구현의 분리된 인스턴스를 수행하도록 하는 것이 바람직할 수도 있다. 그러한 경우에, 원단으로의 송신 전에 디바이스가 다양한 반향 제거된 음성 신호들을 조합 (예를 들어, 믹싱) 하는 것 또한 바람직할 수도 있다.
- [0135] 디바이스 (D100) 의 회의 애플리케이션의 또 다른 예에서, 어레이 (R100) 의 수평 선형 구현이 텔레비전 또는 셋톱박스의 전면 패널 내에 포함된다. 그러한 디바이스는 어레이로부터 대략 1 내지 3 미터 또는 1 내지 4 미터 떨어진 위치의 앞 영역 및 상기 위치로부터의 영역 내에서 말하는 사람 (예를 들어, 텔레비전을 시청하는 시청자) 으로부터의 근접단 소스 신호를 로케이팅 (locating) 하고 반향 제거함으로써 전화기 통신을 지원하도록 구성될 수도 있다. 본원에 개시된 시스템들, 방법들, 및 장치의 적용가능성이 도 8a 내지 16d 에 도시된 특정 예들에 제한되지 않음이 명백히 개시된다.
- [0136] 멀티 마이크로폰 오디오 감지 디바이스 (예를 들어, 디바이스 (D100, D200, D300, D400, D500 또는 D600) 의 동작 가운데, 어레이 (R100) 는 음향 환경에 상응하는 하나 이상의 마이크로폰들의 응답에 기초한 각각의 채널의 멀티채널 신호를 생성한다. 하나의 마이크로폰은 다른 마이크로폰보다 더 직접적으로 특정 소리를 수신할 수도 있어, 단일 마이크로폰을 이용하여 캡처링 (capturing) 될 수 있는 것보다 음향 환경의 더 완벽한 표식을 총체적으로 제공하기 위하여 상응 채널들은 서로 상이하도록 한다.
- [0137] 어레이 (R100) 가 멀티채널 신호 (MCS) 를 생성하기 위하여 마이크로폰들에 의하여 생성되는 신호에 하나 이상



의 처리 동작들을 수행하도록 하는 것이 바람직할 수도 있다. 도 17a 는, 임피던스 매칭, 아날로그-디지털 변환, 이득 제어, 및/또는 아날로그 및/또는 디지털 도메인들에서의 필터링을 포함 (제한이 아니라) 할 수도 있는, 하나 이상의 그러한 동작들을 수행하도록 구성되는 오디오 전처리 스테이지 (AP10) 를 포함하는 어레이 (R100) 의 구현 (R200) 의 블록도를 도시한다.

[0138] 도 17b 는 어레이 (R200) 의 구현 (R210) 의 블록도를 도시한다. 어레이 (R210) 는 아날로그 전처리 스테이지들 (P10a 및 P10b) 을 포함하는 오디오 전처리 스테이지 (AP10) 의 구현 (AP20) 을 포함한다. 일 예에서, 스테이지들 (P10a 및 P10b) 은 상용 마이크로폰 신호에 고역 통과 필터링 동작 (예를 들어, 50, 100, 또는 200 Hz 의 차단주파수로) 을 수행하도록 각각 구성된다.

[0139] 어레이 (R100) 가 멀티채널 신호를 디지털 신호로써, 다시 말해서, 샘플들의 시퀀스으로써 생성하도록 하는 것이 바람직할 수도 있다. 어레이 (R210) 는, 예를 들어, 상용 아날로그 채널을 샘플링하도록 각각 배열되는 아날로그-디지털 변환기들 (ADC) (C10a, 및 C10b) 을 포함한다. 비록 44 kHz 만큼 높은 샘플링 레이트들 또한 사용될 수도 있지만, 음향 애플리케이션들을 위한 보통의 샘플링 레이트들은 8 kHz, 12 kHz, 16 kHz, 및 대략 8 내지 16 kHz 범위의 주파수들을 포함한다. 이 특정 예에서, 어레이 (R210) 는, 멀티채널 신호 (MCS) 의 상용 채널들 (MCS-1, MCS-2) 을 생성하기 위하여 상용 디지털화된 채널에 하나 이상의 처리 동작들 (예를 들어, 반향 소거 (echo cancellation), 잡음 감소, 및/또는 스펙트럼 성형 (spectral shaping)) 을 수행하도록 각각 구성될 수도 있는, 디지털 처리 스테이지들 (P20a 및 P20b) 또한 포함한다. 비록 도 17a 및 17b 는 2-채널 구현들을 도시하나, 동일한 원칙들이 불특정 수의 마이크로폰들과 멀티채널 신호 (MCS) 의 상용 채널들에 연장될 수도 있음이 이해될 것이다.

[0140] 본원에 개시된 방법들 및 장치들은 일반적으로 임의의 트랜시빙 (transceiving) 및/또는 오디오 감지 애플리케이션, 그러한 애플리케이션들의, 특히 모바일 또는 그렇지 않으면 휴대용, 인스턴스들에 적용될 수도 있다. 예를 들어, 본원에 개시된 구성들의 범위는 코드-분할 다중-접속 (CDMA) 오버-더-에어 (over-the-air) 인터페이스를 이용하도록 구성되는 무선 전화 통신 시스템 내에 상주하는 통신 디바이스들을 포함한다. 그럼에도, 당업자에 있어, 본원에 설명된 대로의 기능들을 갖는 방법 및 장치는, 유선 및/또는 무선 (예를 들어, CDMA, TDMA, FDMA 및/또는 TS-SCDMA) 전송 채널들을 통한 음성 패킷망 (VoIP) 을 이용하는 시스템들과 같은, 당업자에 게 알려진 기술들의 넓은 범위를 이용하는 임의의 다양한 통신 시스템들에 상주할 수도 있다.

[0141] 본원에 개시된 통신 디바이스들이, 패킷-교환 (packet-switched) (예를 들어, VoIP 와 같은 프로토콜에 따라 오디오 송신을 수행하도록 배열된 유선 및/또는 무선 네트워크들) 및/또는 회선-교환 (circuit-switched) 인 네트워크들에서의 사용을 위해 적용될 수도 있음이 명백히 관찰되고 여기에 개시된다. 본원에 개시된 통신 디바이스들이, 협대역 코딩 시스템들 (예를 들어, 대략 4 또는 5 kHz 의 범위의 오디오 주파수를 인코딩하는 시스템들) 및/또는 전체 대역 광대역 코딩 시스템들 및 분할 대역 광대역 코딩 시스템들을 포함하는 광대역 코딩 시스템들 (예를 들어, 5 kHz 이상의 오디오 주파수들을 인코딩하는 시스템들) 에서의 사용을 위해 적용될 수도 있음 또한 명백히 관찰되고 여기에 개시된다.

[0142] 설명된 구성들의 기술한 표현은 임의의 당업자가 본원에 개시된 방법들 및 다른 구조들을 만들거나 사용하는 것을 가능하도록 하기 위하여 제공된다. 본원에 도시되고 설명된 플로우차트들, 블록도들, 및 다른 구조들은 오로지 예들이며, 이러한 구조들의 다른 변형들 또한 본 개시물의 범위 내에 있다. 이 구성들에 다양한 수정이 가능하며, 본원에 나타난 일반적 원칙들은 다른 구성들에도 적용될 수도 있다. 따라서, 본 개시물은 상기에 도시된 구성들에 제한되도록 의도된 것이 아니라, 최초 개시물의 일부를 형성하는 출원된 대로의 첨부된 청구항들을 포함하여, 본원에 임의의 방법으로 개시된 원칙들 및 신규한 특성들에 일치하는 가장 넓은 범위를 따르도록 의도된 것이다.

[0143] 당업자는 정보 및 신호들이 임의의 다수의 상이한 기술들 및 테크닉을 이용하여 표현될 수 있음을 인식할 것이다. 예를 들어, 상기 설명을 통해 참조될 수 있는 데이터, 지시들, 명령들, 정보, 신호들, 비트들, 심볼들 및 칩들은 전압들, 전류들, 전자기파들, 자기 필드들 또는 입자들, 광학 필드들 또는 입자들, 또는 그들의 임의의 조합에 의해 표현될 수 있다.

