



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년03월12일
(11) 등록번호 10-1957544
(24) 등록일자 2019년03월06일

(51) 국제특허분류(Int. Cl.)
H04R 3/00 (2006.01) H04R 5/027 (2006.01)
(21) 출원번호 10-2014-7015683
(22) 출원일자(국제) 2012년10월31일
심사청구일자 2017년10월27일
(85) 번역문제출일자 2014년06월10일
(65) 공개번호 10-2014-0089601
(43) 공개일자 2014년07월15일
(86) 국제출원번호 PCT/EP2012/071537
(87) 국제공개번호 WO 2013/068284
국제공개일자 2013년05월16일
(30) 우선권주장
11306472.9 2011년11월11일
유럽특허청(EPO)(EP)
(56) 선행기술조사문헌
EP01737271 A1
US20030016835 A1

(73) 특허권자
돌비 인터네셔널 에이비
네덜란드 1101 씨엔 암스트레담 주이두스트 헤리
커베르그백 1-35 3이 아폴로 빌딩
(72) 발명자
고든, 스벤
독일 30625 하노버 칼 비헤르트 알레 74 리서치
앤드 이노베이션 도이치 톰슨 오에이치취
배케, 요한-마르쿠스
독일 30625 하노버 칼 비헤르트 알레 74 리서치
앤드 이노베이션 도이치 톰슨 오에이치취
크뤼거, 알렉산더
독일 30625 하노버 칼 비헤르트 알레 74 리서치
앤드 이노베이션 도이치 톰슨 오에이치취
(74) 대리인
양영준, 백만기

전체 청구항 수 : 총 10 항

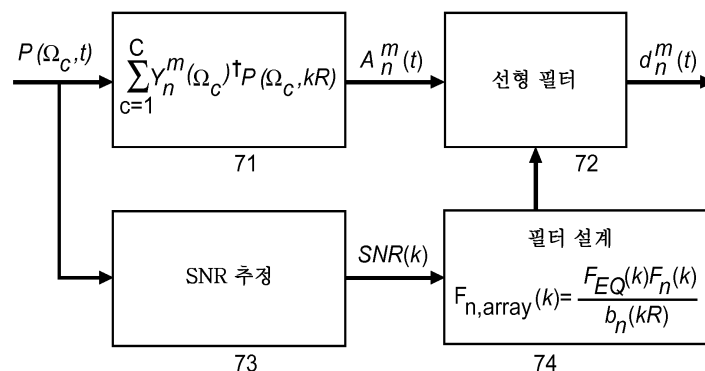
심사관 : 송근배

(54) 발명의 명칭 사운드 필드의 앰비소닉스 표현을 생성하는데 사용되는 강체구 상의 구면 마이크로폰 어레이의 신호들을 프로세싱하는 방법 및 장치

(57) 요약

구면 마이크로폰 어레이는 앰비소닉스 표현($A_n^m(t)$)을 생성하기 위한 3차원 사운드 필드($P(\Omega_c, t)$)을 캡처하는데, 여기서 구의 표면 상의 압력 분포는 어레이의 캡슐들에 의해 샘플링된다. 캡처된 사운드 필드에 대한 마이크로폰의 영향은 마이크로폰 역전달 함수(inverse microphone transfer function)를 사용하여 제거된다. 전달 함수의 역수가 전달 함수에서의 작은 값들에 대해 높은 이득을 초래하고 이들 작은 값은 트랜스듀서 잡음에 의해 영향을 받기 때문에 마이크로폰 어레이의 전달 함수의 등화는 큰 문제이다. 본 발명은 마이크로폰 어레이 캡슐들로부터의 잡음 전력과 평균 사운드 필드 전력 사이의 신호 대 잡음비를 추정하고(73), 확산 사운드 필드에 대해 원점에서의 평균 공간 신호 전력을 계산하고(74), 및 주어진 기준 전력과 원점에서의 시뮬레이션된 전력의 분수의 제곱근으로부터 등화 필터의 주파수 응답을 주파수 도메인에서 설계한다.

대표도



명세서

청구범위

청구항 1

강체구(rigid sphere) 상의 구면 마이크로폰 어레이의 마이크로폰 캡슐 신호들을 프로세싱하는 방법으로서, 상기 마이크로폰 어레이의 표면 상의 압력을 나타내는 상기 마이크로폰 캡슐 신호들을 지향성 계수들과 함께 구면 고조파 또는 앰비소닉스 표현($A_n^m(t)$)으로 변환하는 단계;

상기 마이크로폰 어레이로부터 레코딩된 평면 웨이브의 평균 소스 전력($|P_0(k)|^2$) 및 상기 마이크로폰 어레이에서의 아날로그 프로세싱에 의해 생성된 공간 비상관 잡음을 나타내는 대응 잡음 전력($|P_{noise}(k)|^2$)을 사용하여 상기 마이크로폰 캡슐 신호들의 시변 신호 대 잡음비의 추정($SNR(k)$)을 웨이브 번호(k)당 계산하는 단계;

기준, 에일리어싱, 및 잡음 신호 전력 컴포넌트를 사용하여 확산 사운드 필드에 대한 원점에서의 평균 공간 신호 전력을 웨이브 번호(k)당 계산하고, 주어진 기준 전력과 상기 원점에서의 상기 평균 공간 신호 전력의 분수의 제곱근으로부터 등화 필터의 주파수 응답을 형성하고, 적용된 전달 함수($F_{n,array}(k)$)를 획득하기 위해 상기 시변 신호 대 잡음비 추정($SNR(k)$)으로부터 유도된 잡음 최소화 필터의 이산 유한 웨이브 번호들(k)에서의 각 차수(n)에 대한 전달 함수 및 상기 마이크로폰 어레이의 역전달 함수에 상기 등화 필터의 상기 주파수 응답을 웨이브 번호(k)당 곱하는 단계; 및

선형 필터 프로세싱을 사용하여 상기 적용된 전달 함수($F_{n,array}(k)$)를 상기 구면 고조파 또는 앰비소닉스 표현($A_n^m(t)$)에 적용하여 상기 구면 고조파 또는 앰비소닉스 표현의 적용된 지향성 시간 도메인 계수들($d_n^m(t)$)을 도출하는 단계를 포함하고, 여기서, n 은 앰비소닉스 차수를 나타내고, 인덱스 n 은 0부터 유한 차수로 진행하고, m 은 차수(degree)를 나타내며, 인덱스 m 은 인덱스 n 마다 $-n$ 부터 n 으로 진행하는, 방법.

청구항 2

제1항에 있어서, 상기 잡음 전력($|P_{noise}(k)|^2$)은 $|P_0(k)|^2 = 0$ 이도록 임의의 사운드 소스가 없는 무소음 환경에서 획득되는, 방법.

청구항 3

제1항에 있어서, 상기 평균 소스 전력($|P_0(k)|^2$)은 마이크로폰 캡슐들에서의 압력의 기대값과 상기 마이크로폰 캡슐들에서 측정된 평균 신호 전력의 비교에 의해 상기 마이크로폰 캡슐들에서 측정된 압력($P_{mic}(\Omega_c, k)$)으로부터 추정되는, 방법.

청구항 4

제1항에 있어서,

상기 어레이의 상기 전달 함수($F_{n,array}(k)$)는 주파수 도메인에서 결정되고,

FFT(Fast Fourier transform)를 사용하여 상기 구면 고조파 또는 앰비소닉스 표현($A_n^m(t)$)의 계수들을 주파수

도메인으로 변환한 후 상기 전달 함수($F_{n,array}(k)$)를 곱하는 단계;

지향성 시간 도메인 계수들($d_n^m(t)$)을 획득하기 위한 상기 곱하는 단계에 의해 계산된 곱의 역 FFT, 또는, 시간 도메인에서의 FIR(Finite Impulse Response) 필터에 의한 근사화를 수행하는 단계를 포함하고,

역 FFT를 수행하는 단계;

순환 시프트를 수행하는 단계;

대응 전달 함수를 스무딩하기 위해 최종 필터 임펄스 응답에 테이퍼링 윈도우를 적용하는 단계;

n 과 m 의 조합마다 결과로서 생긴 필터 계수들과 상기 구면 고조파 또는 앰비소닉스 표현($A_n^m(t)$)의 계수들의 컨볼루션을 수행하는 단계를 포함하는, 방법.

청구항 5

제1항에 있어서, 상기 등화 필터의 전달 함수는,

$$F_{EQ}(k) = \sqrt{\frac{E\{|w_{ref}(k)|^2\}}{E\{|w'_{ref}(k)+w'_{alias}(k)|^2\}+E\{|w'_{noise}(k)|^2\}}}$$

에 의해 결정되고, E 는 기대값을 나타내고, $w_{ref}(k)$ 는 웨이브 번호(k)에 대한 기준 가중치이고, $w'_{ref}(k)$ 는 웨이브 번호(k)에 대한 최적화된 기준 가중치이고, $w'_{alias}(k)$ 는 웨이브 번호(k)에 대한 최적화된 에일리어싱 가중치이고, $w'_{noise}(k)$ 는 웨이브 번호(k)에 대한 최적화된 잡음 가중치이며, 상기 최적화는 상기 구면 마이크로폰 어레이에서의 잡음 상승에 대해 잡음이 감소되었음을 의미하는, 방법.

청구항 6

장체구 상의 구면 마이크로폰 어레이의 마이크로폰 캡슐 신호들을 프로세싱하는 장치로서,

상기 마이크로폰 어레이의 표면 상의 압력을 나타내는 상기 마이크로폰 캡슐 신호들을 지향성 계수들과 함께 구면 고조파 또는 앰비소닉스 표현($A_n^m(t)$)으로 변환하기 위한 수단;

상기 마이크로폰 어레이로부터 레코딩된 평면 웨이브의 평균 소스 전력($|P_0(k)|^2$) 및 상기 마이크로폰 어레이에서의 아날로그 프로세싱에 의해 생성된 공간 비상관 잡음을 나타내는 대응 잡음 전력($|P_{noise}(k)|^2$)을 사용하여 상기 마이크로폰 캡슐 신호들의 시변 신호 대 잡음비의 추정($SNR(k)$)을 웨이브 번호(k)당 계산하기 위한 수단;

기준, 에일리어싱, 및 잡음 신호 전력 컴포넌트를 사용하여 확산 사운드 필드에 대한 원점에서의 평균 공간 신호 전력을 웨이브 번호(k)당 계산하고, 주어진 기준 전력과 상기 원점에서의 평균 공간 신호 전력의 분수의 제곱근으로부터 등화 필터의 주파수 응답을 형성하고, 적용된 전달 함수($F_{n,array}(k)$)를 획득하기 위해, 상기 시변 신호 대 잡음비 추정($SNR(k)$)으로부터 유도된 잡음 최소화 필터의 이산 유한 웨이브 번호들(k)에서의 각 차수(n)에 대한 전달 함수 및 상기 마이크로폰 어레이의 역전달 함수에 상기 등화 필터의 상기 주파수 응답을 웨이브 번호(k)당 곱하기 위한 수단; 및

선형 필터 프로세싱을 사용하여 상기 적응된 전달 함수($F_{n,array}(k)$)를 상기 구면 고조파 또는 앰비소닉스 표현($A_n^m(t)$)에 적용하여 상기 구면 고조파 또는 앰비소닉스 표현의 적응된 지향성 시간 도메인 계수들($d_n^m(t)$)을 도출하기 위한 수단을 포함하고, 여기서, n 은 앰비소닉스 차수를 나타내고, 인덱스 n 은 0부터 유한 차수로 진행하고, m 은 차수(degree)를 나타내고, 인덱스 m 은 인덱스 n 마다 $-n$ 부터 n 으로 진행하는, 장치.

청구항 7

제6항에 있어서, 상기 잡음 전력($|P_{noise}(k)|^2$)은 $|P_0(k)|^2 = 0$ 이도록 임의의 사운드 소스가 없는 무소음 환경에서 획득되는, 장치.

