



등록특허 10-2460817



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2022년10월31일  
(11) 등록번호 10-2460817  
(24) 등록일자 2022년10월26일

- (51) 국제특허분류(Int. Cl.)  
*H04S 3/00* (2006.01) *G10L 19/008* (2014.01)
- (52) CPC특허분류  
*H04S 3/008* (2020.05)  
*G10L 19/008* (2020.08)
- (21) 출원번호 10-2021-7034751(분할)
- (22) 출원일자(국제) 2014년11월18일  
심사청구일자 2021년11월25일
- (85) 번역문제출일자 2021년10월26일
- (65) 공개번호 10-2021-0132744
- (43) 공개일자 2021년11월04일
- (62) 원출원 특허 10-2016-7014251  
원출원일자(국제) 2014년11월18일  
심사청구일자 2019년11월12일
- (86) 국제출원번호 PCT/EP2014/074903
- (87) 국제공개번호 WO 2015/078732  
국제공개일자 2015년06월04일
- (30) 우선권주장  
13306629.0 2013년11월28일  
유럽특허청(EPO)(EP)
- (56) 선행기술조사문현  
JP2013507796 A  
W02013171083 A1

(73) 특허권자  
돌비 인터네셔널 에이비  
네덜란드 1097디피 암스테르담 앱알. 트레우블란  
7  
(72) 발명자  
크롭, 홀거  
독일, 베데마르크 30900, 크로이츠하이스터 31  
아벨링, 스테판  
독일, 슈바름슈테트 29690, 춤 키에세 11  
(74) 대리인  
양영준, 백만기

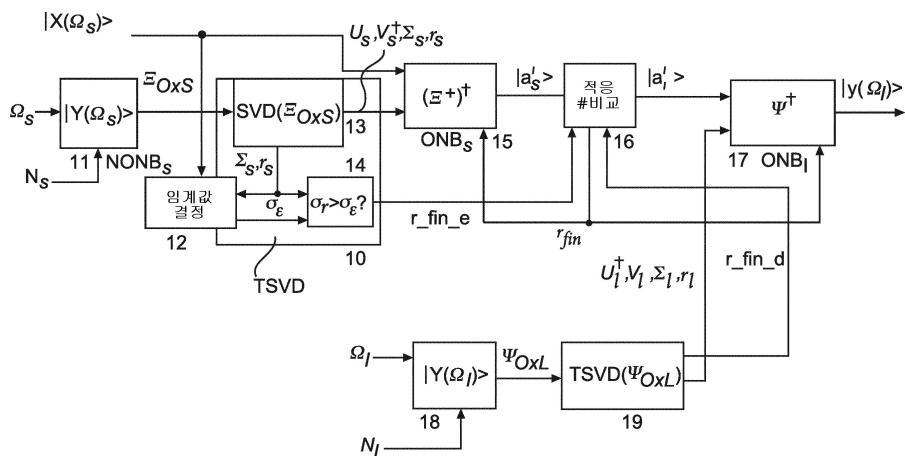
전체 청구항 수 : 총 9 항

심사관 : 권영학

(54) 발명의 명칭 특이 값 분해를 사용하여 고차 앰비소닉스 인코딩 및 디코딩하기 위한 방법 및 장치

**(57) 요약**

특이 값 분해를 이용한 HOA 신호들의 인코딩 및 디코딩은, 음원 방향 값들 및 앰비소닉스 차수에 기초하여 구현  
(뒷면에 계속)