[0144] 본원에 개시된 구성의 구현을 위한 중요 설계 요구사항들은, 8 kHz 보다 더 높은 (예를 들어, 12, 16, 또는 44 kHz) 샘플링 레이트들의 음성 통신을 위한 애플리케이션들과 같은, 계산 집중 애플리케이션들에 있어서 특히, 처리 지연 및/또는 계산 복잡도 (보통 초당 백만 단위 명령어들 또는 MIPS 로 측정되는) 를 최소화하는 것을 포함할 수도 있다.

[0145] 본원에 개시된 장치 (예를 들어, A100, A104, A106, A108, MF100, A200) 의 구현의 다양한 엘리먼트들은, 의도된 애플리케이션에 적합하다 여겨지는 하드웨어, 소프트웨어, 및/또는 펌웨어의 임의의 조합 내에 구현될 수도 있다. 예를 들어, 그러한 엘리먼트는, 예를 들어, 동일한 칩 상에 또는 칩셋의 2 이상의 칩들 가운데 상주하는, 전자 및/또는 광학 디바이스들로써 제조될 수도 있다. 그러한 디바이스의 한 예는, 트랜지스터들 또는 로직 게이트들과 같은, 로직 엘리먼트들의 고정 어레이 또는 프로그래머블 어레이이고, 임의의 이 엘리먼트들은 하나 이상의 그러한 어레이들로 구현될 수도 있다. 임의의 2 이상의, 또는 모든 이 엘리먼트들은 동일 어레이 또는 어레이들 내에 구현될 수도 있다. 그러한 어레이 또는 어레이들은 하나 이상의 칩들 내에 (예를 들어, 2 이상의 칩들을 포함하는 칩들 내에) 구현될 수도 있다.

[0146] 본원에 개시된 장치 (예를 들어, 장치 (A100, A104, A106, A108, MF100, A200)) 의 다양한 구성들의 하나 이상의 엘리먼트들은, 마이크로프로세서들, 내장형 프로세서들, IP 코어들, 디지털 신호 프로세서들, FPGA (필드 프로그래머블 게이트 어레이), ASSP, 및 ASIC (주문형 집적 회로) 과 같은 로직 엘리먼트들의, 하나 이상의 고정 어레이들 또는 프로그래머블 어레이들에서 실행되도록 배열된 명령어들의 하나 이상의 세트들의 전체 또는 부분으로써 구현될 수도 있다. 본원에 개시된 장치의 구현의 임의의 다양한 엘리먼트들은, 하나 이상의 컴퓨터들 (예를 들어, "프로세서" 라 불리기도 하는, 명령어들의 하나 이상의 세트들 또는 시퀀스들을 실행하도록 프로그래밍된 하나 이상의 어레이들을 포함하는 기계) 로써 구현될 수도 있고, 이 엘리먼트들의 2 이상, 또는 전부는 그러한 동일 컴퓨터 또는 컴퓨터들 내에 구현될 수도 있다.

[0147] 본원에 개시된 처리를 위한 프로세서 또는 다른 수단들은, 예를 들어, 동일 칩 또는 칩셋의 2 이상의 칩들에 상주하는 하나 이상의 전자 및/또는 광학 디바이스들로써 제조될 수도 있다. 그러한 디바이스의 한 예는, 트랜지스터들 또는 로직 게이트들과 같은 로직 엘리먼트들의, 고정 또는 프로그래머블 어레이이고, 이 임의의 엘리먼트들은 하나 이상의 그러한 어레이들로 구현될 수도 있다. 그러한 어레이 또는 어레이들은 하나 이상의 칩들 (예를 들어, 2 이상의 칩들을 포함하는 칩셋 내에) 내에 구현될 수도 있다. 그러한 어레이들의 예들은, 마이크로프로세서들, 내장형 프로세서들, IP 코어들, 디지털 신호 프로세서들, FPGA, ASSP, 및 ASIC 과 같은 로직 엘리먼트들의, 고정 또는 프로그래머블 어레이들을 포함한다. 본원에 개시된 처리를 위한 프로세서 또는 다른 수단들은, 컴퓨터들 (예를 들어, 명령어들의 하나 이상의 세트들 또는 시퀀스들을 실행하도록 프로그래밍된 하나 이상의 어레이들을 포함하는 기계) 또는 다른 프로세서들으로써 구현될 수도 있다. 본원에 설명된 프로세서가, 프로세서가 내장된 디바이스 또는 시스템 (예를 들어, 오디오 감지 디바이스) 의 다른 동작에 관련된 태스크와 같은, 코히런트 검출 과정과 직접적으로 연관되지 않은, 태스크들을 수행하거나 명령어들의 다른 세트들을 실행하도록 이용되는 것이 가능하다. 또한 본원에 개시된 방법의 일부가 오디오 감지 디바이스의 프로세서에 의해 수행되고 방법의 다른 일부가 하나 이상의 다른 프로세서들의 제어 하에 수행되는 것이 가능하다.

[0148] 당업자는 본원에 개시된 구성들과 관련하여 설명된 다양한 예시적 모듈들, 로직 블록들, 회로들, 및 테스트들과 동작들이, 전자 하드웨어, 컴퓨터 소프트웨어, 또는 양자의 조합들으로써 구현될 수도 있음을 인식할 수 있다. 그러한 모듈들, 로직 블록들, 회로들, 및 동작들은, 범용 프로세서, 디지털 신호 프로세서, ASIC 또는 ASSP, FPGA 또는 다른 프로그래머블 로직 디바이스, 이산 게이트 또는 트랜지스터 로직, 이산 하드웨어 컴포넌트들, 또는 본원에 개시된 구성들을 생성하도록 설계된 이들의 임의의 조합으로, 구현되거나 수행될 수도 있다. 예를 들어, 그러한 구성은, 하드와이어드 회로, ASCII 내에 제조된 회로 구성, 또는 비휘발성 저장장치에 로딩된 펌웨어 프로그램이나, 범용 프로세서 또는 다른 디지털 신호 처리 유닛과 같은 로직 엘리먼트들의 어레이에 의해 실행가능한 명령어들인, 머신-판독가능 코드로써 데이터 저장 매체에 로딩 되거나 데이터 저장 매체로부터 로딩된 소프트웨어 프로그램의 적어도 일부로써 구현될 수도 있다. 범용 프로세서는 마이크로프로세서일 수도 있지만, 다른 방법으로, 그 프로세서는 임의의 종래 프로세서, 제어기, 마이크로 제어기, 또는 상태 머신일 수도 있다. 또한, 프로세서는 컴퓨팅 디바이스들의 결합, 예를 들어, DSP 와 마이크로프로세서의 결합, 복수의 마이크로프로세서들, DSP 코어와 결합된 하나 이상의 마이크로프로세서들 또는 임의의 기타 다른 구성물로 구현될 수도 있다. 소프트웨어 모듈은 RAM 메모리, 플래쉬 메모리, ROM 메모리, PROM 메모리, EPROM 메모리, EEPROM 메모리, 레지스터, 하드 디스크, 착탈형 디스크, CD-ROM, 또는 당업계에 알려진 임의의 다른 형태의 저장 매체에 상주할 수도 있다. 예시적 저장 매체는 프로세서에 커플링되어, 프로세서는 저장 매체로부터 정보를 읽을 수 있고, 저장매체에 정보를 기록할 수 있다. 다른 방법으로, 저장 매체는 프로세서와 일체형일 수도 있다. 프로세서 및 저장 매체는 ASIC 내에 상주할 수도 있다. ASIC 은 사용자 단말에 상주할 수도 있다. 다른 방법으로는, 프로세서 및 저장 매체는 사용자 단말에 이산 컴포넌트들으로써 상주할 수도 있다.