청구항 8

제6항에 있어서, 상기 평균 소스 전력($|P_0(k)|^2$)은 마이크로폰 캡슐들에서의 압력의 기대값과 상기 마이크로폰 캡슐들에서 측정된 평균 신호 전력의 비교에 의해 상기 마이크로폰 캡슐들에서 측정된 압력($P_{mic}(\Omega_c, k)$)으로부터 추정되는, 장치.

청구항 9

제6항에 있어서,

상기 어레이의 상기 전달 함수($F_{n,array}(k)$)는 주파수 도메인에서 결정되고,

FFT(Fast Fourier transform)를 사용하여 상기 구면 고조파 또는 앰비소닉스 표현($A_n^m(t)$)의 계수들을 주파수 도메인으로 변환한 후 상기 전달 함수($F_{n,array}(k)$)를 곱하는 것; 및

지향성 시간 도메인 계수들($d_n^m(t)$)을 획득하기 위한 상기 곱하는 것에 의해 계산된 곱의 역 FFT, 또는, 시간 도메인에서의 FIR 필터에 의한 근사화를 수행하는 것

을 포함하고,

역 FFT를 수행하고;

순환 시프트를 수행하고;

대응 전달 함수를 스무딩하기 위해 최종 필터 임펄스 응답에 테이퍼링 윈도우를 적용하고;

n 과 m 의 조합마다 결과로서 생긴 필터 계수들과 상기 구면 고조파 또는 앰비소닉스 표현($A_n^m(t)$)의 계수들의 컨볼루션을 수행하는 것

을 포함하는, 장치.

청구항 10

제6항에 있어서, 상기 등화 필터의 전달 함수는,

$$F_{EQ}(k) = \sqrt{\frac{E\{|w_{ref}(k)|^2\}}{E\{|w'_{ref}(k)+w'_{alias}(k)|^2\}+E\{|w'_{noise}(k)|^2\}}}$$

에 의해 결정되고, E 는 기대값을 나타내고, $w_{ref}(k)$ 는 웨이브 번호(k)에 대한 기준 가중치이고, $w'_{ref}(k)$ 는 웨이브 번호(k)에 대한 최적화된 기준 가중치이고, $w'_{alias}(k)$ 는 웨이브 번호(k)에 대한 최적화된 에일리 어싱 가중치이고, $w'_{noise}(k)$ 는 웨이브 번호(k)에 대한 최적화된 잡음 가중치이며, 상기 최적화는 상기 구면

마이크로폰 어레이에서의 잡음 상승에 대해 잡음이 감소되었음을 의미하는, 장치.

발명의 설명

기술 분야

[0001] 본 발명은 사운드 필드의 앰비소닉스 표현(Ambisonics representation)을 생성하는데 사용되는 강체구 상의 구면 마이크로폰 어레이의 신호들을 프로세싱하는 방법 및 장치에 관한 것으로서, 인버스 마이크로폰 어레이 응답에 등화 필터(equalisation filter)가 적용된다.

배경 기술

[0002] 구면 마이크로폰 어레이는 3차원 사운드 필드를 캡처하는 능력을 제공한다. 사운드 필드를 저장하고 프로세싱하는 하나의 방식은 앰비소닉스(Ambisonics) 표현이다. 앰비소닉스는 스위트 스팟(sweet spot)이라고도 알려진 원점(point of origin) 주변 영역에서 사운드 필드(sound field)를 설명하기 위한 직교 정규(orthonormal) 구면 함수들을 사용한다. 이러한 설명의 정확도는 앰비소닉스 차수(Ambisonics order)(N)에 의해 결정되며, 여기서 유한 개의 앰비소닉스 계수들은 사운드 필드를 설명한다. 구면 어레이의 최대 앰비소닉스 차수는 마이크로폰 캡슐의 개수에 의해 제한되는데, 이 개수는 앰비소닉스 계수들의 개수($O = (N + 1)^2$)와 같거나 그보다 많아야 한다.

[0003] 앰비소닉스 표현의 하나의 이점은 사운드 필드의 재생이 임의의 주어진 라우드 스피커(loudspeaker) 배열에 개별 적용될 수 있다는 것이다. 게다가, 이러한 표현은 포스트 프로덕션(post production)에서 빔 형성 기법들을 사용하는 서로 다른 마이크로폰 특성들의 시뮬레이션을 가능하게 한다.

[0004] B-포맷은 앰비소닉스의 하나의 공지 예이다. B-포맷 마이크로폰은 사면체 상의 4개의 캡슐이 앰비소닉스 차수가 1인 사운드 필드를 캡처하도록 요구한다.

[0005] 1보다 큰 차수의 앰비소닉스는 HOA(Higher Order Ambisonics)라고 지칭되고, HOA 마이크로폰들은 통상 강체구 상의 구면 마이크로폰 어레이들로서, 예를 들어 mhAcoustics의 아이겐마이크(Eigenmike)이다. 앰비소닉스 프로세싱의 경우, 구의 표면에 대한 압력 분포는 어레이의 캡슐들에 의해 샘플링된다. 그 후 샘플링된 압력은 앰비소닉스 표현으로 변환된다. 이러한 앰비소닉스 표현은 사운드 필드를 설명하지만, 마이크로폰 어레이의 영향을 포함하고 있다. 캡처된 사운드 필드에 대한 마이크로폰들의 영향은 인버스 마이크로폰 어레이 응답을 사용하여 제거되는데, 평면 웨이브의 사운드 필드를 마이크로폰 캡슐들에서 측정되는 압력으로 변환한다. 이는 마이크로폰 어레이의 사운드 필드에 대한 간섭 및 캡슐들의 지향성을 시뮬레이션한다.

발명의 내용

해결하려는 과제

[0006] 구면 마이크로폰 어레이에 의해 캡처되는 재구성된 앰비소닉스 신호의 왜곡된 스펙트럼 전력은 등화되어야 한다. 한편, 이러한 왜곡은 공간 에일리어싱(aliasing) 신호 전력에 의해 일어난다. 다른 한편, 강체구(rigid sphere) 상의 구면 마이크로폰 어레이에 대한 잡음 감소로 인해, 고차 계수들은 구면 고조파 표현에서 누락되고, 이들 누락된 계수는 특히 빔 형성 응용들에 대해 재구성된 신호의 전력 스펙트럼을 언밸런싱한다.

[0007] 본 발명에 의해 해결될 문제점은 구면 마이크로폰 어레이에 의해 캡처되는 재구성된 앰비소닉스 신호의 스펙트럼 전력의 왜곡을 감소시키고, 스펙트럼 전력을 등화하는 것이다. 이 문제점은 청구항 제1항에 개시된 방법에 의해 해결된다. 이 방법을 활용한 장치가 청구항 제2항에 개시된다.

과제의 해결 수단

[0008] 본 발명의 프로세싱은 재구성된 앰비소닉스 신호의 주파수 스펙트럼을 밸런싱하는 필터를 결정하는 역할을 한다. 필터링되고 재구성된 앰비소닉스 신호의 신호 전력은 분석되며, 이로써 평균 공간 에일리어싱 전력 및 누락된 고차 앰비소닉스 계수들의 영향은 앰비소닉스 디코딩 및 빔 형성 응용들에 대하여 설명된다. 이들 결과로부터, 재구성된 앰비소닉스 신호의 평균 주파수 스펙트럼을 밸런싱하는 사용하기 편한 등화 필터가 유도되는데, 즉 레코딩의 사용된 디코딩 계수들 및 신호 대 잡음비(SNR)에 따라, 원점에서의 평균 전력이 추정된다.

- [0009] 등화 필터는 다음으로부터 획득된다.
- [0010] - 마이크로폰 어레이 캡슐로부터 평균 사운드 필드 전력과 잡음 전력 사이의 신호 대 잡음비의 추정.
- [0011] - 확산 사운드 필드에 대한 원점에서의 평균 공간 신호 전력의 웨이브 번호(k)당 계산. 이러한 시뮬레이션은 모든 신호 전력 컴포넌트들(기준, 에일리어싱, 및 잡음)을 포함한다.
- [0012] - 등화 필터의 주파수 응답은 원점에서 주어진 기준 전력과 계산된 평균 공간 신호 전력의 분수의 제곱근으로부터 형성된다.
- [0013] - 적응된 전달 함수($F_{n,array}(k)$)를 획득하기 위해, 신호 대 잡음비 추정으로부터 유도된 잡음 최소화 필터의 (이산 유한 웨이브 번호들(k)에서의 각 차수(n)에 대한) 전달 함수 및 마이크로폰 어레이의 역전달 함수에 등화 필터의 주파수 응답을 (웨이브 번호(k) 당) 곱함.
- [0014] 최종 필터는 레코딩된 사운드 필드의 구면 고조파 표현 또는 재구성 신호들에 적용된다. 이러한 필터의 설계는 계산 복잡도가 매우 높다. 이롭게도, 계산이 복잡한 프로세싱은 고정 필터 설계 파라미터들의 계산을 사용함으로써 감소될 수 있다. 이들 파라미터는 주어진 마이크로폰 어레이에 대해 일정하고, 룩업 테이블에 저장될 수 있다. 이는 관리 가능한 계산 복잡성을 갖는 시변 적응성 필터(time-variant adaptive filter) 설계를 용이하게 한다.
- [0015] 필터는 고주파수에서 상승된 평균 신호 전력을 제거한다는 이점을 갖는다. 또한, 필터는 저주파수에서 구면 고조파 표현에서 빔 형성 디코더의 주파수 응답을 뺄런싱한다. 본 발명의 필터를 사용하지 않으면, 사운드들을 레코딩하는 구면 마이크로폰 어레이로부터의 재구성된 사운드가 언밸런싱되는데, 그 이유는 레코딩된 사운드 필드의 전력이 모든 주파수 서브대역에서 정확히 재구성되지 않기 때문이다.
- [0016] 원칙적으로, 본 발명의 방법은 강체구(rigid sphere) 상의 구면 마이크로폰 어레이의 마이크로폰 캡슐 신호들을 프로세싱하는데 적합하며, 상기 방법은 다음의 단계들을 포함한다.
- [0017] - 마이크로폰 어레이의 표면 상의 압력을 표현하는 마이크로폰 캡슐 신호들을 구면 고조파 또는 앰비소닉스 표현($A_n^m(t)$)으로 변환하는 단계;
- [0018] - 마이크로폰 어레이로부터 레코딩된 평면 웨이브의 평균 소스 전력($|P_0(k)|^2$) 및 마이크로폰 어레이에서의 아날로그 프로세싱에 의해 생성된 공간 비상관 잡음을 표현하는 대응 잡음 전력($|P_{noise}(k)|^2$)을 사용하여 마이크로폰 캡슐 신호들의 시변 신호 대 잡음비의 추정($SNR(k)$)을 웨이브 번호(k)당 계산하는 단계;
- [0019] - 기준, 에일리어싱, 및 잡음 신호 전력 컴포넌트를 사용하여 확산 사운드 필드에 대한 원점에서의 평균 공간 신호 전력을 웨이브 번호(k)당 계산하고,
- [0020] 주어진 기준 전력과 원점에서의 평균 공간 신호 전력의 분수의 제곱근으로부터 등화 필터의 주파수 응답을 형성하고,
- [0021] 적응된 전달 함수($F_{n,array}(k)$)를 획득하기 위해, 신호 대 잡음비 추정($SNR(k)$)으로부터 유도된 잡음 최소화 필터의 (이산 유한 웨이브 번호들(k)에서의 각 차수(n)에 대한) 전달 함수 및 마이크로폰 어레이의 역전달 함수에 등화 필터의 주파수 응답을 (웨이브 번호(k)당) 곱하는 단계;
- [0022] - 선형 필터 프로세싱을 사용하여 적응된 전달 함수($F_{n,array}(k)$)를 구면 고조파 표현($A_n^m(t)$)에 적용하여 적응된 지향성 계수들($d_n^m(t)$)을 도출하는 단계.