**대 표 도**

고조파의 대응하는 케트 벡터들  $(|Y(\Omega_s)\rangle)$  및 인코더 모드 행렬  $(\Xi_{Oxs})$  을 형성(11)하는 것을 포함한다. 오디오 입력 신호  $(|x(\Omega_s)\rangle)$ 로부터 특이 임계치  $(\sigma_\varepsilon)$  가 결정된다. 인코더 모드 행렬에 대해서 특이 값 분해(13)는 임계값과 비교되어, 최종 인코더 모드 행렬 등급  $(r_{fin_e})$  으로 이어지는 관련된 특이 값들을 얻기 위하여 실행된다. 스피커들의 방향 값들  $(\Omega_i)$  및 디코더 앰비소닉스 차수  $(N_i)$ 에 기초하여, 대응하는 케트 벡터들  $(|Y(\Omega_i)\rangle)$  및 디코더 모드 행렬  $(\Psi_{\chi L})$  이 형성된다(18). 디코더 모드 행렬에 대해서 특이 값 분해(19)가 실행되어, 최종 디코더 모드 행렬 등급  $(r_{fin_d})$  을 제공한다. 최종 인코더 및 디코더 모드 행렬 등급들로부터, 최종 모드 행렬 등급이 결정되고, 이 최종 모드 행렬 등급과 인코더 측의 특이 값 분해로부터, 인코더 모드 행렬  $(\Xi_{Oxs})$  의 수반 의사 역행렬  $(\Xi^+)^{\dagger}$  및 앰비소닉스 케트 벡터  $(|\alpha'_s\rangle)$  가 계산된다. 앰비소닉스 케트 벡터의 성분들의 개수는 적응된 앰비소닉스 케트 벡터  $(|\alpha'_s\rangle)$  를 제공하기 위해 최종 모드 행렬 등급에 따라 감소(16)된다. 적응된 앰비소닉스 케트 벡터, 디코더 측면 특이 값 분해의 출력값들 및 최종 모드 행렬 등급으로부터, 수반 디코더 모드 행렬  $(\Psi^{\dagger})$  이 계산되고, 결과적으로 모든 스피커들에 대한 출력 신호들의 케트 벡터  $(|y(\Omega_i)\rangle)$  를 얻는다.

## (52) CPC특허분류

H04S 2420/11 (2013.01)

## 명세서

### 청구범위

#### 청구항 1

고차 앰비소닉스(Higher Order Ambisonics; HOA) 디코딩을 위한 방법으로서,

라우드스피커들에 대한 구면 고조파들(spherical harmonics)의 상태를 기술하는 벡터들에 관한 정보를 수신하는 단계;

상기 구면 고조파들의 상태를 기술하는 벡터들을 결정하는 단계 - 상기 벡터들은 특이값 분해(Singular Value Decomposition)에 기초하여 결정되었고, 상기 벡터들은 상기 벡터들에 관련된 정보의 행렬에 기초함 -; 및

상기 구면 고조파들의 상태를 기술하는 벡터들에 기초하여, 벡터 기반 신호들의 결과적인 HOA 표현을 결정하는 단계

를 포함하고,

상기 벡터들에 관련된 정보의 행렬은 음원들(sound sources)의 방향에 기초하여 적응되었고, 상기 행렬은 상기 벡터들에 관련된 다수의 선형 독립 열들 및 행들을 제공하는 등급(rank)에 기초하는

고차 앰비소닉스 디코딩을 위한 방법.

#### 청구항 2

제1항에 있어서, 라우드스피커들의 방향 값들( $\Omega_i$ ) 및 디코더 앰비소닉스 차수( $N_i$ )에 관한 정보를 수신하는 단계, 및

상기 라우드스피커들의 방향 값들( $\Omega_i$ ) 및 상기 디코더 앰비소닉스 차수( $N_i$ )에 기초하여 상기 방향 값들( $\Omega_i$ )에

대응하는 방향들에 위치하는 라우드스피커들에 대한 벡터 및 디코더 모드 행렬( $\Psi_{o\&L}$ )을 결정하는 단계

를 더 포함하는 고차 앰비소닉스 디코딩을 위한 방법.

#### 청구항 3

제2항에 있어서,

상기 디코더 모드 행렬( $\Psi_{o\&L}$ )의 특이값 분해에 기초하여 두 개의 대응하는 디코더 단위 행렬들( $U_i^\dagger, V_i$ ) 및 상

기 디코더 모드 행렬( $\Psi_{o\&L}$ )의 최종 등급( $r_{final}$ ) 및 특이값들을 포함하는 디코더 대각 행렬( $\Sigma_i$ )을 결정하는 단계

를 더 포함하는 고차 앰비소닉스 디코딩을 위한 방법.