[0149] 본원에 개시된 다양한 방법들 (예를 들어, 방법 (M100, M102)) 은 프로세서와 같은 로직 엘리먼트들의 어레이에 의하여 수행될 수도 있고, 본원에 개시된 장치의 다양한 엘리먼트들은 그러한 어레이에서 실행되도록 설계된 모듈들로 구현될 수도 있음을 알 수 있다. 본원에 사용된, 용어 "모듈" 또는 "서브-모듈" 은 임의의 방법, 장치, 디바이스, 유닛, 또는 소프트웨어, 하드웨어, 또는 펌웨어 형의 컴퓨터 명령어들 (예를 들어, 논리적 표현들) 을 포함하는 컴퓨터-판독가능 데이터 저장 매체를 지칭할 수 있다. 다중 모듈들 또는 다중 시스템들이 하나의 모듈 또는 시스템으로 조합될 수 있고 하나의 모듈 또는 하나의 시스템은 동일한 기능들을 수행하기 위하여 다중 모듈들 또는 다중 시스템들로 분리될 수 있다. 소프트웨어 또는 다른 컴퓨터-실행가능한 명령어 들로 구현시, 프로세스의 엘리먼트들은, 루틴들, 프로그램들, 오브젝트들, 컴포넌트들, 데이터 구조들, 및 그와 같은, 관련된 태스크들을 수행하기 위해 필수적인 코드 세그먼트들이다. 용어 "소프트웨어" 는, 소스 코드, 어셈블리 언어 코드, 기계어, 이진 코드, 펌웨어, 매크로 코드, 마이크로 코드, 로직 엘리먼트들의 어레이에 의해 실행가능한 명령어들의 임의의 하나 이상의 세트들 또는 시퀀스들, 및 그러한 예들의 임의의 조합을 포함하는 것으로 이해되어야 한다. 프로그램 또는 코드 세그먼트들은 프로세서 판독가능 매체에 저장되거나 반송 파에 실린 컴퓨터 데이터 신호에 의하여 송신 매체 또는 통신 링크를 통해 송신될 수 있다.

[0150] 또한, 본원에 개시된 방법들, 방식들, 및 기술들의 구현들은, 로직 엘리먼트들의 어레이를 포함하는 기계 (예를 들어, 프로세서, 마이크로프로세서, 마이크로컨트롤러, 또는 다른 유한 상태 기계) 에 의해 판독가능하거나/하고 실행가능한 명령어들의 하나 이상의 세트들로 명백히 구현 (예를 들어, 본원에 열거된 바와 같은 하나 이상의 컴퓨터-판독가능 매체 내에) 될 수도 있다. 용어 "컴퓨터 판독가능 매체" 는 휘발성, 비휘발성, 탈착가능, 탈착 불가능 매체를 포함하는, 정보를 저장 또는 전송할 수 있는 임의의 매체를 포함할 수도 있다. 컴퓨터 판독가능 매체의 예들은 전기 회로, 컴퓨터 판독가능 저장 매체 (예를 들어, ROM, EROM, 플래시 메모리, 또는 다른 반도체 저장 디바이스; 플로피 디스켓, 하드 디스크 또는 다른 자기 저장장치; CD-ROM/DVD 또는 다른 광 디스크 저장장치), 송신 매체 (예를 들어, 광섬유 매체, 무선 주파수 (RF) 링크), 또는 원하는 정보를 얻기 위해 액세스될 수 있는 임의의 다른 매체를 포함한다. 컴퓨터 데이터 신호는, 전자 네트워크 채널들, 광섬유, 공기, 전자기, RF 링크들, 등과 같은, 전송 매체 너머로 전파될 수 있는 임의의 신호를 포함할 수도 있다. 코드 세그먼트들은 인터넷 또는 인트라넷과 같은 컴퓨터 네트워크들을 통해 다운로드될 수도 있다. 임의의 경우에, 본 개시물의 범위는 그러한 구현들에 의하여 제한되지 않는 것으로 이해되어야 한다.

[0151] 본원에 설명된 방법들의 태스크들 각각은 하드웨어, 프로세서에 의해 실행가능한 소프트웨어 모듈, 또는 그들의 조합으로 직접 구현될 수도 있다. 본원에 개시된 방법의 구현의 일반적 애플리케이션에 있어, 로직 엘리먼트들의 어레이 (예를 들어, 로직 게이트들) 는 방법의 다양한 태스크들의 하나, 하나 이상, 또는 모두를 수행하도록 구성된다. 하나 이상 (가능하면 전부) 의 태스크들은 또한, 로직 엘리먼트들의 어레이 (예를 들어, 프로세서, 마이크로프로세서, 마이크로컨트롤러, 또는 다른 유한 상태 머신) 를 포함하는 머신 (예를 들어, 컴퓨터) 에 의하여 판독가능하거나/하고 실행가능한, 컴퓨터 프로그램 제품 (예를 들어, 디스크들, 플래시 또는 다른 비휘발성 메모리 카드들, 반도체 메모리 칩들, 등과 같은 하나 이상의 데이터 저장 매체) 에 구현된, 코드 (예를 들어, 명령어들의 하나 이상의 세트들) 로써 구현될 수도 있다. 본원에 개시된 방법의 구현의 태스크들은 또한 하나 이상의 그러한 어레이 또는 머신에 의해 수행될 수도 있다. 이러한 구현들 또는 다른 구현들에서, 태스크들은 셀룰러 전화기 또는 그러한 통신 가능성을 갖는 다른 디바이스와 같은 무선 통신을 위한 디바이스 내에서 수행될 수도 있다. 그러한 디바이스는 회로 스위칭 및/또는 패킷 스위칭 네트워크들과 (예를 들어, VoIP 와 같은 하나 이상의 프로토콜들을 이용하여) 통신하도록 구성될 수도 있다. 예를 들어, 그러한 디바이스는 인코딩된 프레임들을 송신 및/또는 수신하도록 구성된 RF 회로망을 포함할 수도 있다.

[0152] 본원에 개시된 다양한 방법들이, 핸드셋, 헤드셋, 또는 개인 정보 단말 (PDA) 과 같은 휴대용 통신 디바이스에 의하여 수행될 수도 있고, 본원에 설명된 다양한 장치들이 그러한 디바이스 내에 포함될 수도 있음이 명백히 개시된다. 보통의 실시간 (예를 들어, 온라인) 애플리케이션은 그러한 모바일 디바이스를 이용하여 수행되는 전화 대화이다.

[0153] 하나 이상의 예시적인 실시형태에서, 설명된 동작들은 하드웨어, 소프트웨어, 펌웨어, 또는 이들의 임의의 조합으로 구현될 수도 있다. 소프트웨어로 구현되는 경우, 그러한 동작들은 하나 이상의 명령들 또는 컴퓨터 판독가능 매체 상의 코드 상에 저장되거나 이를 통해 송신될 수도 있다. 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 이용 가능한 매체일 수도 있다. 용어 "컴퓨터 판독가능 매체" 는 컴퓨터 판독가능 저장 매체 및 통신 (예를 들어, 송신) 매체 양자를 포함한다. 비 제한적인 예시의 방식으로, 컴퓨터 판독가능 저장 매체는 반도체 메모리 (이에 제한되는 것이 아니라, 동적 또는 정적 RAM, ROM, EEPROM, 및/또는 플래시 RAM 을 포함할 수도 있는), 또는 강유전성, 자기 저항성, 오보닉 (ovonic), 중합, 또는 위상 변이 메모

리; CD-ROM 또는 다른 광 디스크 저장장치; 및/또는 자기 디스크 저장장치 또는 다른 자기 저장 디바이스들과 같은 저장 엘리먼트들의 어레이를 포함할 수 있다. 그러한 저장 매체는 컴퓨터에 의해 액세스 될 수 있는 명령어들 또는 데이터 구조들의 형태로 정보를 저장할 수도 있다. 통신 매체는, 한 곳으로부터 다른 곳으로 컴퓨터 프로그램의 전송을 용이하게 하는 임의의 매체를 포함하여, 명령어들 또는 데이터 구조들의 형태로 원하는 프로그램 코드를 운반하는데 사용될 수 있고, 컴퓨터에 의하여 액세스될 수 있는 임의의 매체를 포함한다.