- [0023] 원칙적으로, 본 발명의 장치는 강체구 상의 구면 마이크로폰 어레이의 마이크로폰 캡슐 신호들을 프로세싱하는데 적합하며, 상기 장치는 다음을 포함한다.
- [0024] - 마이크로폰 어레이의 표현에 대한 압력을 표현하는 마이크로폰 캡슐 신호들을 구면 고조파 또는 앰비소닉스 표현($A_n^m(t)$)으로 변환하는데 적응된 수단;
- [0025] - 마이크로폰 어레이로부터 레코딩된 평면 웨이브의 평균 소스 전력($|P_0(k)|^2$) 및 마이크로폰 어레이에서의 아날로그 프로세싱에 의해 생성된 공간 비상관 잡음을 표현하는 대응 잡음 전력($|P_{noise}(k)|^2$)을 사용하여 마이크로폰 캡슐 신호들의 시변 신호 대 잡음비의 추정($SNR(k)$)을 웨이브 번호(k)당 계산하는데 적응된 수단;
- [0026] - 기준, 에일리어싱, 및 잡음 신호 전력 컴포넌트를 사용하여 확산 사운드 필드에 대해 원점에서의 평균 공간 신호 전력을 웨이브 번호(k)당 계산하고,
- [0027] 주어진 기준 전력과 원점에서의 평균 공간 신호 전력의 분수의 제곱근으로부터 등화 필터의 주파수 응답을 형성하고,
- [0028] 적응된 전달 함수($F_{n,array}(k)$)를 획득하기 위해, 신호 대 잡음비 추정($SNR(k)$)으로부터 유도된 잡음 최소 필터의, 이산 유한 웨이브 번호들(k)에서의 각 차수(n)에 대한 전달 함수 및 마이크로폰 어레이의 역전달 함수에 등화 필터의 주파수 응답을 웨이브 번호(k)당 곱하는데 적응된 수단;
- [0029] - 선형 필터 프로세싱을 사용하여 적응된 전달 함수($F_{n,array}(k)$)를 구면 고조파 표현($A_n^m(t)$)에 적용하여 적응된 지향성 계수들($d_n^m(t)$)을 도출하는데 적응된 수단.
- [0030] 본 발명의 추가적인 유리한 실시예들은 개별 종속항들에 개시되어 있다.

도면의 간단한 설명

- [0031] 본 발명의 예시적인 실시예들은 다음의 첨부 도면을 참조하여 설명된다.
- 도 1은 강체구 상에 32개의 캡슐이 있는 마이크로폰 어레이를 위한 최종 라우드 스피커 가중치로부터의 기준, 에일리어싱, 및 잡음 컴포넌트들의 전력을 도시한다.
- 도 2는 $SNR(k) = 20\text{dB}$ 인 경우 잡음 감소 필터를 도시한다.
- 도 3은 종래의 앰비소닉스 디코더를 사용하여 도 2의 최적화 필터를 따르는 가중 컴포넌트들의 평균 전력을 도시한다.
- 도 4는 $D_n^m(\Omega_l) = Y_n^m(\Omega_{[0,0]T})$ 인 경우 빔 형성을 사용하여 잡음 최적화 필터가 적용된 후의 가중 컴포넌트들의 평균 전력을 도시한다.
- 도 5는 종래의 앰비소닉스 디코더 및 20dB의 $SNR(k)$ 의 경우 최적화된 어레이 응답을 도시한다.
- 도 6은 빔 형성 디코더 및 20dB의 $SNR(k)$ 의 경우 최적화된 어레이 응답을 도시한다.
- 도 7은 본 발명에 따른 적응형 앰비소닉스 프로세싱을 위한 블록도를 도시한다.
- 도 8은 종래의 앰비소닉스 디코딩을 사용하여 잡음 최적화 필터($F_n(k)$) 및 필터($F_{EQ}(k)$)가 적용된 후의 최종 가중치의 평균 전력을 도시하며, 이에 따라 최적화된 가중치, 기준 가중치, 및 잡음 가중치의 전력이 비교된

다.

도 9는 빔 형성 디코더를 사용하여 잡음 최적화 필터($F_n(k)$) 및 필터($F_{EQ}(k)$)가 적용된 후의 가중 컴포넌트들의 평균 전력을 도시하며, 여기서 $D_n^m(\Omega_l) = Y_n^m(\Omega_{[0,0]^T})$ 이고, 이에 따라 최적화된 가중치, 기준 가중치, 및 잡음 가중치의 전력이 비교된다.

발명을 실시하기 위한 구체적인 내용

[0032] 구면 마이크로폰 어레이 프로세싱 - 앰비소닉스 이론

[0033] 앰비소닉스 디코딩은 평면 웨이브의 사운드 필드를 방출하는 라우드 스피커를 가정함으로써 규정된다(M.A. Poletti, "Three-Dimensional Surround Sound Systems Based on Spherical Harmonics", Journal Audio Engineering Society, vol.53, no.11, pages 1004-1025, 2005 참조).

수학식 1

[0034]
$$w(\Omega_l, k) = \sum_{n=0}^N \sum_{m=-n}^n D_n^m(\Omega_l) d_n^m(k)$$

[0035] L개의 라우드 스피커들의 배열은 앰비소닉스 계수들($d_n^m(k)$)에 저장된 3차원 사운드 필드를 재구성한다. 프로세싱은 웨이브 번호(k)마다 별도로 수행된다.

수학식 2

[0036]
$$k = \frac{2\pi f}{c_{\text{sound}}}$$

[0037] 여기서 f는 주파수이고, c_{sound} 는 사운드의 속도이다. 인덱스(n)는 0부터 유한 차수(N)로 진행하지만, 인덱스(m)는 인덱스(n)마다 $-n$ 로부터 n 로 진행한다. 이에 따라 계수들의 총 개수는 $O = (N + 1)^2$ 이다. 라우드 스피커 위치는 구면 좌표계에서 방향 벡터($\Omega_l = [\theta_l, \phi_l]^T$)에 의해 정의되고, $[\cdot]^T$ 는 벡터의 전치 버전(transposed version)을 나타낸다.

[0038] 수학식 1은 앰비소닉스 계수들($d_n^m(k)$)의 라우드 스피커 가중치들($w(\Omega_l, k)$)로의 변환을 정의한다. 이들 가중치는 라우드 스피커들의 구동 함수이다. 모든 스피커 가중치들의 슈퍼пози션은 사운드 필드를 재구성한다.

[0039] 디코딩 계수들($D_n^m(\Omega_l)$)은 일반적으로 앰비소닉스 디코딩 프로세싱을 설명한다. 이는 Morag Agmon, Boaz Rafaely, "Beamforming for a Spherical-Aperture Microphone", IEEEI, pages 227-230, 2008의 섹션 3 (ω_{nm}^*) 및 섹션 3.2에서의 전술한 M.A. Poletti 논문에서 주어진 모드 매칭 디코딩 행렬의 행들에 도시된 빔 패턴의 공액 복소 계수들을 포함한다. Johann-Markus Batke, Florian Keiler, "Using VBAP-Derived Panning Functions for 3D Ambisonics Decoding", Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics, 6-7 May 2010, Paris, France의 섹션 4에 설명된 상이한 방식의 프로세싱은 임의의 3차원 라우드 스피커 배열을 위한 디코딩 행렬을 계산하기 위한 벡터 기반 진폭 패닝(panning)을 사용한다. 이들 행렬의 행 원소들은 또한 계수들($D_n^m(\Omega_l)$)에 의해 설명된다.

[0040] 앰비소닉스 계수들($d_n^m(k)$)은 Boaz Rafaely, "Plane-wave decomposition of the sound field on a sphere

by spherical convolution", J. Acoustical Society of America, vol.116, no.4, pages 2149-2157, 2004의 섹션 3에 설명된 바와 같이 평면 웨이브의 슈퍼포지션으로 항상 분해될 수 있다. 이에 따라 분석은 방향(Ω_s)으로부터 충돌된 평면 웨이브의 계수들로 제한될 수 있다.

수학식 3

$$d_n^m \text{ plane}(k) = P_0(k) Y_n^m(\Omega_s)^*$$

평면 웨이브의 사운드 필드를 방출하는 라우드 스피커들의 가정을 위해 평면 웨이브의 계수들($d_n^m \text{ plane}(k)$) 이 정의된다. 원점에서의 압력은 웨이브 번호(k)에 대해 $P_0(k)$ 에 의해 정의된다. 공액 복소 구면 고조파($Y_n^m(\Omega_s)^*$)는 평면 웨이브의 지향성 계수들을 나타낸다. 전술한 M.A. Poletti 논문에서 주어진 구면 고조파($Y_n^m(\Omega_s)$)의 정의가 사용된다.

구면 고조파는 앰비소닉스 표현들의 직교 정규 베이스 함수들로서 다음의 수학적식을 충족시킨다.

수학식 4

$$\delta_{n-n'} \delta_{m-m'} = \int_{\Omega \in S^2} Y_n^m(\Omega) Y_{n'}^{m'}(\Omega)^* d\Omega$$

여기서 δ_q 는 델타 임펄스로서 다음의 수학적식으로 표현된다.

수학식 5

$$\delta_q = \begin{cases} 1, & q = 0 \text{ 인 경우} \\ 0, & \text{그 밖의 경우} \end{cases}$$

구면 마이크로폰 어레이는 구 표면 상의 압력을 샘플링하는데, 여기서 샘플링 지점들의 개수는 앰비소닉스 계수들의 개수($O = (N + 1)^2$) 이상이어야 한다. 앰비소닉스 차수가 N 인 경우, 또한 샘플링 지점들은 구의 표면에 걸쳐 균일하게 분포되어야 하는데, O 개의 지점들의 최적 분포가 차수 $N = 1$ 인 경우에만 정확히 알려진다. 더 높은 차수의 경우, 구의 샘플링의 양호한 근사화가 존재한다(2007년 2월 1일에 방문된 mh acoustics 홈페이지 <http://www.mhacoustics.com> 및 F. Zotter, "Sampling Strategies for Acoustic Holography/Holophony on the Sphere", Proceedings of the NAG-DAGA, 23-26 March 2009, Rotterdam를 참조한다).

최적의 샘플링 지점들(Ω_c)의 경우, 수학식 4로부터의 적분은 수학식 6으로부터의 이산 합과 동등하다.

수학식 6

$$\delta_{n-n'} \delta_{m-m'} = \frac{4\pi}{C} \sum_{c=1}^C Y_n^m(\Omega_c) Y_{n'}^{m'}(\Omega_c)^*$$

여기서, $C \geq (N + 1)^2$ 인 경우 $n' \leq N$ 및 $n \leq N$ 이고, C 는 캡슐들의 총 개수이다.

[0051] 비-최적 샘플링 지점들에 관한 안정한 결과들을 이루기 위해, 공액 복소 구면 고조파는 $L \times O$ 구면 고조파 행렬(\underline{Y})로부터 획득된 의사 역행렬(\underline{Y}^\dagger)의 열로 대체될 수 있는데, 여기서 구면 고조파($Y_n^m(\Omega_c)$)의 O 개의 계수들이 \underline{Y} 의 행 요소들이며, 전술된 Moreau/Daniel/Bertet 논문의 섹션 3.2.2를 참조한다.

수학식 7

[0052]
$$\underline{Y}^\dagger = (\underline{Y}^H \underline{Y})^{-1} \underline{Y}^H$$

[0053] 다음에서, \underline{Y}^\dagger 의 열 요소들이 $Y_n^m(\Omega_c)^\dagger$ 로 표시되며, 이로써 수학식 6으로부터의 직교 정규 조건이 다음의 수학식에 대해 충족된다.

수학식 8

[0054]
$$\delta_{n-n'} \delta_{m-m'} = \sum_{c=1}^C Y_n^m(\Omega_c) Y_{n'}^{m'}(\Omega_c)^\dagger$$

[0055] 여기서, $C \geq (N+1)^2$ 인 경우 $n' \leq N$ 및 $n \leq N$ 이다.