#### 청구항 4

제2항에 있어서, 상기 라우드스피커들에 대한 구면 고조파들의 벡터들( $|Y(\Omega_i)\rangle$ ) 및 상기 디코더 모드 행렬

( $\Psi_{o\&L}$ )은, 선형 연산 및 오디오 입력 신호( $|x(\Omega_s)\rangle$ ) 내의 소스 위치들을 라우드스피커 출력 신호들의 벡터

( $|y(\Omega_i)\rangle$ ) 내의 라우드스피커들의 위치들에 매핑하는 것을 포함하는 대응하는 패닝 함수( $f_i$ )에 기초하는

고차 앰비소닉스 디코딩을 위한 방법.

#### 청구항 5

고차 앰비소닉스(HOA) 디코딩을 위한 장치로서,

라우드스피커들에 대한 구면 고조파들의 상태를 기술하는 벡터들에 관한 정보를 수신하기 위한 수신기; 및

구면 고조파들의 상태를 기술하는 벡터들을 결정하도록 구성되는 프로세서 - 상기 벡터들은 특이값 분해에 기초

하여 결정되었고, 상기 벡터들은 상기 벡터들에 관련된 정보의 행렬에 기초하여, 상기 프로세서는 상기 구면 고조파들의 상태를 기술하는 벡터들에 기초하여 벡터 기반 신호들의 결과적인 HOA 표현을 결정하도록 추가로 구성됨 - ,

를 포함하고,

상기 벡터들에 관련된 정보의 행렬은 음원들의 방향에 기초하여 적응되었고, 상기 행렬은 상기 벡터들에 관련된 다수의 선형 독립 열들 및 행들을 제공하는 등급에 기초하는

고차 앰비소닉스 디코딩을 위한 장치.

## 청구항 6

제5항에 있어서, 상기 프로세서는 라우드스피커들의 방향 값들( $\Omega_i$ ) 및 디코더 앰비소닉스 차수( $N_i$ )에 관한 정보를 수신하고, 상기 라우드스피커들의 방향 값들( $\Omega_i$ ) 및 상기 디코더 앰비소닉스 차수( $N_i$ )에 기초하여 상기 방향 값들( $\Omega_i$ )에 대응하는 방향들에 위치하는 라우드스피커들에 대한 벡터 및 디코더 모드 행렬( $\Psi_{o_xL}$ )을 결정하도록 더 구성되는 고차 앰비소닉스 디코딩을 위한 장치.

## 청구항 7

제6항에 있어서, 상기 프로세서는 상기 디코더 모드 행렬( $\Psi_{o_xL}$ )의 특이값 분해에 기초하여 두 개의 대응하는 디코더 단위 행렬들( $U_i^\dagger, V_i$ ) 및 상기 디코더 모드 행렬( $\Psi_{o_xL}$ )의 최종 등급( $r_{fin_d}$ ) 및 특이 값들을 포함하는 디코더 대각 행렬( $\Sigma_i$ )을 결정하도록 더 구성되는 고차 앰비소닉스 디코딩을 위한 장치.

## 청구항 8

제6항에 있어서, 상기 라우드스피커들에 대한 구면 고조파들의 벡터들( $|Y(\Omega_i)$ ) 및 상기 디코더 모드 행렬( $\Psi_{o_xL}$ )은, 선형 연산 및 오디오 입력 신호( $|x(\Omega_s)$ ) 내의 소스 위치들을 라우드스피커 출력 신호들의 벡터( $|y(\Omega_i)$ ) 내의 라우드스피커들의 위치들에 매핑하는 것을 포함하는 대응하는 패닝 함수( $f_i$ )에 기초하는 고차 앰비소닉스 디코딩을 위한 장치.