또한, 임의의 연결은 컴퓨터 판독가능 매체로 적절히 불릴 수 있다. 예를 들어, 소프트웨어가 동축 케이블, 광섬유 케이블, 연선 (twisted pair), DSL (digital subscriber line), 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술을 이용하여 웹사이트, 서버, 또는 다른 원격 소스로부터 송신되면, 동축 케이블, 광섬유 케이블, 연선, DSL, 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술은 매체의 정의 내에 포함된다. 본 명세서에서 이용된 바와 같이, 디스크 (disk) 및 디스크 (disc) 는 콤팩트 디스크 (CD), 레이저 디스크, 광학 디스크, DVD (digital versatile disc), 플로피 디스크 및 블루-레이 디스크를 포함하는데, 여기서 디스크 (disk) 는 보통 자기적으로 데이터를 재생하는 반면, 디스크 (disc) 는 레이저를 이용하여 광학적으로 데이터를 재생한다. 상기의 조합은 또한 컴퓨터 판독가능 매체의 범위 내에 포함되어야 한다.

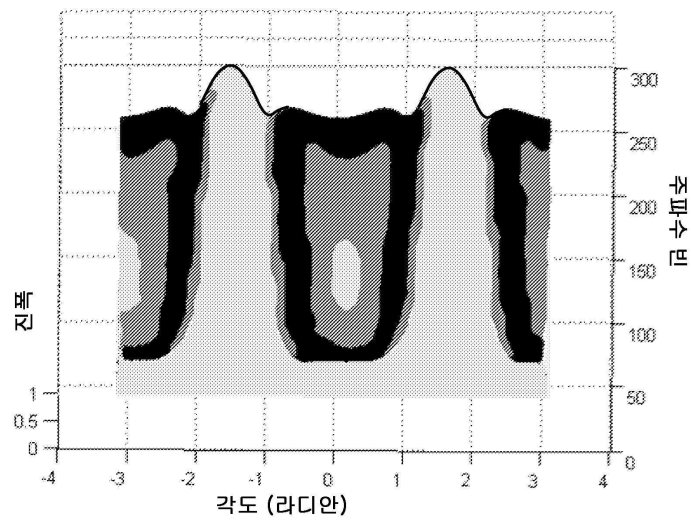
[0154] 본원에 설명된 음향 신호 처리 장치는, 특정 동작들을 제어하기 위하여 음성 입력을 수용하거나, 배경 잡음으로부터 원하는 잡음의 분리를 통해 이익을 얻을 수도 있는, 통신 디바이스들과 같은, 전자 디바이스에 병합될 수도 있다. 많은 애플리케이션들이 다중 방향들로부터 기원하는 배경 소리로부터 깨끗한 원하는 소리의 강화 또는 분리로 이익을 얻을 수도 있다. 그러한 애플리케이션들은, 음성 인식 및 검출, 음성 강화 및 분리, 음성-가동 제어, 및 유사한 것들과 같은 가능성들을 병합하는 전자 또는 컴퓨팅 디바이스들의 인간-기계 인터페이스들을 포함할 수도 있다. 그러한 음향 신호 처리 장치를 제한된 처리 가능성만을 제공하는 디바이스에 적합하도록 구현하는 것이 바람직할 수도 있다.

[0155] 본원에 설명된 모듈들, 엘리먼트들, 및 디바이스들의 다양한 구현들의 엘리먼트들은, 예를 들어, 동일 칩 상에 또는 칩셋의 2 이상의 칩들 가운데 상주하는 전자 및/또는 광학 디바이스들일 수도 있다. 그러한 디바이스의 한 예는, 트랜지스터 또는 게이트들과 같은 로직 엘리먼트들의, 고정 또는 프로그래머블 어레이이다. 본원에 설명된 장치의 다양한 구현들의 하나 이상의 엘리먼트들은 또한, 마이크로프로세서들, 내장형 프로세서들, IP 코어들, 디지털 신호 프로세서들, FPGA들, ASSP들, 및 ASIC들과 같은 로직 엘리먼트들의, 하나 이상의 고정 어레이들 또는 프로그래머블 어레이들에서 실행되도록 배열된 명령어들의 하나 이상의 세트들의 전체 또는 부분으로써 구현될 수도 있다.

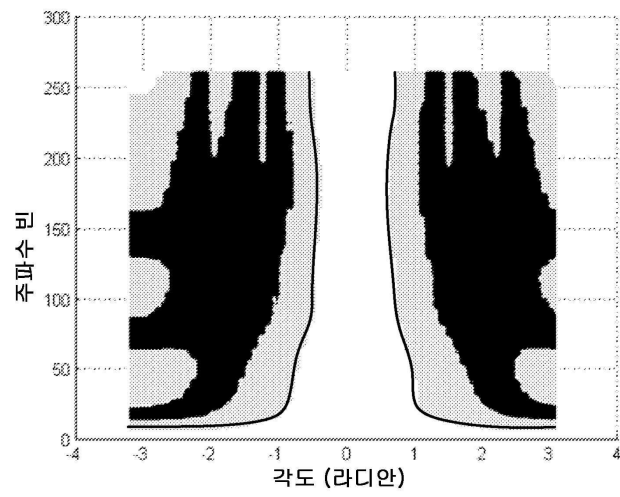
[0156] 본원에 설명된 장치의 구현의 하나 이상의 엘리먼트들이, 장치가 내장된 디바이스 또는 시스템의 다른 동작과 관련된 태스크와 같은, 장치의 동작과 직접적으로 관련되지 않은, 태스크들을 수행하거나 명령어들의 다른 세트들을 실행하는 것이 가능하다. 그러한 장치의 구현의 하나 이상의 엘리먼트들이 공통된 구조 (예를 들어, 상이한 시간에서의 상이한 엘리먼트들, 상이한 시간에서의 상이한 엘리먼트들에 상응하는 태스크들을 수행하도록 실행되는 명령어들의 세트, 또는 상이한 시간에 상이한 엘리먼트들을 위한 동작들을 수행하는 전자 및/또는 광학 디바이스들의 배치에 상응하는 코드의 부분을 실행하도록 이용되는 프로세서) 를 갖도록 하는 것 또한 가능하다.