[0056] 구면 마이크로폰 어레이가 구의 표면 상에 거의 균일하게 분포된 캡슐들을 갖고 캡슐들의 개수가 O 보다 크다고 가정하면, 다음의 수학식이 유효한 표현이 된다.

수학식 9

[0057]
$$Y_n^m(\Omega_c)^\dagger \approx \frac{4\pi}{C} Y_n^m(\Omega_c)^*$$

[0058] 구면 마이크로폰 어레이 프로세싱 - 프로세싱의 시뮬레이션

[0059] 강체구(rigid sphere)(단단하고 고정됨) 상의 구면 마이크로폰 어레이들을 위한 전체 HOA 프로세싱 체인은 캡슐에서의 압력 추정, HOA 계수들의 계산, 및 라우드 스피커 가중치들에 대한 디코딩을 포함한다. 구면 고조파 표현에서의 마이크로폰 어레이의 설명은 주어진 디코더를 위한 원점에서 평균 스펙트럼 전력의 추정을 가능하게 한다. 모드 매칭 앰비소닉스 디코더 및 단순 빔 형성 디코더를 위한 전력이 평가된다. 스위트 스팟에서의 추정 평균 전력은 등화 필터를 설계하는데 사용된다.

[0060] 다음의 섹션은 $w(k)$ 를 기준 가중치($w_{\text{ref}}(k)$), 공간 에일리어싱 가중치($w_{\text{alias}}(k)$), 및 잡음 가중치($w_{\text{noise}}(k)$)로 분해하는 것을 설명한다. 에일리어싱은 유한 차수(N)를 위해 연속적인 사운드 필드의 샘플링에 의해 일어나고, 잡음은 캡슐마다 도입된 공간 비상관 신호 부분들을 시뮬레이션한다. 공간 에일리어싱은 주어진 마이크로폰 어레이를 위해 제거될 수 없다.

[0061] 구면 마이크로폰 어레이 프로세싱 - 캡슐 신호들의 시뮬레이션

[0062] 강체구의 표면 상의 마이크로폰 어레이에 대한 충돌 평면 웨이브의 전달 함수는 전술한 M.A. Poletti 논문의 섹션 2.2의 수학식 19에서 다음의 수학식에 의해 규정된다.

수학식 10

$$b_n(kR) = \frac{4\pi i^{n+1}}{(kR)^2 \left. \frac{d h_n^{(1)}(kr)}{d kr} \right|_{kr=kR}}$$

[0063]

[0064] 여기서, $h_n^{(1)}(kr)$ 는 제1 종류의 한켈 함수(Hankel function)이고, 반지름(r)은 구의 반지름(R)과 같다. 전달 함수는 강체구 상의 압력을 분산시키는 물리적 원리로부터 유도되는데, 이는 강체구의 표면에 반지름 방향 속도가 사라진다는 것을 의미한다. 다시 말하면, 진입 및 분산된 사운드 필드의 반지름 방향 유도의 슈퍼포지션은 영이며, "Fourier Acoustics"이라는 책의 섹션 6.10.3를 참조한다.

[0065] 이에 따라, Ω_s 로부터 충돌하는 평면 웨이브에 대한 위치(Ω)에서의 구의 표면 상의 압력은 Moreau/Daniel/Bertet 논문의 섹션 3.2.1의 수학식 21에서 다음의 수학식에 의해 주어진다.

수학식 11

$$\begin{aligned} P(\Omega, kR) &= \sum_{n=0}^{\infty} \sum_{m=-n}^n b_n(kR) Y_n^m(\Omega) d_n^m(k) \\ &= \sum_{n=0}^{\infty} \sum_{m=-n}^n b_n(kR) Y_n^m(\Omega) Y_n^m(\Omega_s)^* P_0(k) \end{aligned}$$

[0066]

[0067] 등방성 잡음 신호($P_{\text{noise}}(\Omega_c, k)$)는 트랜스듀서 잡음을 시뮬레이션하기 위해 추가되는데, 여기서 “등방성(isotropic)”은 캡슐들의 잡음 신호들이 공간적으로 상관되지 않는다는 점을 의미하며, 이는 시간 도메인에서의 상관을 포함하지 않는다.

[0068] 압력은 마이크로폰 어레이의 최대 차수(N)에 대해 계산된 압력($P_{\text{ref}}(\Omega_c, kR)$)과 나머지 차수로부터의 압력으로 분리될 수 있고, 전술한 Rafaely "Analysis and design ..." 논문에서의 섹션 7의 수학식 24를 참조한다. 마이크로폰 어레이의 차수는 이들 신호 컴포넌트를 재구성하는데 충분하지 않기 때문에 나머지 차수들로부터의 압력($P_{\text{alias}}(\Omega_c, kR)$)은 공간 에일리어싱 압력이라고 지칭된다. 이에 따라, 캡슐(\mathcal{C})에 레코딩된 총 압력은 다음의 수학식에 의해 정의된다.

수학식 12

$$P(\Omega_c, kR) = P_{\text{ref}}(\Omega_c, kR) + P_{\text{alias}}(\Omega_c, kR) + P_{\text{noise}}(\Omega_c, k) \quad (12a)$$

$$\begin{aligned} &= \sum_{n=0}^N \sum_{m=-n}^n b_n(kR) Y_n^m(\Omega_c) Y_n^m(\Omega_s)^* P_0(k) \\ &\quad + \sum_{n=N+1}^{\infty} \sum_{m=-n}^n b_n(kR) Y_n^m(\Omega_c) Y_n^m(\Omega_s)^* P_0(k) \\ &\quad + P_{\text{noise}}(\Omega_c, k) . \end{aligned} \quad (12b)$$

[0069]

[0070] 구면 마이크로폰 어레이 프로세싱 - 앰비소닉스 인코딩

[0071] 앰비소닉스 계수들($d_n^m(k)$)은 수학식 13a에 주어진 수학식 11의 도치에 의해 캡슐에서의 압력으로부터 획득되고, 전술한 Moreau/Daniel/Bertet 논문의 섹션 3.2.2의 수학식 26을 참조한다. 구면 고조파($Y_n^m(\Omega_c)$)는 수

함식 8을 사용하여 $Y_n^m(\Omega_c)^\dagger$ 에 의해 반전되고, 전달 함수($b_n(kR)$)는 그 역에 의해 등화된다.

수학식 13

$$d_n^m(k) = \sum_{c=1}^C \frac{Y_n^m(\Omega_c)^\dagger P(\Omega_c, kR)}{b_n(kR)} \quad (13a)$$

$$= \sum_{c=1}^C \frac{Y_n^m(\Omega_c)^\dagger (P_{\text{ref}}(\Omega_c, kR) + P_{\text{alias}}(\Omega_c, kR) + P_{\text{noise}}(\Omega_c, k))}{b_n(kR)} \quad (13b)$$

$$= d_{n \text{ ref}}^m(k) + d_{n \text{ alias}}^m(k) + d_{n \text{ noise}}^m(k) \quad (13c)$$

[0072]

[0073] 앰비소닉스 계수들($d_n^m(k)$)은 수학식 13b 및 13c에 도시된 바와 같이 수학식 13a 및 12a를 사용하여 기준 계수들($d_{n \text{ ref}}^m(k)$), 에일리어싱 계수들($d_{n \text{ alias}}^m(k)$), 및 잡음 계수들($d_{n \text{ noise}}^m(k)$)로 분리될 수 있다.

[0074]

구면 마이크로폰 어레이 프로세싱 - 앰비소닉스 디코딩

[0075]

최적화는 원점에서의 최종 라우드 스피커 가중치 $w(k)$ 를 사용한다. 모든 스피커들이 원점까지 동일한 거리를 갖는다고 가정하는데, 결과적으로 모든 라우드 스피커 가중치들에 대한 합은 $w(k)$ 가 된다. 수학식 14는 수학식 1 및 수학식 13b로부터 $w(k)$ 를 제공하는데, 여기서 L 은 라우드 스피커들의 개수이다.

수학식 14

$$w(k) = \sum_{l=1}^L \sum_{n=0}^N \sum_{m=-n}^n D_n^m(\Omega_l) \times \sum_{c=1}^C \frac{Y_n^m(\Omega_c)^\dagger (P_{\text{ref}}(\Omega_c, kR) + P_{\text{alias}}(\Omega_c, kR) + P_{\text{noise}}(\Omega_c, k))}{b_n(kR)} \quad (14a)$$

$$= w_{\text{ref}}(k) + w_{\text{alias}}(k) + w_{\text{noise}}(k) \quad (14b)$$

[0076]

[0077] 수학식 14b에서 알 수 있듯이, $w(k)$ 는 3개의 가중치($w_{\text{ref}}(k)$, $w_{\text{alias}}(k)$, 및 $w_{\text{noise}}(k)$)로 분리될 수 있다. 간략화를 위해, 진술된 Rafaely "Analysis and design ..." 논문의 섹션 7의 수학식 24에 주어진 위치 오차는 여기에 고려되지 않는다.

[0078]

디코딩에서, 기준 계수들은 차수(n)의 합성 생성된 평면 웨이브가 생성할 가중치들이다. 다음 수학식 15a에서 수학식 12b로부터의 기준 압력($P_{\text{ref}}(\Omega_c, kR)$)은 수학식 14a에 대입되며, 이로써 압력 신호들($P_{\text{alias}}(\Omega_c, kR)$ 및 $P_{\text{noise}}(\Omega_c, k)$)이 무시된다(즉, 0으로 설정됨).

수학식 15

$$w_{\text{ref}}(k) = \sum_{l=1}^L \sum_{n=0}^N \sum_{m=-n}^n D_n^m(\Omega_l) \times \sum_{n'=0}^N \sum_{m'=-n'}^{n'} Y_{n'}^{m'}(\Omega_s)^* \frac{b_{n'}(kR)}{b_n(kR)} \sum_{c=1}^C Y_n^m(\Omega_c)^\dagger Y_{n'}^{m'}(\Omega_c) P_0(k) \quad (15a)$$

$$= \sum_{l=1}^L \sum_{n=0}^N \sum_{m=-n}^n D_n^m(\Omega_l) Y_n^m(\Omega_s)^* P_0(k) \\ = \sum_{l=1}^L \sum_{n=0}^N \sum_{m=-n}^n D_n^m(\Omega_l) d_{\text{plane}}^m(k) \quad (15b)$$

[0079]

[0080] c , n' , 및 m' 에 대한 합은 수학식 8을 사용하여 삭제될 수 있는데, 이로써 수학식 15a는 수학식 3으로부터의 앰비소닉스 표현에서의 평면 웨이브의 가중치들의 합으로 간략화될 수 있다. 이에 따라, 에일리어싱 신호 및 잡음 신호가 무시되면, 차수(N)의 평면 웨이브의 이론적 계수들은 마이크로폰 어레이 레코딩으로부터 완전히 재구성될 수 있다.

[0081] 잡음 신호의 최종 가중치($w_{\text{noise}}(k)$)는 수학식 14a로부터 그리고 수학식 12b로부터의 $P_{\text{noise}}(\Omega_c, k)$ 만을 사용하여 다음의 수학식에 의해 주어진다.

수학식 16

$$w_{\text{noise}}(k) = \sum_{l=1}^L \sum_{n=0}^N \sum_{m=-n}^n D_n^m(\Omega_l) \times \sum_{c=1}^C \frac{Y_n^m(\Omega_c)^\dagger P_{\text{noise}}(\Omega_c, k)}{b_n(kR)}$$

[0082]

[0083] 수학식 14a에서 수학식 12b로부터의 $P_{\text{alias}}(\Omega_c, kR)$ 의 항을 대체하고 다른 압력 신호들을 무시하는 것은 결과적으로 다음의 수학식을 초래한다.