## 청구항 9

컴퓨터 상에서 수행될 때, 제1항에 따른 방법을 수행하는 명령어들을 포함하는, 비일시적 컴퓨터 관독가능 매체 상에 저장되는 컴퓨터 프로그램.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 특이 값 분해(Singular Value Decomposition)를 사용하여 고차 앰비소닉스(Higher Order Ambisonics) 인코딩 및 디코딩하기 위한 방법 및 장치에 관한 것이다.

### 배경 기술

[0002] 고차 앰비소닉스(HOA)는 3차원 사운드를 나타낸다. 다른 기술은 파면 합성(wave field synthesis)(WFS) 또는 22.2와 같은 채널 기반 접근 방식이다. 채널 기반 방법과는 달리, HOA 표현은 특정 스피커 셋업과는 독립적인 장점을 제공한다. 그러나, 이러한 유연성은 특정한 스피커 셋업에서 HOA 표현의 재생을 위해 요구되는 디코딩 처리가 따른다. 요구되는 스피커들의 개수가 일반적으로 매우 큰 WFS 방식에 비해, HOA는 단지 몇 개의 스피커로 구성된 셋업을 위하여 렌더링될 수도 있다. HOA의 또 다른 장점은, 동일한 표현이 헤드폰에 대해 입체 렌더링 하기 위해 어떠한 변경 없이 사용될 수도 있다는 것이다.

[0003] HOA는 절단된 구면 고조파(SH) 팽창에 의해 복합 고조파 평면파 진폭의 공간 밀도의 표현에 기초한다. 각각의 팽창 계수는 시간 도메인 함수로 동일하게 표현될 수 있는 각주파수(angular frequency)의 함수이다. 따라서, 일반성의 손실 없이, 완전한 HOA 음장 표현은 실제로  $O$  시간 도메인 함수를 구성하는 것으로 가정할 수 있으며, 여기서,  $O$  은 팽창 계수들의 개수를 나타낸다.

[0004] 이러한 시간 도메인 함수들은 HOA 계수 시퀀스들 또는 다음의 HOA 채널들로서 동일하게 지칭될 것이다. HOA 표현은 HOA 계수를 포함하는 HOA 데이터 프레임들의 시간적 시퀀스로서 표현될 수 있다. HOA 표현의 공간 해상도는 확장의 증가하는 최대 차수  $N$  으로 향상된다. 3D의 경우에는 팽창 계수들의 개수  $O$  는 차수  $N$  으로 2차식으로 증가하는데, 특히,  $O = (N + 1)^2$  이다.

### 복소 벡터 공간

[0006] 앰비소닉스는 복소 함수들을 처리해야 한다. 따라서, 복소 벡터 공간들에 기초한 표기법이 도입된다. 이는 3차원 'xyz' 좌표계로부터 공지된 진정한 기하학적 벡터들을 표현하지 않는 추상적인 복소 벡터들을 연산한다. 대신에, 각각의 복소 벡터는 물리계에서 있을 수 있는 상태를 설명하고, d 성분들  $x_i$  을 갖는 d-차원 공간에서 열 벡터들에 의해 형성되며, 디랙(Dirac)에 따라 이들 열-지향적인 벡터들은  $|x\rangle$  로서 표시되는 케트 벡터들이라 불리운다. d-차원 공간에 있어서, 임의의  $|x\rangle$  는 성분들  $x_i$  및 d 정규 직교 기준 벡터들  $|e_i\rangle$  에 의해 형성된다.

$$|x\rangle = x_1 |e_1\rangle + x_2 |e_2\rangle + \dots + x_d |e_d\rangle = \sum_{i=1}^d x_i |e_i\rangle \quad . \quad (1)$$

[0007] 여기서, d-차원 공간은 정규 'xyz' 3D 공간이 아니다.

[0009] 케트 벡터의 공액 복소는 브라 벡터  $|x\rangle^* = \langle x|$  라 불리운다. 브라 벡터들은 행-기반 디스크립션(row-based description)을 나타내고, 원래의 케트 공간, 브라 공간의 이중 공간을 형성한다.