도면

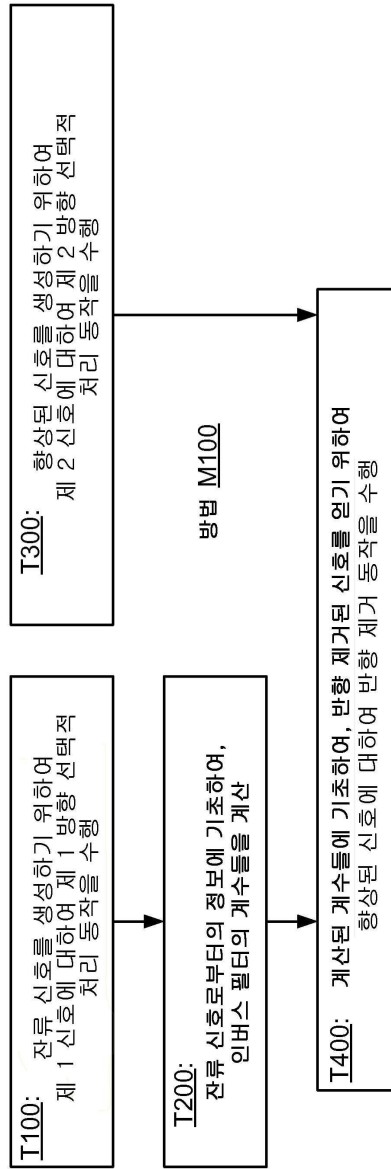
도면1a



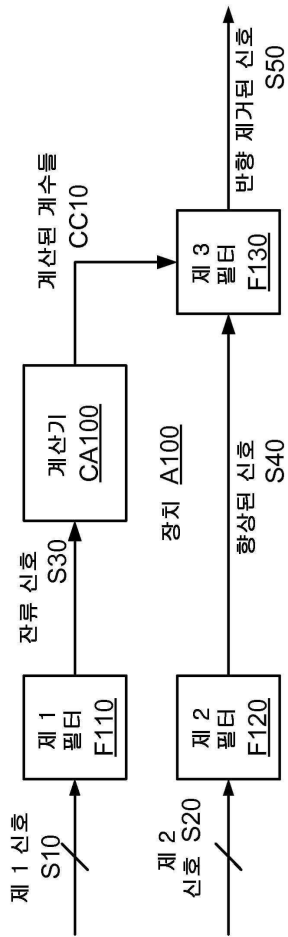
도면1b



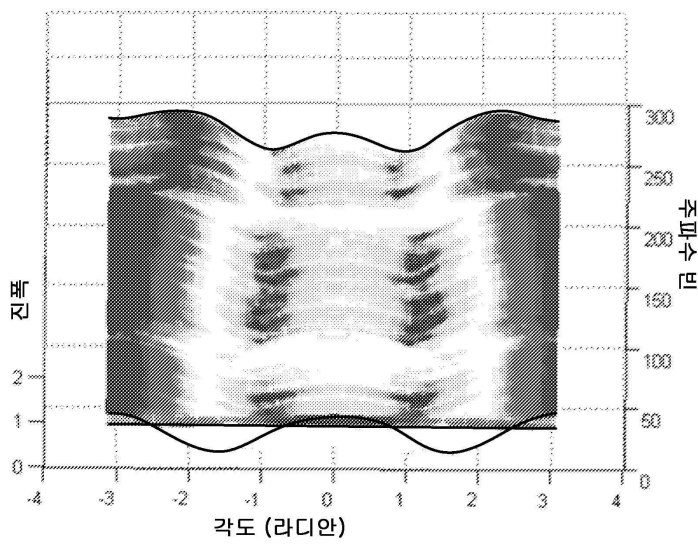
도면2a



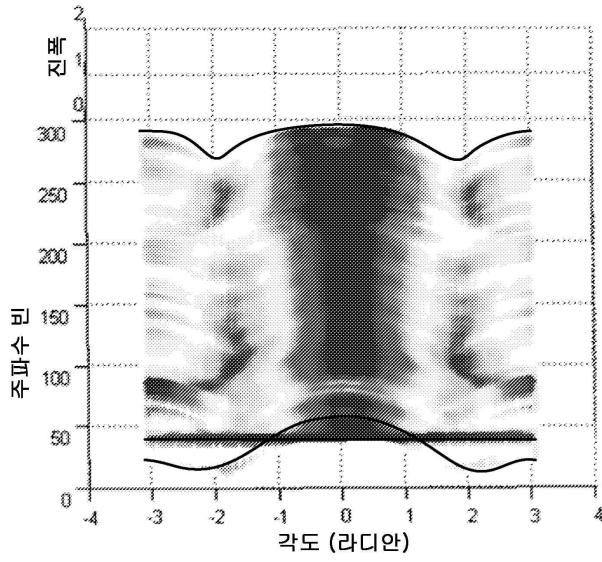
도면2b



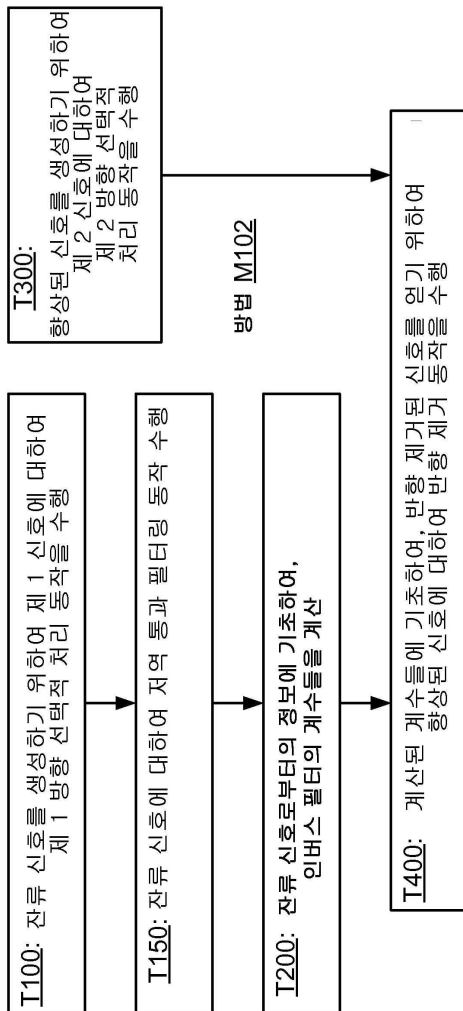
도면3a



도면3b

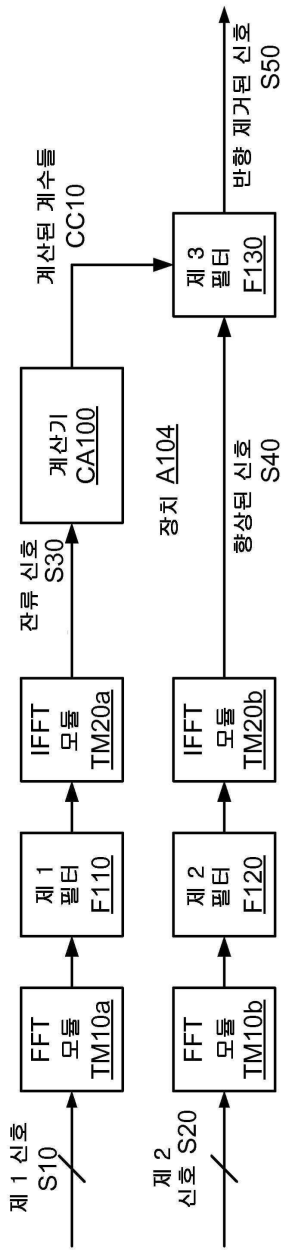


도면4a

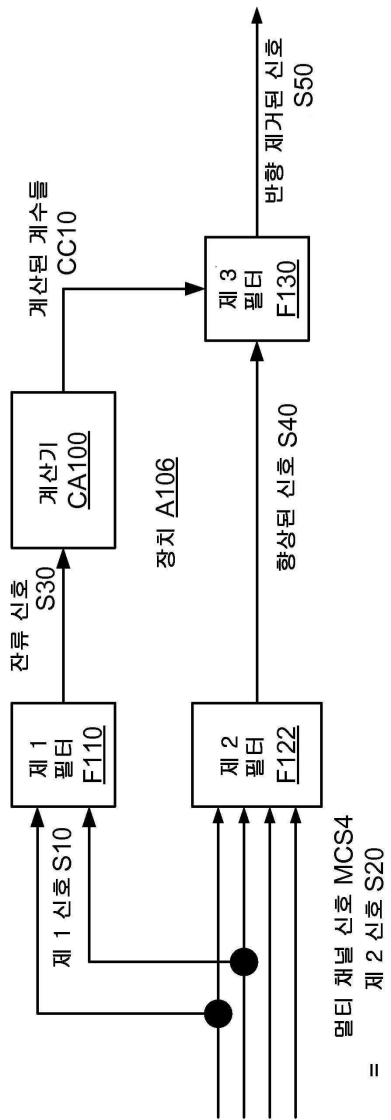




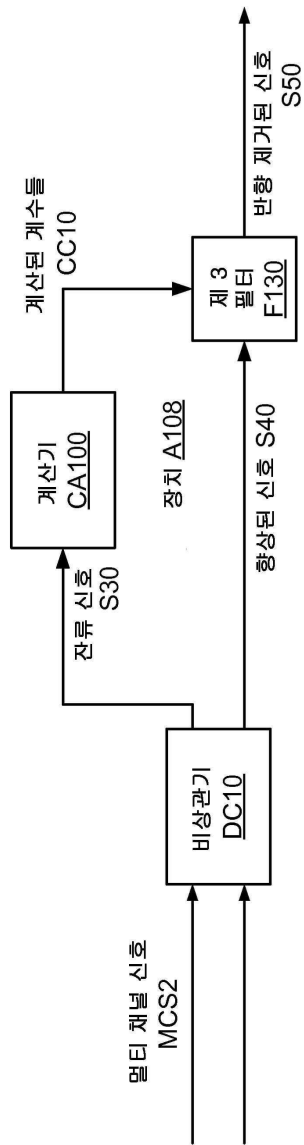
도면4b



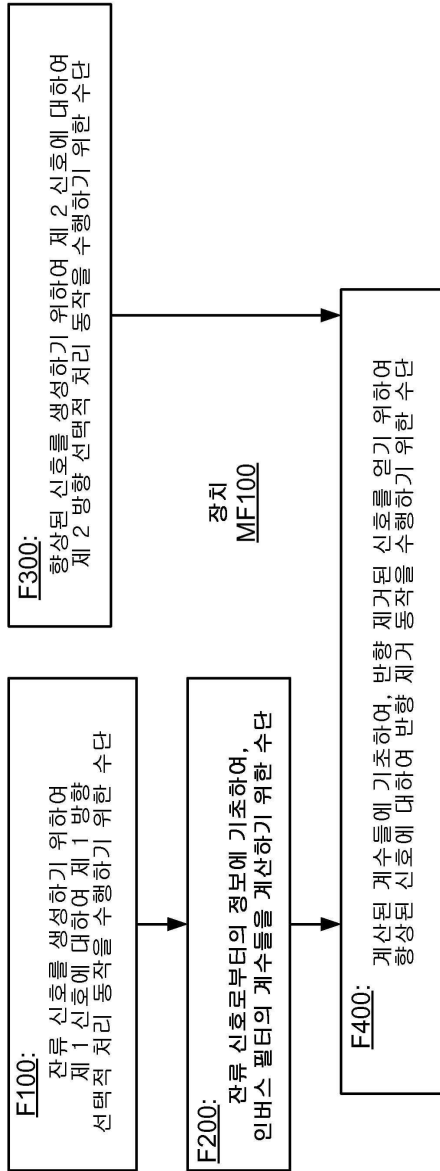
도면5a



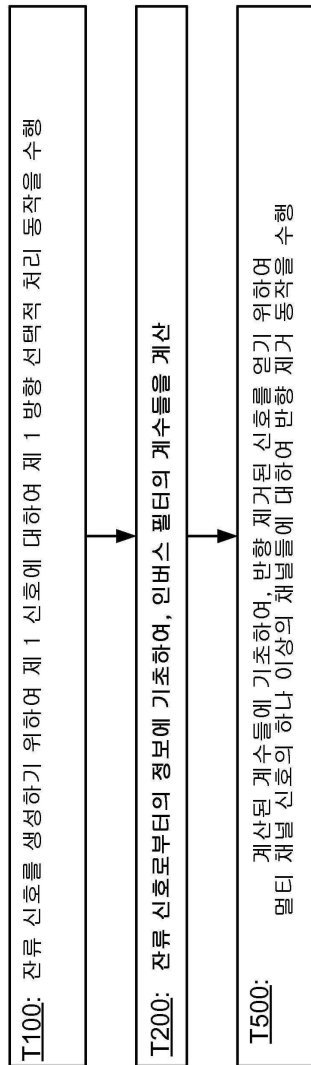
도면5b