수학식 17

$$w_{\text{alias}}(k) = \sum_{l=1}^L \sum_{n=0}^N \sum_{m=-n}^n D_n^m(\Omega_l) \times \sum_{n'=N+1}^{\infty} \sum_{m'=-n'}^{n'} Y_{n'}^{m'}(\Omega_s)^* \frac{b_{n'}(kR)}{b_n(kR)} \sum_{c=1}^C Y_n^m(\Omega_c)^\dagger Y_{n'}^{m'}(\Omega_c) P_0(k)$$

[0084]

[0085] 결과로서 생긴 에일리어싱 가중치($w_{\text{alias}}(k)$)는 인덱스(n')가 N 보다 크기 때문에 수학식 8로부터의 직교 정규 조건에 의해 간략화될 수 없다.

[0086] 에일리어스 가중치의 시뮬레이션은 충분한 정확도로 캡슐 신호들을 표현하는 앰비소닉스 차수를 요구한다. 전술한 Moreau/Daniel/Bertet 논문의 섹션 2.2.2의 수학식 14에서, 앰비소닉스 사운드 필드 재구성을 위한 절단 오차(truncation error)의 분석이 주어진다. 다음의 수학식에 대해 설명한다.

수학식 18

$$N_{\text{opt}} = \lceil kR \rceil$$

[0087]

[0088] 사운드 필드의 합리적인 정확도가 획득될 수 있고, 여기서 ' $\lceil \cdot \rceil$ '는 가장 가까운 정수로의 올림(rounding-

up)을 나타낸다. 이러한 정확도는 시뮬레이션의 주파수 상한 f_{\max} 에 사용된다. 따라서, 다음의 수학적식의 앰비소닉스 차수는 각각의 웨이브 번호의 에일리어싱 압력의 시뮬레이션에 사용된다.

수학적식 19

$$N_{\max} = \left\lceil \frac{2\pi f_{\max} R}{c_{\text{sound}}} \right\rceil$$

이는 결과적으로 주파수 상한에서의 정확도가 수용 가능하고, 정확도는 심지어 저주파수의 경우에도 증가한다.

구면 마이크로폰 어레이 프로세싱 - 라우드 스피커 가중치의 분석

도 1은 강체구 상에 32개의 캡슐이 있는 마이크로폰 어레이에 대해 방향($\Omega_s = [0,0]^T$)으로부터의 평면 웨이브에 대한 최종 라우드 스피커 가중치로부터의 가중 컴포넌트 a) $w_{\text{ref}}(k)$, b) $w_{\text{noise}}(k)$ 및 c) $w_{\text{alias}}(k)$ 의 전력을 도시한다(전술된 Agmon/Rafaely 논문으로부터 아이겐마이크(Eigenmike)가 시뮬레이션에 사용되었음). 마이크로폰 캡슐들은 직교 정규 조건들이 성취될 수 있도록 $R = 4.2$ cm인 구의 표면에 균일하게 분포되어 있다. 이 어레이에 의해 지원되는 최대 앰비소닉스 차수(N)는 4이다. 전술한 M.A. Poletti 논문에 설명된 모드 매칭 프로세싱은 Joerg Fliege, Ulrike Maier, "A Two-Stage Approach for Computing Cubature Formulae for the Sphere", Technical report, 1996, Fachbereich Mathematik, Universitat Dortmund, Germany에 따라 균일하게 분포된 25개의 라우드 스피커 위치들에 대한 디코딩 계수들($D_n^m(\Omega_l)$)을 획득하는데 사용된다. 노드 번호들은 <http://www.mathematik.uni-dortmund.de/lx/research/projects/fliege/nodes/nodes.html>에 도시되어 있다.

기준 가중치($w_{\text{ref}}(k)$)의 전력은 전체 주파수 범위에 대해 일정하다. 최종 잡음 가중치($w_{\text{noise}}(k)$)는 저주파수에서 높은 전력을 보여주며, 높은 주파수에서 감소한다. 잡음 신호 또는 전력은 20dB의 분산을 갖는(즉, 평면 웨이브의 전력보다 20dB 낮음) 정상 분포된 언바이어싱된 의사-랜덤 잡음에 의해 시뮬레이션된다. 에일리어싱 잡음($w_{\text{alias}}(k)$)은 저주파수에서 무시될 수 있지만, 주파수가 상승함에 따라 증가하여, 10kHz 초파에서 기준 전력을 초과한다. 에일리어싱 전력 곡선의 기울기는 평면 웨이브 방향에 종속한다. 그러나, 평균적인 추세는 모든 방향들에 대해 일치한다.

2개의 오차 신호($w_{\text{noise}}(k)$ 및 $w_{\text{alias}}(k)$)는 상이한 주파수 범위에서 기준 가중치를 왜곡한다. 게다가, 오차 신호들은 서로 독립적이다. 그러므로, 2단계 등화 프로세싱이 처리된다. 제1 단계에서, 잡음 신호는 동일한 출원인 및 동일한 발명자에 의해 동일 자로 출원된 내부 참조번호가 PD110039인 유럽 출원에 개시된 방법을 사용하여 보상된다. 제2 단계에서, 전체 신호 전력은 에일리어싱 신호 및 제1 프로세싱 단계를 고려하여 등화된다.

제1 단계에서, 기준 가중치와 왜곡된 기준 가중치 사이의 평균 제곱 오차는 모든 진입 평면 웨이브 방향에 대해 최소화된다. 앰비소닉스 표현의 차수에 의해 공간 대역 제한된 후 $w_{\text{alias}}(k)$ 가 정정될 수 없기 때문에, 에일리어싱 신호로부터의 가중치($w_{\text{alias}}(k)$)는 무시된다. 이는 샘플링되고 대역 제한된 시간 신호로부터 에일리어싱이 제거될 수 없는 시간 도메인 에일리어싱과 등가이다.

제2 단계에서, 재구성된 가중치의 평균 전력은 모든 평면 웨이브 방향에 대해 추정된다. 이하, 재구성된 가중치의 전력을 기준 가중치의 전력에 밸런싱하는 필터가 설명된다. 이러한 필터는 스위트 스팟에서만 전력을 등화한다. 그러나, 에일리어싱 오차는 고주파수를 위한 사운드 필드 표현을 여전히 방해한다.

[0097] 마이크로폰 어레이의 공간 주파수 한도는 공간 에일리어싱 주파수라고 지칭된다. 공간 에일리어싱 주파수는 다음의 수학식에 의해 캡슐들의 거리(W003/061336 A1을 참조)로부터 계산된다

수학식 20

$$f_{\text{alias}} = \frac{c_{\text{sound}}}{2R0.73}$$

[0098]

[0099] 여기서, 반지름(R)이 4.2cm인 아이겐마이크(Eigenmike)의 경우 대략 5594Hz이다.

[0100] 최적화 - 잡음 감소

[0101] 잡음 감소는 내부 참조번호가 PD110039인 전술한 유럽 출원에 설명되는데, 평균 사운드 필드 전력과 트랜스듀서 잡음 사이의 신호 대 잡음비($SNR(k)$)가 추정된다. 추정된 $SNR(k)$ 로부터, 다음의 최적화 필터가 설계될 수 있다.

수학식 21

$$F_n(k) = \frac{|b_n(kR)|^2}{|b_n(kR)|^2 + \frac{(4\pi)^2}{C} SNR(k)}$$

[0102]

[0103] 전달 함수($F_n(k)$)의 파라미터들은 마이크로폰 캡슐들의 개수에 종속하고 웨이브 번호(k)에 대한 신호 대 잡음비에 종속한다. 필터는 앰비소닉스 디코더에 독립적인데, 이는 3차원 앰비소닉스 디코딩 및 지향성 빔 형성에 유효하다는 것을 의미한다. $SNR(k)$ 는 내부 참조번호가 PD110039인 전술한 유럽 출원으로부터 획득될 수 있다. 필터는 저주파수에 대해 앰비소닉스 표현의 차수를 제한하는 고대역 통과 필터이다. 필터의 차단 주파수는 더 높은 $SNR(k)$ 의 경우 감소한다. 20dB의 $SNR(k)$ 에 대한 필터의 전달 함수들($F_n(k)$)은 앰비소닉스 차수가 0 내지 4인 경우에 각각 도 2a 내지 도 2e에 도시되는데, 여기서 전달 함수들은 차단 주파수를 높은 차수로 증가시킴에 따라 각각의 차수(n)에 대해 고대역 통과 특성을 갖는다. 차단 주파수들은 전술한 Moreau/Daniel/Bertet 논문에서 섹션 4.1.2에 전술된 바와 같이 정규화 파라미터(λ)와 함께 감소된다. 이에 따라, 높은 $SNR(k)$ 가 저주파수에 대해 고차 앰비소닉스 계수들을 획득하도록 요구된다.

[0104] 최적화된 가중치($w'(k)$)는 다음의 수학식으로부터 계산된다.

수학식 22

$$\begin{aligned} w'(k) &= \sum_{n=0}^N \sum_{m=-n}^n \sum_{l=1}^L D_n^m(\Omega_l) \\ &\times \frac{F_n(k)}{b_n(kR)} \sum_{c=1}^C Y_n^m(\Omega_c)^\dagger (P_{\text{ref}}(\Omega_c, kR) + P_{\text{alias}}(\Omega_c, kR) + P_{\text{noise}}(\Omega_c, k)) \\ &= w'_{\text{ref}}(k) + w'_{\text{alias}}(k) + w'_{\text{noise}}(k) . \end{aligned}$$

[0105]

[0106] $w'_{\text{noise}}(k)$ 의 최종 평균 전력은 다음의 섹션에서 평가된다.

[0107] 최적화 - 스펙트럼 전력 등화

[0108] 최적화된 가중치($w'(k)$)의 평균 전력은 제곱 크기 기대값으로부터 획득된다. 잡음 가중치($w'_{\text{noise}}(k)$)는 잡음 전력이 수학적식 23a에 도시된 바와 같이 독립적으로 계산될 수 있도록 가중치들($w'_{\text{ref}}(k)$ 및 $w'_{\text{alias}}(k)$)에 공간적으로 상관되지 않는다. 기준 및 에일리어싱 가중치의 전력은 수학적식 23b로부터 유도된다. 수학적식 22, 15a, 및 17의 조합은 결과적으로 수학적식 23c를 도출하는데, 수학적식 22에서 $w'_{\text{noise}}(k)$ 는 무시된다. 제곱 크기의 확장은 수학적식 4를 사용하여 수학적식 23c 및 23d를 간략하게 만든다.

수학적식 23

$$E\{|w'(k)|^2\} = E\{|w'_{\text{ref}}(k) + w'_{\text{alias}}(k)|^2\} + E\{|w'_{\text{noise}}(k)|^2\} \quad (23a)$$

$$E\{|w'_{\text{ref}}(k) + w'_{\text{alias}}(k)|^2\} = \frac{1}{4\pi} \int_{\Omega_s \in S^2} |w'_{\text{ref}}(k) + w'_{\text{alias}}(k)|^2 d\Omega_s \quad (23b)$$

$$= \frac{1}{4\pi} \int_{\Omega_s \in S^2} \left| \sum_{n=0}^N \sum_{m=-n}^n \sum_{l=1}^L D_n^m(\Omega_l) \right. \\ \left. \times \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} Y_{n'}^{m'}(\Omega_s)^* \frac{F_n(k)b_{n'}(kR)}{b_n(kR)} \sum_{c=1}^C Y_n^m(\Omega_c)^{\dagger} Y_{n'}^{m'}(\Omega_c) P_0(k) \right|^2 d\Omega_s \quad (23c)$$

$$= \frac{|P_0(k)|^2}{4\pi} \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} \left| \sum_{n=0}^N \sum_{m=-n}^n \sum_{l=1}^L D_n^m(\Omega_l) \right. \\ \left. \times \frac{F_n(k)b_{n'}(kR)}{b_n(kR)} \sum_{c=1}^C Y_n^m(\Omega_c)^{\dagger} Y_{n'}^{m'}(\Omega_c) \right|^2 \quad (23d)$$

$$E\{|w'_{\text{noise}}(k)|^2\} = \frac{4\pi}{C} \sum_{n=0}^N \sum_{m=-n}^n \frac{|\sum_{l=1}^L D_n^m(\Omega_l)|^2 |P_{\text{noise}}(k)|^2 |F_n(k)|^2}{|b_n(kR)|^2} \quad (23e)$$

[0109]

[0110] 최적화된 오차 가중치($w'_{\text{noise}}(k)$)의 전력은 수학적식 23e에 주어진다. $E\{|w'_{\text{noise}}(k)|^2\}$ 의 유도는 내부 참조번호가 PD110039인 전술한 유럽 출원에 설명되어 있다.