[0010] 이러한 디랙 표기법(Dirac notation)은 앰비소닉스 관련된 오디오 시스템을 위한 다음 설명들에 사용될 것이다.

[0011] 내적(inner product)은 복소 스칼라 값을 초래하는 동일한 차원의 브라 및 케트 벡터에서 구축될 수 있다. 랜덤 벡터  $|x\rangle$  가 정규 직교 벡터 기준에서 자신의 성분들에 의해 설명되는 경우, 특정 베이스, 예를 들어,  $|e_i\rangle$  로의  $|x\rangle$  의 음향 방사에 대한 특정 성분은 내적으로 주어진다.

$$x_i = \langle x | e_i \rangle = \langle x | e_i \rangle \quad . \quad (2)$$

[0013] 두 개의 바 대신에 오직 하나의 바가 브라와 케트 벡터 사이에서 고려된다.

[0014] 동일한 기준에서 서로 다른 벡터들  $|x\rangle$  및  $|y\rangle$  에 있어서, 내적은 브라  $|x\rangle$  를  $|y\rangle$  의 케트로 곱하여 얻어진다,

$$\langle x | y \rangle = \sum_{i=1}^d \langle x | e_i \rangle \cdot \sum_{j=1}^d y_j |e_j \rangle = \sum_{i,j=1}^d x_i^* y_j \langle e_i | e_j \rangle = \sum_{i,j=1}^d x_i^* y_j = \sum_{i,j=1}^d y_j^* x_i \quad . \quad (3)$$

[0016] 차원  $m \times 1$ 의 케트와 차원  $1 \times n$ 의 브라 벡터가 외적에 의해 곱해진 경우,  $m$  행들 및  $n$  열들을 갖는 행렬  $A$ 이 얻어진다.

$$A = |x\rangle \langle y| \quad . \quad (4)$$

[0017]

## [0018] 앰비소닉스 행렬들

앰비소닉스-기반 디스크립션은 시변 행렬들로 완벽한 음장을 매핑하는 데 필요한 종속성을 고려한다. 고차 앰비소닉스(HOA) 인코딩 또는 디코딩 행렬들에 있어서, 행들(열들)의 개수는 음원 또는 사운드 싱크로부터 특정 방향들에 관련된다. 인코더측에서,  $S$  음원들의 변형 개수가 고려되며, 여기서,  $s = 1, \dots, S$  이다. 각각의 음원들 ( $s$ )은 원점으로부터의 개별 거리  $r_s$ , 개별 방향  $\Omega_s = (\theta_s, \phi_s)$  을 가질 수 있으며, 여기서,  $\theta_s$  는  $z$ -축으로부터 시작하는 경사 각도를 설명하고,  $\phi_s$  는  $x$ -축으로부터 시작하는 방위 각도를 설명한다. 대응하는 시간 의존 신호  $x_s(t)$  는 개별 시간 동작(behaviour)을 갖는다.

[0020] 단순화를 위해서, 방향성의 부분만이 고려된다(반경 의존성은 베셀 함수들에 의해 설명될 것이다). 이후에, 특정 방향  $\Omega_s$  가 열 벡터  $|Y_n^m(\Omega_s)\rangle$ 에 의해 설명되고, 여기서,  $n$ 은 앰비소닉스 정도를 나타내고,  $m$ 은 앰비소닉스 차수  $N$ 의 인덱스이다. 대응하는 값은  $m = 1, \dots, N$  및  $n = -m, \dots, 0, \dots, m$ 에서 각각 실행된다.