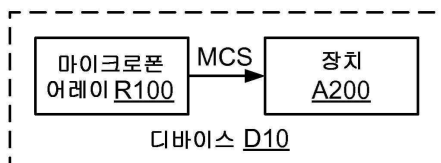
도면6a



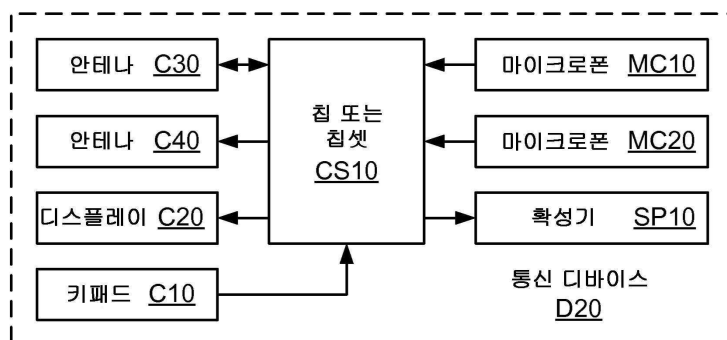
도면6b



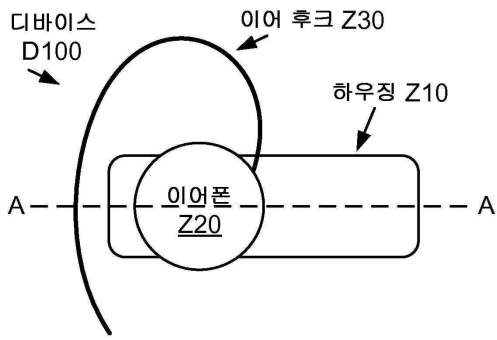
도면7a



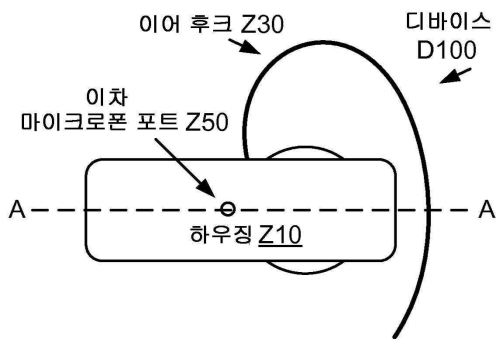
도면7b



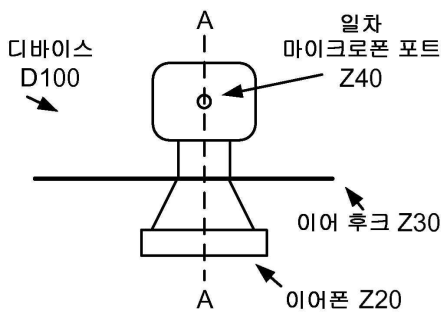
도면8a



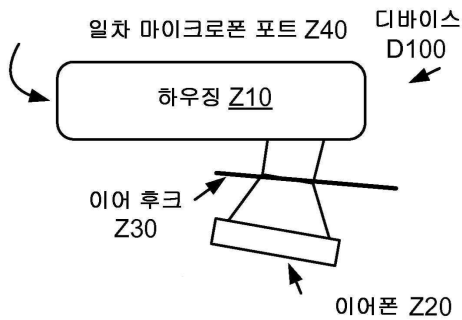
도면8b



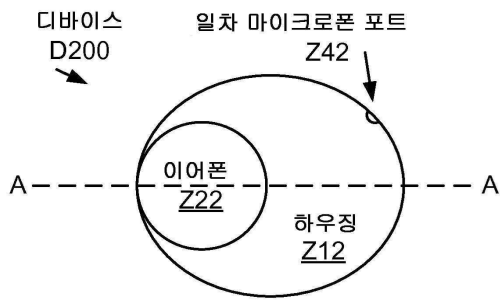
도면8c



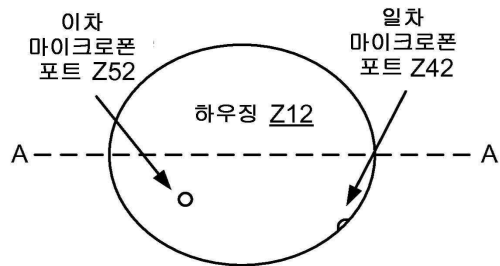
도면8d



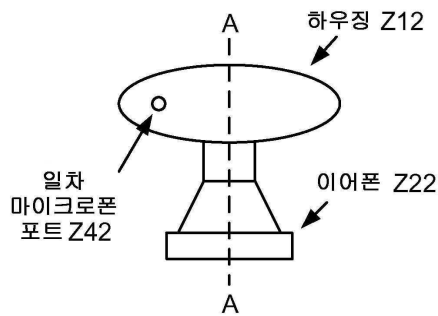
도면9a



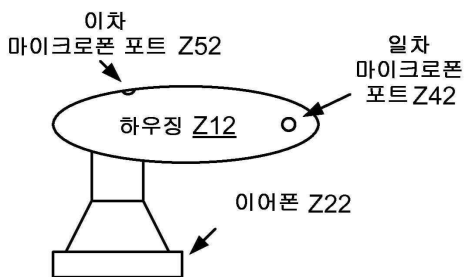
도면9b



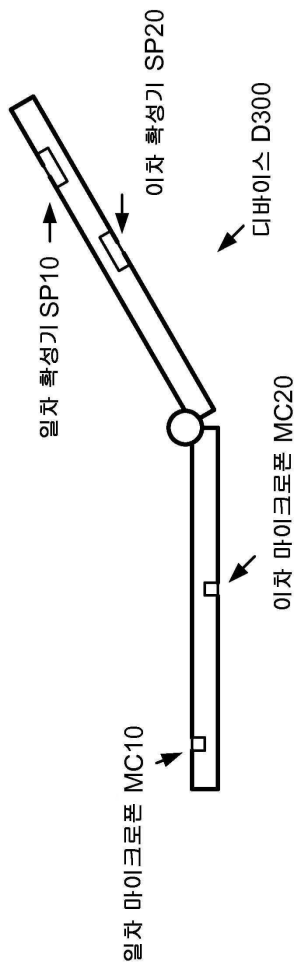