[0111] 결과로서 생긴 전력은 사용된 디코딩 프로세싱에 종속적이다. 그러나, 종래의 3차원 앰비소닉스 디코딩의 경우, 모든 방향이 라우드 스피커 배열에 의해 커버된다고 가정한다. 이 경우, 0보다 큰 차수를 갖는 계수들은 수학적식 23에 주어진 디코딩 계수들($D_n^m(\Omega_l)$)의 합에 의해 삭제된다. 이는 원점에서의 압력이 0차 신호와 등가라는 것을 의미하는데, 이로써 저주파수에서 누락된 고차 계수들이 스위트 스팟에서의 전력을 감소시키지 않는다.

[0112] 이는 특정 방향으로부터의 사운드만이 재구성되기 때문에 앰비소닉스 표현의 빔 형성의 경우 상이하다. 여기에서, $D_n^m(\Omega_l)$ 의 모든 계수들이 원점에서의 전력에 기여하도록 하나의 라우드 스피커가 사용된다. 이에 따라, 저주파수에 대한 경감된 고차 계수들은 고주파수들에 비해 가중치($w'(k)$)의 전력을 변경한다.

[0113] 이는 차수(N)를 변경함으로써 수학적식 24에 주어지는 기준 가중치의 전력에 대해 완벽히 설명될 수 있다.

수학식 24

$$E\{|w_{\text{ref}}(k)|^2\} = \frac{|P_0(k)|^2}{4\pi} \sum_{n=0}^N \sum_{m=-n}^n |\sum_{l=1}^L D_n^m(\Omega_l)|^2$$

[0114]

[0115] 수학식 24의 유도는 내부 참조번호가 PD110039인 전술한 유럽 출원에 제공된다. 전력은 $D_n^m(\Omega_l)$ 의 제곱 크기의 합과 등가이며, 이로써 하나의 라우드 스피커(l)의 경우 차수(N)에 따라 전력이 증가한다.

[0116] 그러나, 앰비소닉스 디코딩의 경우, 모든 라우드 스피커 디코딩 계수들($D_n^m(\Omega_l)$)의 합은 고차 계수들을 제거함으로써 0차 계수들만이 스위치 스피커에서의 전력에 기여한다. 이에 따라, 저주파수에서 누락된 HOA 계수들은 앰비소닉스 디코딩을 위한 것이 아닌 빔 형성을 위한 $w'(k)$ 의 전력을 변경한다.

[0117] 잡음 최적화 필터로부터 획득된 $w'(k)$ 의 평균 전력 컴포넌트들은 종래의 앰비소닉스 디코딩을 위해 도 3에 도시된다. 도 3b는 기준+에일리어스 전력을 도시하고, 도 3c는 잡음 전력을 도시하고, 도 3a는 양자 모두의 합을 도시한다. 잡음 전력은 1kHz의 주파수까지 -35 dB로 감소된다. 1 kHz 초과에서 잡음 전력은 -10dB까지 선형으로 증가한다. 최종 잡음 전력은 8 kHz의 주파수까지 $P_{\text{noise}}(\Omega_c, k) = -20$ dB보다 더 작다. 총 전력은 10 kHz 초과에서 10 dB만큼 상승되는데, 이는 에일리어싱 전력에 의해 일어난다. 10 kHz 초과에서, 마이크로폰 어레이의 HOA 차수는 반지름이 R 인 구의 표면 상의 압력 분포를 충분히 설명하지 않는다. 그 결과, 획득된 앰비소닉스 계수들에 의해 일어나는 평균 전력은 기준 전력보다 크다.

[0118] 도 4는 $L=1$ 인 경우 디코딩 계수들($D_n^m(\Omega_l) = Y_n^m(\Omega_{[0,0]})$)에 대한 $w'(k)$ 의 전력 컴포넌트들을 도시한다. 이는 전술한 Agmon/Rafaely 논문에 도시된 바와 같이 방향($\Omega = [0,0]^T$)으로의 빔 형성으로서 해석될 수 있다. 도 4b는 기준+에일리어스 전력을 도시하고, 도 4c는 잡음 전력을 도시하고, 도 4a는 양자 모두의 합을 도시한다. 전력은 저주파수에서 고주파수로 증가하고, 3kHz 내지 6kHz까지 거의 일정하게 유지되고, 그 후 다시 상당히 증가한다. 3kHz가 도 2e에 도시된 4차 계수들에 대해 대략 $F_n(k)$ 의 차단 주파수이기 때문에 제1 증가는 고차 계수들의 경감에 의해 일어난다. 제2 증가는 앰비소닉스 디코딩에 대해 논의된 바와 같이 공간 에일리어싱 전력에 의해 일어난다.

[0119] 이제, $w'(k)$ 의 평균 전력에 대한 등화 필터가 결정된다. 이러한 필터는 사용된 디코딩 계수들($D_n^m(\Omega_l)$)에 강력히 종속하며, 따라서 이들 디코딩 계수들($D_n^m(\Omega_l)$)이 알려진 경우에만 사용될 수 있다.

[0120] 종래의 앰비소닉스 디코딩의 경우, 수학식 25와 같이 가정될 수 있다.

수학식 25

$$\sum_{l=1}^L D_n^m(\Omega_l) = \delta_n \delta_m$$

[0121]

[0122] 다만, 적용된 앰비소닉스 디코더들이 이러한 가정을 거의 이행할 것이라는 점이 보장되어야 한다.

[0123] 실제 값(real-valued) 등화 필터($F_{\text{EQ}}(k)$)는 수학식 26a에 주어진다. 이는 $w_{\text{ref}}(k)$ 의 기준 전력에 대해 $w'(k)$ 의 평균 전력을 보상한다. 수학식 26b에서 수학식 23e 및 27은 $F_{\text{EQ}}(k)$ 또한 $\text{SNR}(k)$ 의 함수라고 수학식 26b에서 보여주기 위해 사용된다.

수학식 26

$$E\{|w_{\text{ref}}(k)|^2\} = E\{|F_{\text{EQ}}(k)(w'_{\text{ref}}(k) + w'_{\text{alias}}(k))|^2\} + E\{|F_{\text{EQ}}(k)w'_{\text{noise}}(k)|^2\}$$

$$F_{\text{EQ}}(k) = \sqrt{\frac{E\{|w_{\text{ref}}(k)|^2\}}{E\{|w'_{\text{ref}}(k) + w'_{\text{alias}}(k)|^2\} + E\{|w'_{\text{noise}}(k)|^2\}}} \quad (26a)$$

$$= \sqrt{\frac{|P_0(k)|^2 E\{|w_{\text{ref}}(k)|^2\}}{|P_0(k)|^2 E\{|w'_{\text{ref}}(k) + w'_{\text{alias}}(k)|^2\} + \frac{4\pi}{C} \sum_{n=0}^N \sum_{m=-n}^n \frac{[\sum_{l=1}^L D_n^m(\Omega_l)]^2 |P_{\text{noise}}(k)|^2 |F_n(k)|^2}{|b_n(kR)|^2}}}} \quad (26b)$$

$$= \sqrt{\frac{E\{|w_{\text{ref}}(k)|^2\}}{E\{|w'_{\text{ref}}(k) + w'_{\text{alias}}(k)|^2\} + \frac{4\pi}{C} \sum_{n=0}^N \sum_{m=-n}^n \frac{[\sum_{l=1}^L D_n^m(\Omega_l)]^2 |F_n(k)|^2}{|b_n(kR)|^2 \text{SNR}(k)}}}} \quad (26c)$$

[0124]

수학식 27

$$|P_0(k)|^2 E\{|w'(k)|^2\} = E\{|w(k)|^2\}$$

[0125]

[0126] 문제점은 필터($F_{\text{EQ}}(k)$)가 필터($F_n(k)$)에 종속하며, $\text{SNR}(k)$ 의 변경마다 양측 필터 모두가 재설계되어야 한다는 것이다. 에일리어싱 및 기준 오차($E\{|w'_{\text{ref}}(k) + w'_{\text{alias}}(k)|^2\}$)의 전력을 시뮬레이션하는데 사용되는 높은 앰비소닉스 차수로 인해 필터 설계의 계산 복잡도가 높다. 적응형 필터링의 경우, 이러한 복잡도는 주어진 마이크로폰 어레이에 대해 일정한 필터 설계 계수들의 세트를 생성하기 위해 단 한번 계산이 복잡한 프로세싱을 수행함으로써 감소될 수 있다. 수학식 28에서, 이들 필터 계수의 유도가 제공된다.

수학식 28

$$A_{n'n}^{m'} = \sum_{l=1}^L \sum_{m=-n}^n D_n^m(\Omega_l) \times \frac{b_{n'}(kR)}{b_n(kR)} \sum_{c=1}^C Y_n^m(\Omega_c)^\dagger Y_{n'}^{m'}(\Omega_c) \quad (28a)$$

$$E\{|w'_{\text{ref}}(k) + w'_{\text{alias}}(k)|^2\}$$

$$= \frac{1}{4\pi} \sum_{n'=0}^\infty \sum_{m'=-n'}^{n'} \sum_{n=0}^N \sum_{m=-n}^n F_n(k) A_{nn'}^{m'} F_{n'}(k)^* A_{nn''}^{m'}{}^* \quad (28b)$$

$$= \frac{1}{4\pi} \sum_{n=0}^N \sum_{n''=0}^N F_{n''}(k)^* F_n(k) \sum_{n'=0}^\infty \sum_{m'=-n'}^{n'} A_{nn'}^{m'} A_{nn''}^{m'}{}^* = \quad (28c)$$

$$\frac{1}{4\pi} \sum_{n=0}^N \sum_{n''=n}^N \begin{cases} F_{n''}(k)^* F_n(k) \sum_{n'=0}^\infty \sum_{m'=-n'}^{n'} A_{nn'}^{m'} A_{nn''}^{m'}{}^*, & \text{for } n = n'' \\ 2 \text{ real}\{F_{n''}(k)^* F_n(k) \sum_{n'=0}^\infty \sum_{m'=-n'}^{n'} A_{nn'}^{m'} A_{nn''}^{m'}{}^*\}, & \text{else} \end{cases} \quad (28d)$$

[0127]

[0128] 수학식 28d에 설명된 바와 같이, $E\{|w'_{\text{ref}}(k) + w'_{\text{alias}}(k)|^2\}$ 의 매우 복잡한 계산이 0부터 N 까지 n 의 합과 n 부터 N 까지의 n'' 에 대한 종속 합(dependent sum)으로 분리될 수 있다. 이들 합의 각 요

소는 필터($F_n(k)$), 자신의 공역 복소 값, $A_{nm}^{m'}$ 의 곱의 n' 및 m' 에 대한 무한 합, 및 자신의 공역 복소 값의 곱이다. 무한 합은 $n' = N_{max}$ 로 진행되는 유한 합에 의해 근사화된다. 이들 덧셈의 결과들은 n 과 n'' 의 조합마다 일정한 필터 설계 계수들을 제공한다. 이들 계수는 주어진 어레이를 위해 한번 계산되고, 시변 신호 대 잡음비 적응형 필터 설계를 위한 룩업 테이블에 저장될 수 있다.

[0129] 최적화 - 최적화된 앰비소닉스 프로세싱

[0130] 앰비소닉스 마이크로폰 어레이 프로세싱의 실제 구현에서, 최적화된 앰비소닉스 계수들($d_{n, opt}^m(k)$)은 수학식 29로부터 획득된다.