[0021] 일반적으로, 특정 HOA 디스크립션은  $N$ 에 의존하는 2D 또는 3D 경우에 각각의 케트 벡터  $|Y_n^m(\Omega_s)\rangle$ 에 대한 성분들  $O$ 의 개수를 제한한다.

$$O = \begin{cases} 2N+1, & 2D \\ (N+1)^2, & 3D \end{cases} \quad (5)$$

[0022] [0023] 하나 이상의 음원에 있어서, 차수  $n$ 의  $s$  개별 벡터들  $|Y_n^m(\Omega_s)\rangle$ 이 조합되는 경우 모든 방향들이 포함된다. 이는,  $O \times S$  모드 성분들을 포함하는 모드 행렬  $\Xi$ 로 이어지며, 예를 들어,  $\Xi$ 의 각각의 열은 특정 방향을 나타낸다.

$$\Xi = \begin{bmatrix} Y_0^0(\Omega_1) & \cdots & Y_0^0(\Omega_S) \\ Y_1^{-1}(\Omega_1) & \cdots & Y_1^{-1}(\Omega_S) \\ \vdots & \ddots & \vdots \\ Y_N^N(\Omega_1) & \cdots & Y_N^N(\Omega_S) \end{bmatrix} \quad (6)$$

[0024] [0025] 모든 신호 값들은 각각의 개별 음원 신호  $x_s(kT)$ 의 시간 의존성을 고려하는 신호 벡터  $|x(kT)\rangle$ 에서 조합되지만,  $\frac{1}{T}$  공통 샘플율로 샘플링된다.

$$|x(kT)\rangle = \begin{bmatrix} x_1(kT) \\ x_2(kT) \\ \vdots \\ x_s(kT) \end{bmatrix} \quad (7)$$

[0026] [0027] 다음에서, 단순화를 위해,  $|x(kT)\rangle$ 와 같은 시변 신호들에서, 샘플 번호  $k$ 는 더 이상 설명되지 않으며, 즉 무시될 것이다. 이어서,  $|x\rangle$ 는 식(8)에 도시된 것처럼 모드 행렬  $\Xi$  과 곱해진다. 이는 모든 신호 성분들이 동일한 방향  $\Omega_s$ 의 대응하는 열과 선형으로 조합되고, 식(5)에 따라  $O$  앰비소닉스 모드 성분들 또는 계수들

에 의해 케트 벡터  $(|a'_s\rangle)$ 로 이어진다.

$$|a_s\rangle = \Xi|x\rangle \quad . \quad (8)$$

[0028] [0029] 디코더는  $l$  스피커 신호들  $|y\rangle$ 의 전용 번호에 의해 표시되는 음장  $|a_l\rangle$ 을 재생하는 작업을 한다. 이에 따라, 스피커 모드 행렬  $\Psi$ 은 구형 고조파 기반 단위 벡터들  $|Y_n^m(\Omega_l)\rangle$  {식 (6)과 유사}의 L 분리된 열들, 예를 들어, 각각의 스피커 방향에 대해 하나의 케트로 구성된다.

$$\Omega_l: |a_l\rangle = \Psi|y\rangle \quad . \quad (9)$$

[0030] [0031] 모드의 개수가 스피커들의 개수와 동일한 이차 행렬들에 있어서,  $|y\rangle$ 는 반전된 모드 행렬  $\Psi$ 에 의해 결정될 수 있다. 행들 및 열들의 개수가 서로 다른 임의의 행렬인 일반적인 경우에, 스피커 신호들  $|y\rangle$ 은 의사 역행렬 (pseudo inverse)에 의해 결정될 수 있다. 참조: 엠. 에이. 폴레티(M.A. Poletti), "3D 서라운드 사운드 시스템에 대한 구면 고조파 접근 방법", 포럼 어쿠스틱(Forum Acusticum), 부다페스트, 2005. 이때,  $\Psi$ 의 의사 역행렬  $\Psi^+$ 을 사용한다:

$$|y\rangle = \Psi^+|a_l\rangle \quad . \quad (10)$$

[0032] [0033] 인코더 및 디코더측에서 설명되는 음장들은 거의 동일, 즉  $|a_s\rangle \approx |a_l\rangle$  것으로 가정한다. 그러나, 스피커 위치들은 음원 위치들과 서로 다를 수 있는데, 예를 들어, 유한 앰비소닉스 차수에 대해서,  $|x\rangle$ 에 의해 설명되는 실수값의 음원 신호들과  $|y\rangle$ 에 의해 설명되는 스피커 신호들은 서로 다르다. 따라서,  $|y\rangle$ 에 대해서  $|x\rangle$ 을 맵핑하는 패닝 행렬  $G$ 가 사용될 수 있다. 이때, 식 (8) 및 (10)으로부터, 인코더 및 디코더의 체인 연산은 다음과 같다.