도면9c



도면9d

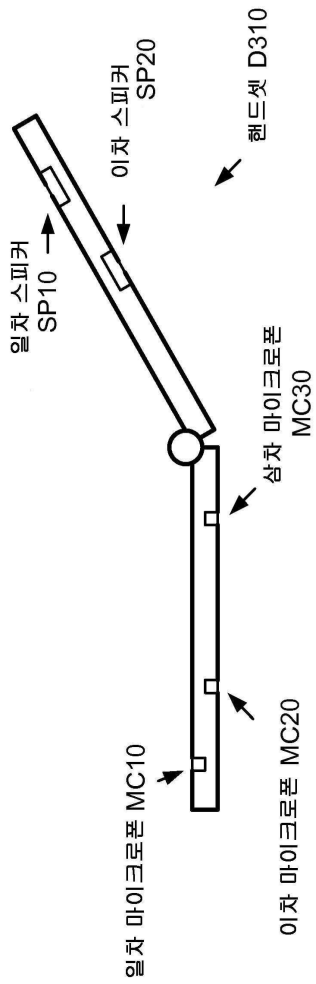


도면10a

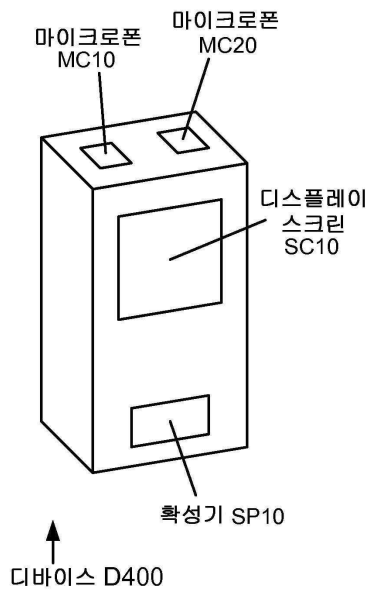




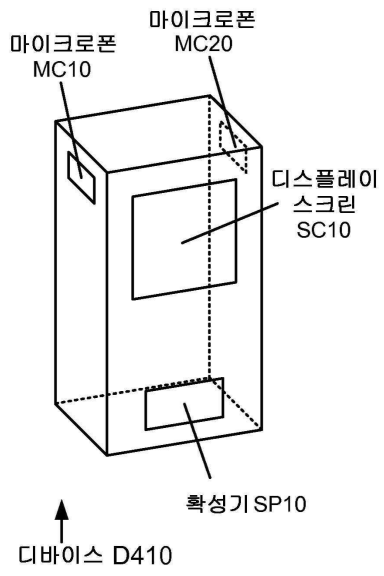
도면10b



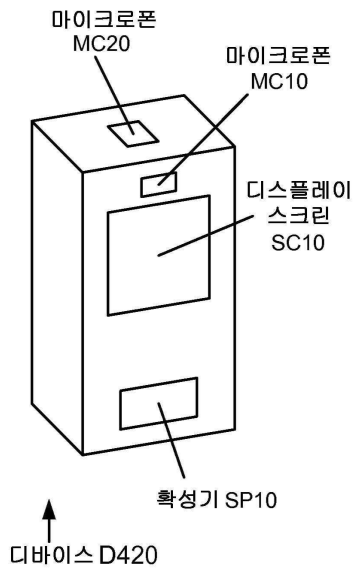
도면11a



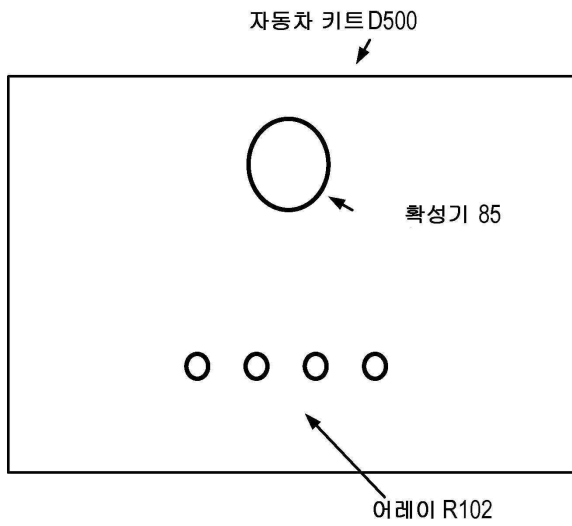
도면11b



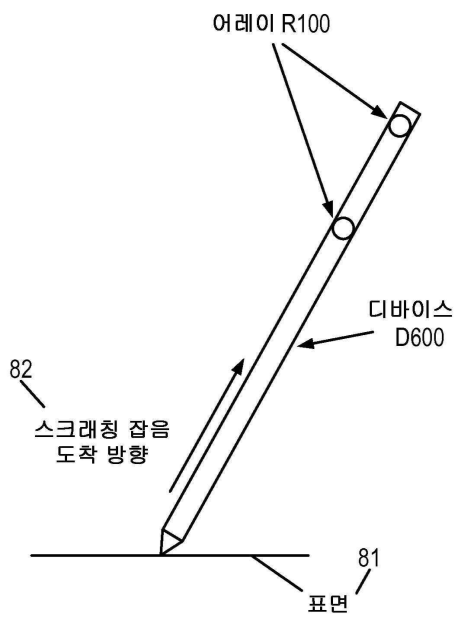
도면11c



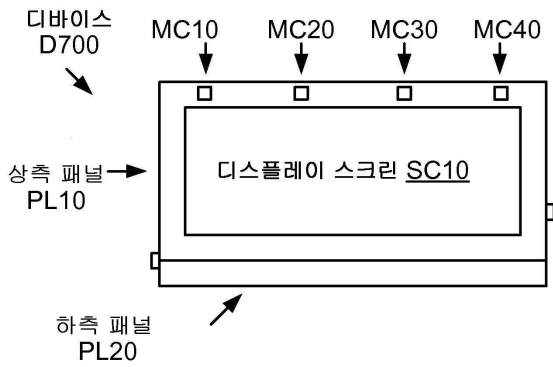
도면12a



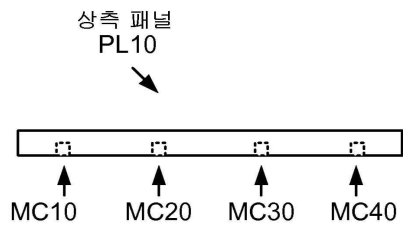
도면12b



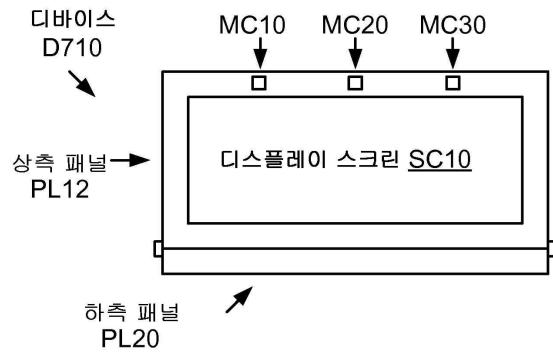
도면13a



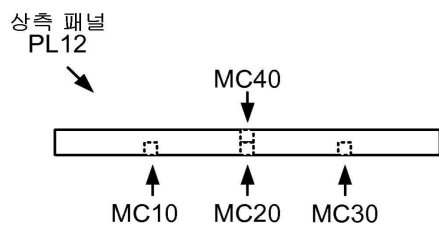
도면13b