수학식 29

$$d_{n, opt}^m(k) = \frac{F_{EQ}(k)F_n(k)}{b_n(kR)} \sum_{c=1}^C Y_n^m(\Omega_c)^\dagger P(\Omega_c, kR)$$

[0131]

[0132] 이는 캡슐들(C)에 대한 합 및 각각의 차수(n) 및 웨이브 번호(k)에 대한 적응형 전달 함수를 포함한다. 이 합은 구의 표면에 대한 샘플링된 압력 분포를 앰비소닉스 표현으로 변환하며, 광대역(wide-band) 신호들의 경우, 시간 도메인에서 수행될 수 있다. 이러한 프로세싱 단계에서는 시간 도메인 압력 신호들($P(\Omega_c, t)$)을 제1 앰비소닉스 표현($A_n^m(t)$)으로 변환한다.

[0133] 제2 프로세싱 단계에서, 수학식 30의 최적화된 전달 함수는 제1 앰비소닉스 표현($A_n^m(t)$)으로부터 지향성 정보들을 재구성한다.

수학식 30

$$F_{n, array}(k) = \frac{F_{EQ}(k)F_n(k)}{b_n(kR)}$$

[0134]

[0135] 전달 함수($b_n(kR)$)의 역수는 $A_n^m(t)$ 를 지향성 계수들($d_n^m(t)$)로 변환하는데, 여기서 샘플링된 사운드 필드가 구의 표면에 분산되었던 평면 웨이브들의 슈퍼пози션에 의해 생성된다고 가정한다. 계수들($d_n^m(t)$)은 전술한 Rafaely "Plane-wave decomposition ..." 논문의 수학식 14, 섹션 3에 설명된 사운드 필드의 평면 웨이브 분해를 나타내고, 이러한 표현은 기본적으로 앰비소닉스 신호들의 전송에 사용된다. $SNR(k)$ 에 따라, 최적화 전달 함수($F_n(k)$)는 잡음에 의해 커버된 HOA 계수들을 제거하기 위해 고차 계수들의 기여도를 감소시킨다. 재구성 신호의 전력은 알려져 있거나 가정된 디코딩 프로세싱을 위한 필터($F_{EQ}(k)$)에 의해 등화된다.

[0136] 결과적으로, 제2 프로세싱 단계는 설계된 시간 도메인 필터와 $A_n^m(t)$ 의 컨볼루션이다. 종래의 앰비소닉스 디코딩을 위한 최종 최적화 어레이들은 도 5에 도시되고, 빔 형성 디코더 예시를 위한 최종 최적화 어레이 응답들

은 도 6에 도시되어 있다. 도 5 및 도 6에서, 전달 함수 a) 내지 e)는 앰비소닉스 차수 0 내지 4에 각각 대응한다.

[0137]

계수들($A_n^m(t)$)의 프로세싱은 선형 필터링 동작으로서 간주될 수 있는데, 여기서 필터의 전달 함수는 $F_{n,array}(k)$ 에 의해 결정된다. 이는 시간 도메인에서뿐 아니라 주파수 도메인에서 수행될 수 있다. FFT는 전달 함수($F_{n,array}(k)$)에 의한 연속적인 곱셈을 위해 계수들($A_n^m(t)$)을 주파수 도메인으로 변환하는데 사용될 수 있다. 그 곱의 역 FFT는 결과적으로 시간 도메인 계수들($d_n^m(t)$)을 도출한다. 이러한 전달 함수 프로세싱은 오버랩 추가 방법 또는 오버랩 세이프 방법을 사용하는 고속 컨볼루션이라고도 알려져 있다.

[0138]

다른 방법으로, 선형 필터는 FIR 필터에 의해 근사화될 수 있는데, FIR 필터의 계수들은 역 FFT를 이용하여 시간 도메인으로 변환하고, 순환 시프트를 수행하고, 최종 필터 임펄스 응답에 테이퍼링 윈도우를 적용하여 대응 전달 함수를 스무딩함으로써 전달 함수($F_{n,array}(k)$)로부터 계산될 수 있다. 그 후 선형 필터링 프로세스는 전달 함수($F_{n,array}(k)$)의 시간 도메인 계수들과 n 와 m 의 각 조합에 대한 계수들($A_n^m(t)$)의 컨볼루션에 의해 시간 도메인에서 수행된다.

[0139]

본 발명의 적응형 블록 기반 앰비소닉스 프로세싱이 도 7에 도시된다. 상부 신호 경로에서, 마이크로폰 캡슐 신호들의 시간 도메인 압력 신호들($P(\Omega_c, t)$)은 수학식 13a를 사용하여 단계 또는 스테이지(71)에서 앰비소닉스 표현($A_n^m(t)$)으로 변환되는데, 이로써 마이크로폰 전달 함수($b_n(kR)$)에 의한 나눗셈이 수행되지 않고($A_n^m(t)$ 가 $d_n^m(t)$ 대신 계산됨), 단계/스테이지(72)에서 대신 수행된다. 그 후, 단계/스테이지(72)는 계수들($d_n^m(t)$)을 획득하기 위해 시간 도메인 또는 주파수 도메인에서 설명된 선형 필터링 동작을 수행함으로써 마이크로폰 어레이 응답이 $A_n^m(t)$ 로부터 제거된다. 제2 프로세싱 경로는 전달 함수($F_{n,array}(k)$)의 자동 적응형 필터 설계에 사용된다. 단계/스테이지(73)는 고려된 기간 동안(즉, 샘플 블록) 시간 대 잡음비($SNR(k)$)의 추정을 수행한다. 유한 개의 이산 웨이브 번호들(k)에 대한 주파수 도메인에 수행된다.

이에 따라 관련 압력 신호들($P(\Omega_c, t)$)은 예를 들어 FFT를 사용하여 주파수 도메인으로 변환되어야 한다.

이 $SNR(k)$ 값은 2개의 전력 신호($|P_{noise}(k)|^2$ 및 $|P_0(k)|^2$)에 의해 특징된다. 잡음 신호의 전력($|P_{noise}(k)|^2$)은 주어진 어레이에 대해 일정하며 캡슐에 의해 생성된 잡음을 나타낸다. 평면 웨이브의 전력($|P_0(k)|^2$)은 압력 신호들($P(\Omega_c, t)$)로부터 추정된다. 이 추정은 내부 참조번호가 PD110039인 전술한 유럽 출원에서의 SNR 추정 섹션에 더 설명되어 있다. 추정된 $SNR(k)$ 로부터, 수학식 30, 26c, 21 및

10을 사용하여 주파수 도메인에서의 단계/스테이지(74)에서 전달 함수($F_{n,array}(k)$)($n \leq N$)이 설계된다. 필터 설계는 위너 필터(Wiener filter) 및 인버스 어레이 응답 또는 역전달 함수($1/b_n(kR)$)를 사용할 수 있다. 그 후, 필터 구현은 스텝/스테이지(72)의 시간 또는 주파수 도메인에서 대응 선형 필터 프로세싱에 적합하다.

[0140]

본 발명의 프로세싱의 결과들은 다음과 같이 설명된다. 이에 따라, 수학식 26c로부터의 등화 필터

($F_{EQ}(k)$)는 기대값($E\{|w'(k)|^2\}$)에 적용된다. 도 3으로부터의 종래 앰비소닉스 디코딩 및 도 4로부터의 빔 형성의 예시들에 대한 $E\{|w'(k)|^2\}$ 의 최종 전력, 기준 전력($E\{|w_{ref}(k)|^2\}$), 및 최종 잡음 전력이 설명된다. 종래의 앰비소닉스 디코더를 위한 최종 전력 스펙트럼은 도 8에 도시되어 있고, 빔 형성 디코더를 위한 최종 전력 스펙트럼은 도 9에 도시되는데, 여기서 곡선 a) 내지 c)는 각각 $|w_{opt}|^2$, $|w_{ref}|^2$, 및 $|w_{noise}|^2$ 을 도시한다.

[0141] 기준 가중치의 전력과 최적화 가중치의 전력은 동일하며, 이로써 최종 가중치는 밸런싱된 파수 스펙트럼을 갖는다. 저주파수에서, 스위트 스팟에서의 최종 신호 대 잡음비는 20 dB의 주어진 $SNR(k)$ 에 비해 종래의 앰비소닉스 디코딩의 경우 증가하며, 빔 형성 디코딩의 경우 감소했다. 고주파수에서, 신호 대 잡음비는 양자 디코더 모두에 대한 주어진 $SNR(k)$ 과 동일하다. 그러나, 빔 형성 디코딩의 경우 고주파수에서의 SNR은 저주파수에서의 SNR보다 크지만, 앰비소닉스 디코더의 경우 고주파수에서의 SNR은 저주파수에서의 SNR보다 작다. 빔 형성 디코더의 저주파수에서 작은 SNR은 누락된 고차 계수들에 의해 일어난다. 도 9에서, 평균 잡음 전력은 도 1에 비해 감소된다. 한편, 신호 전력은 또한 최적화-스펙트럼 전력 등화 섹션에서 논의된 바와 같이 누락된 고차 계수들로 인해 저주파수에서 감소한다. 그 결과, 신호와 잡음 전력 사이의 거리가 작아진다.

[0142] 또한, 최종 SRS는 사용된 디코딩 계수들($D_n^m(\Omega_l)$)에 강하게 종속적이다. 예시적인 빔 패턴은 강한 고차 계수들을 갖는 좁은 빔 패턴이다. 더 넓은 빔을 갖는 빔 패턴을 생성하는 디코딩 계수들은 SNR을 증가시킬 수 있다. 이들 빔은 낮은 차수에서 강한 계수들을 갖는다. 저주파수에서 제한된 차수에 적합하도록 여러 주파수 대역을 위한 상이한 디코딩 계수들을 사용함으로써 더 나은 결과들이 성취될 수 있다.

[0143] 최종 SNR을 최소화하는 최적화된 빔 성형을 위한 다른 방법들이 존재하는데, 디코딩 계수들($D_n^m(\Omega_l)$)은 특정 조향 방향(steering direction)을 위한 수치 최적화에 의해 획득된다. Shefeng, S. Haohai, U.P. Svensson, M. Xiaochuan, J.M. Hovem, "Optimal Modal Beamforming for Spherical Microphone Arrays", IEEE Transactions on Audio, Speech, and language processing, vol.19, no.2, pages 361-371에 제시된 최적의 모달 빔 형성 및 M. Agmon, B. Rafaely, J. Tabrikian, "Maximum Directivity Beamformer for Spherical-Aperture Microphones", 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics WASPAA '09, Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, pages 153-156, 18-21 October 2009, New Paltz, NY, USA에 논의된 최대 지향성 빔 형성은 최적화된 빔 형성을 위한 2개의 예시이다.

[0144] 예시적인 앰비소닉스 디코더는 모드 매칭 프로세싱을 사용하고, 각각의 라우드 스피커 가중치는 빔 형성 예에서 사용되는 디코딩 계수들로부터 계산된다. 라우드 스피커들이 구의 표면에 균일하게 분포되기 때문에 Ω_c 에서 라우드 스피커를 위한 디코딩 계수들은 $D_n^m(\Omega_l) = Y_n^m(\Omega_{\Omega_c})$ 에 의해 정의된다. 라우드 스피커 신호들은 빔 형성 디코더 예시에 대해 동일한 SNR을 갖는다. 그러나, 한편, 원점에서의 라우드 스피커 신호들의 슈퍼포지션은 매우 우수한 SNR을 초래한다. 다른 한편으로, 청취 위치가 스위트 스팟 외부로 이동하면 SNR은 낮아진다.

[0145] 그 결과들에 따르면, 전술된 최적화는 종래의 앰비소닉스 디코더를 위해 원점에서 SNR이 증가된 밸런싱 주파수 스펙트럼을 생성하는데, 즉 본 발명의 시변 적응형 필터 설계는 앰비소닉스 레코딩에 유리하다. 본 발명의 프로세싱은 시간에 따라 레코딩의 SNR이 일정하도록 가정하면 시변 필터를 설계하는데 사용될 수도 있다.