$$|y\rangle = G\Psi^+\Xi|x\rangle \quad . \quad (11)$$

#### 선형 합수

[0034] [0035] 다음의 방정식을 간단하게 유지하기 위해, 패닝 행렬은 단락 "발명의 내용" 까지 무시될 것이다. 요구된 기준 벡터들의 개수가 무한한 경우, 별개의 기준에서 연속된 기준으로 변경할 수 있다. 따라서, 합수  $f$ 는 모드 성분들 무한한 개수를 갖는 벡터로서 해석될 수 있다. 이는 케트 벡터들로부터 특정 출력 케트 벡터에 결정적 방식으로 매핑을 수행하기 때문에, 수학적 의미에서 '합수'라고 지칭한다. 이는 합수  $f$ 와 케트  $|x\rangle$  사이의 내적에 의해 설명될 수 있으며, 일반적으로 복소수  $c$ 의 결과를 얻는다.

$$\langle f | (|x\rangle) = \sum_{i=1}^N f_i \cdot x_i = c \quad . \quad (12)$$

[0036] [0037] 합수가 케트 벡터들의 선형 조합을 보존한다면,  $f$ 는 '선형 합수'로 지칭한다.

[0039] 헬미틴(Hermitean) 연산자들에 대한 제한이 존재하는 한, 다음과 같은 특성들이 고려되어야 한다. 헬미틴 연산자들은 항상 다음을 갖는다.

[0040] ● 실제 고유값들.

[0041] ● 서로 다른 고유값들에 대한 직교 고유 함수들의 완전 세트.

[0042] 따라서 모든 함수는 이를 고유 함수들로부터 구축될 수 있다. 참조: 에이취. 보겔(H. Vogel), 씨. 게르센(C. Gerthsen), 에이취. 오. 크네세르(H.O. Kneser), "물리학(Physik)", 스프링거 출판사(Springer Verlag), 1982.

임의의 함수는 복소 상수  $C_n^m$  를 갖는 구형 고조파  $Y_n^m(\theta, \phi)$  의 선형 조합으로서 표현될 수 있다.

$$f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-N}^N C_n^m \cdot Y_n^m(\theta, \phi) \quad (13)$$

$$\langle f(\theta, \phi) | Y_{n'}^{m'}(\theta, \phi) \rangle = \int_0^{2\pi} \int_0^{\pi} f(\theta, \phi)^* Y_{n'}^{m'}(\theta, \phi) \sin \theta d\theta d\phi \quad . \quad (14)$$

[0043] 인덱스들  $n, m$  은 결정론적 방식으로 사용된다. 이것들은 1차원 인덱스  $j$  로 대체되고, 인덱스들  $n', m'$  은 동일한 크기의 인덱스  $i$  로 대체된다. 각각의 하위 공간이 서로 다른  $i, j$  를 갖는 하위 공간에 직교한다는 사실로 인하여, 무한 차원의 공간에서 선형으로 독립적인 정규 직교 단위 벡터들로서 설명될 수 있다.

$$\langle f(\theta, \phi) | Y_i(\theta, \phi) \rangle = \int_0^{2\pi} \int_0^{\pi} \left( \sum_{j=0}^{\infty} C_j Y_j(\theta, \phi) \right) Y_i(\theta, \phi) \sin \theta d\theta d\phi \quad . \quad (15)$$

[0044]  $C_j$  의 상수 값들은 적분 앞에 설정될 수 있다.

$$\langle f(\theta, \phi) | Y_i(\theta, \phi) \rangle = \sum_{j=0}^{\infty} C_j^* \int_0^{2\pi} \int_0^{\pi} Y_j^*(\theta, \phi) Y_i(\theta, \phi) \sin \theta d\theta d\phi \quad . \quad (16)$$