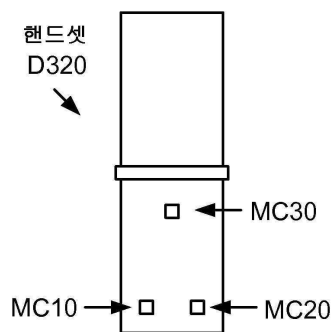
도면13c



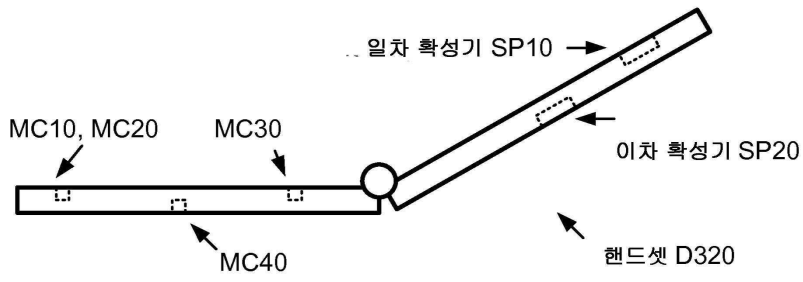
도면13d



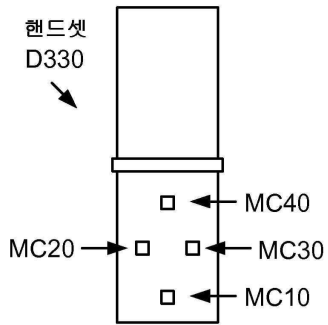
도면14a



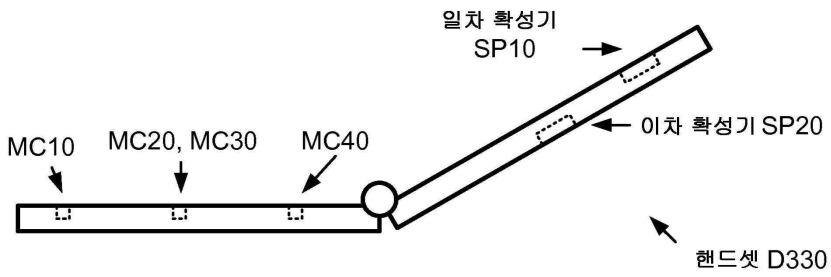
도면14b



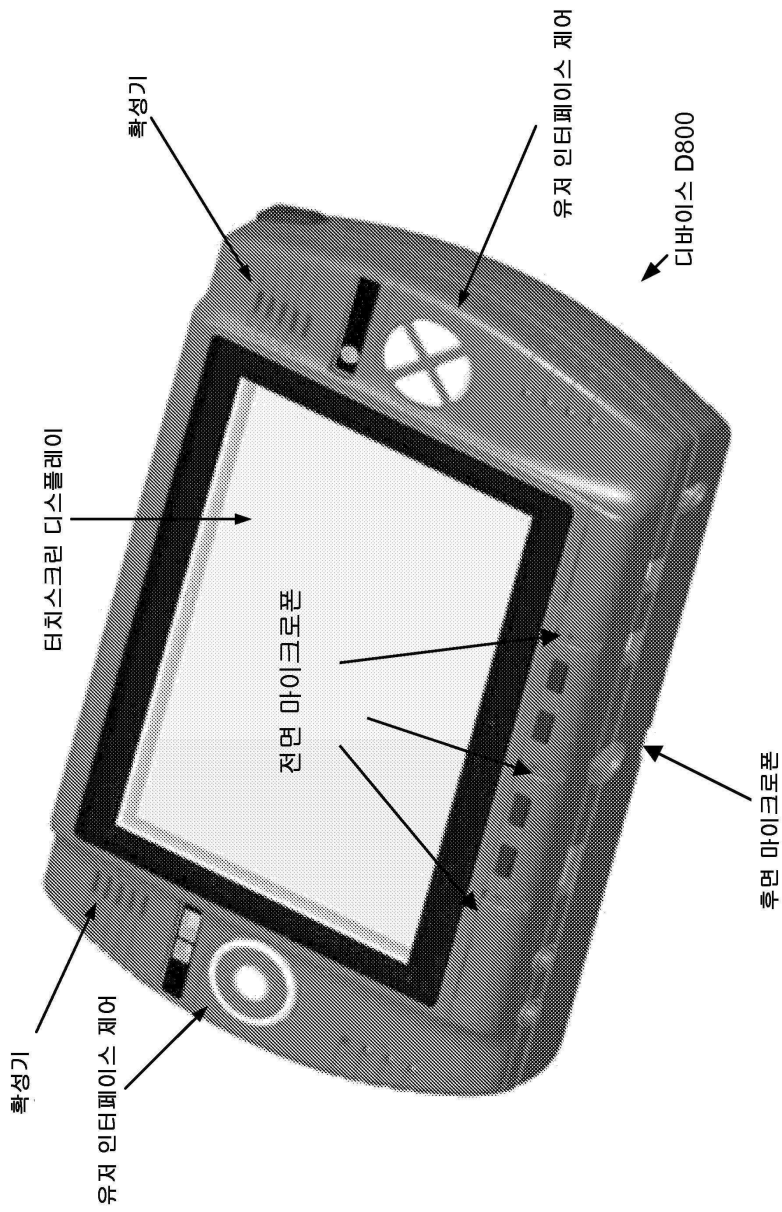
도면14c



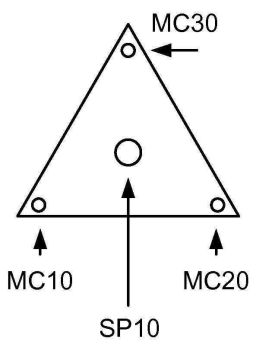
도면14d



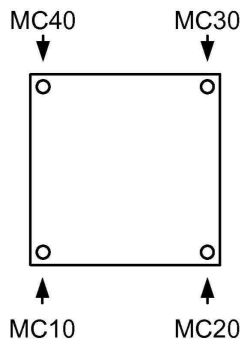
도면15



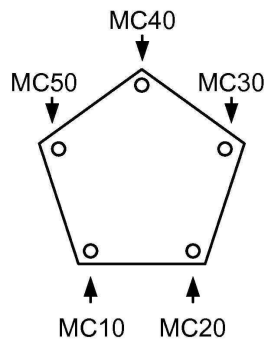
도면16a



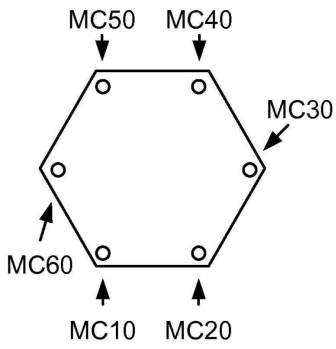
도면16b



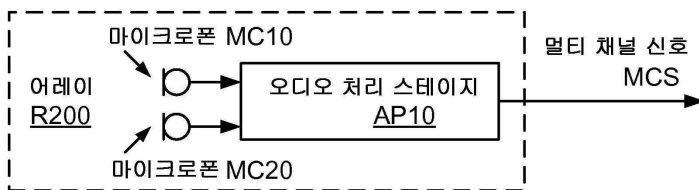
도면16c



도면16d



도면17a



도면17b