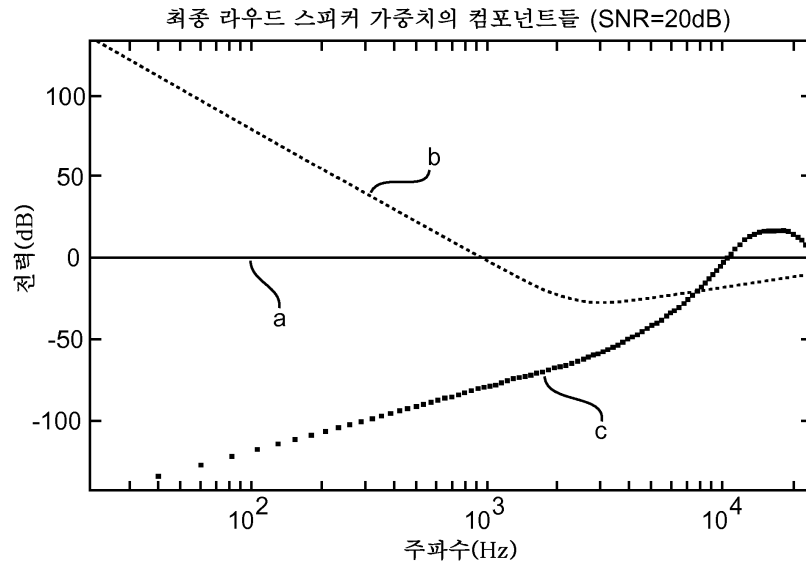
[0146] 빔 형성 디코더들의 경우, 본 발명의 프로세싱은 저주파수에서 낮은 SNR의 단점을 갖는 최종 주파수 스펙트럼을 밸런싱할 수 있다. SNR은 폭이 넓은 빔들을 생성하는 적절한 디코딩 계수들을 선택하거나, 상이한 주파수 서브-대역의 앰비소닉스 차수에 대해 빔 폭을 적응시킴으로써 증가될 수 있다.

[0147]

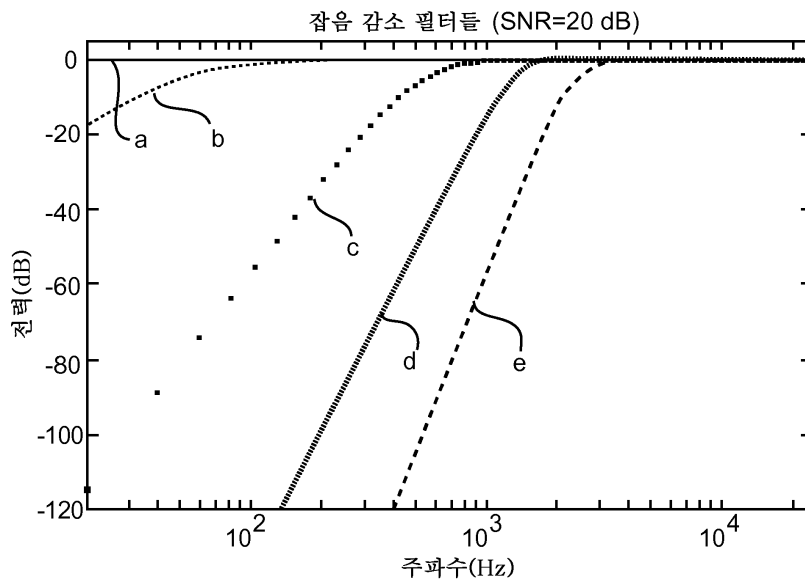
본 발명은 구면 고조파 표현에서 모든 구면 마이크로폰 레코딩에 적용될 수 있으며, 여기서 원점에서의 재생된 스펙트럼 전력은 구면 고조파 계수들을 에일리어싱 또는 누락시킴으로써 언밸런싱된다.

도면

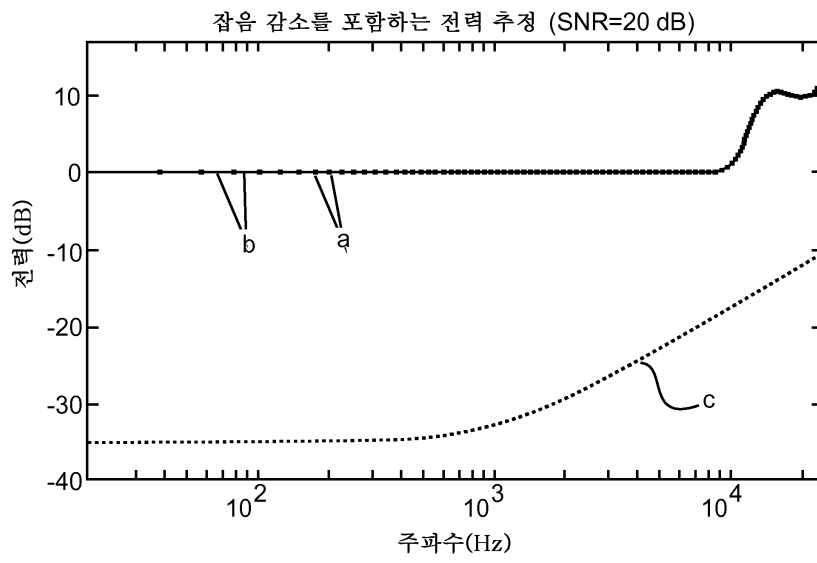
도면1



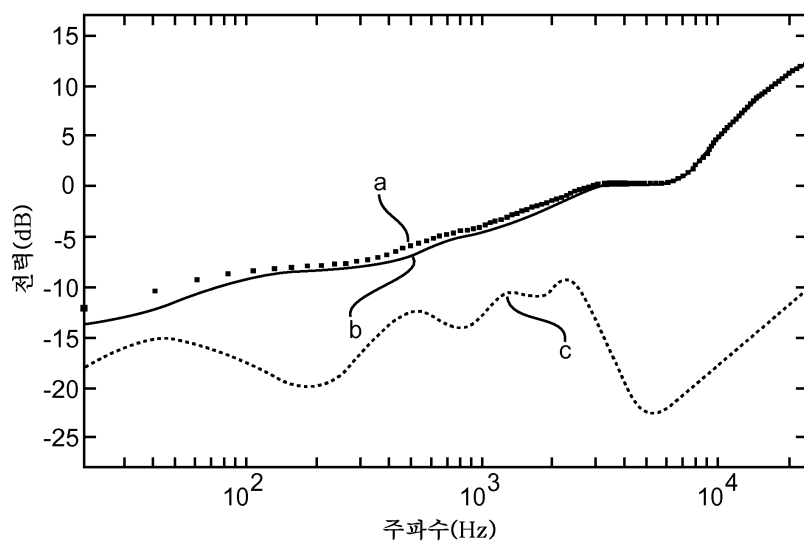
도면2



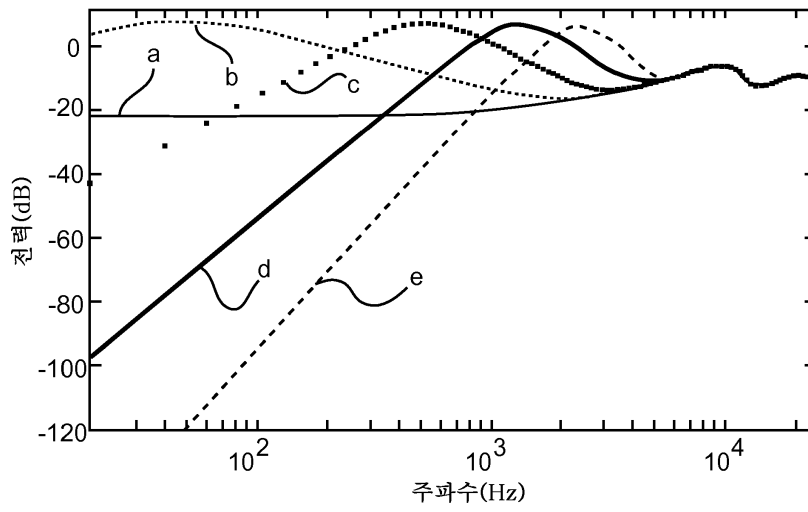
도면3



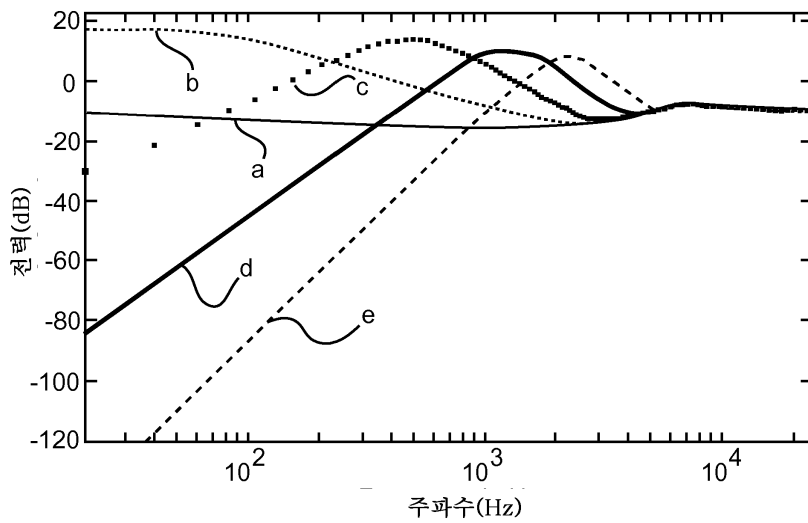
도면4



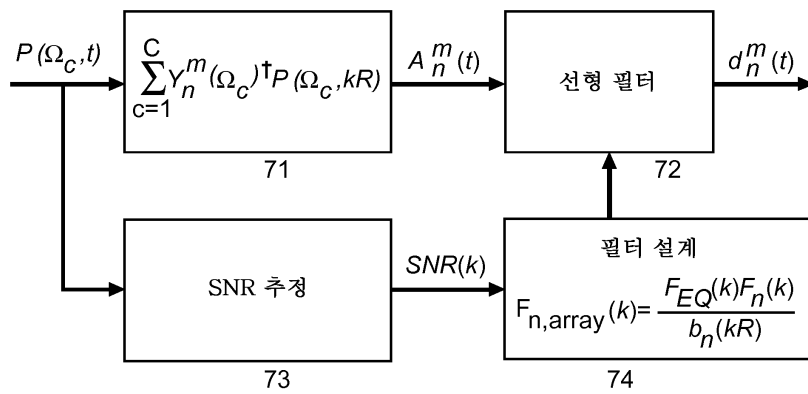
도면5



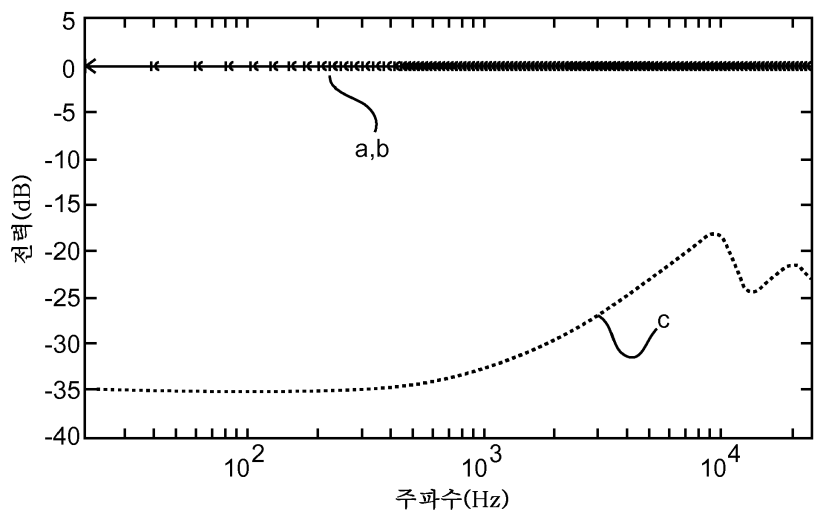
도면6



도면7



도면8



도면9