[0045] 하나의 하위 공간(인덱스  $j$ )으로부터 다른 하위 공간(인덱스  $i$ )으로의 맵핑은, 고유함수들  $Y_j$  및  $Y_i$  上호 직교하는 한, 단지 동일한 인덱스들  $i = j$  에 대한 고조파의 적분이 필요하다.

$$\langle f(\theta, \phi) | Y_i(\theta, \phi) \rangle = \sum_{j=0}^{\infty} C_j^* \langle Y_j(\theta, \phi) | Y_i(\theta, \phi) \rangle \quad . \quad (17)$$

[0051] 필수적인 관점은, 계속되는 디스크립션으로부터 브라/케트 표기법으로의 변경이 존재하는 경우, 적분 해법은 구형 고조파의 브라 및 케트 디스크립션들 사이에서 내적의 합으로 대체될 수 있다.

[0052] 일반적으로, 계속되는 기준에 의한 내적은 케트 기반 웨이브 디스크립션  $|x\rangle$  의 이산 표현을 연속되는 표현으로 맵핑하기 위해 사용될 수 있다. 예를 들어,  $x(ra)$  는 위치 기준(예를 들어, 반경)에서 케트 표현이다

$$ra: x(ra) = \langle ra | x \rangle \quad . \quad (18)$$

[0054] 모드 행렬들  $\Psi$  및  $\Xi$  의 서로 다른 종류를 보면, 특이 값 분해(SVD)는 행렬들 중 임의의 종류를 처리하는 데 사용된다.

[0055] 특이 값 분해

[0056] 특이 값 분해(SVD, 참조: G.H. 골롭, Ch.F. 반 론(Golub, Ch.F. van Loan), "행렬 계산", 존스 흑킨스 대학 출판, 제 3 판, 11. 1996년 10월}는  $m$  행들 및  $n$  열들에 의한 임의의 행렬  $A$ 를 3개의 행렬들  $U$ ,  $\Sigma$ , 및  $V^\dagger$ 로 분해할 수 있으며, 식 (19)을 참조한다. 원래의 형태에 있어서, 행렬들  $U$  및  $V^\dagger$ 은 차원  $m \times m$  및  $n \times n$ 의 단위 행렬들이다. 이러한 행렬들은 직교되고, 복소 단위 벡터들  $|u_i\rangle$  및  $|v_i\rangle^\dagger = \langle v_i|$ 을 각각 나타내는 직교 열로부터 구축된다.

[0057] 복소 공간에서 단위 행렬들은 실제 공간에서 직교 행렬들과 동일하며, 즉 열들은 정규 직교 벡터 기준을 제공한다.

$$A = U \Sigma V^\dagger \quad . \quad (19)$$

[0058] 행렬들  $U$  및  $V$ 는 모든 4개의 하위 공간들에 대한 직교 기준을 포함한다.

[0059] ●  $U$ 의 제 1의  $r$  열들 :  $A$ 의 열 공간

[0060] ●  $U$ 의 최종  $m - r$  열들:  $A^\dagger$ 의 널 공간(null space)

[0061] ●  $V$ 의 제 1의  $r$  열들:  $A$ 의 행 공간

[0062] ●  $V$ 의 최종  $n - r$  열들:  $A$ 의 널 공간

[0063] 행렬  $\Sigma$ 은  $A$ 의 동작을 특성화하기 위해 사용될 수 있는 모든 특이 값을 포함한다. 일반적으로,  $\Sigma$ 는  $r$  대각 요소들  $\sigma_i$  까지 갖는  $m \times n$  사각형 대각 행렬이며, 여기서, 등급  $r$ 은  $A(r \leq \min(m, n))$ 의 선형 독립된 열들 및 행들의 개수를 제공한다. 이는 내림 차수로 특이 값을 포함하며, 예를 들어, 식(20) 및 식(21)에 있어서,  $\sigma_1$ 은 최대값을 갖고,  $\sigma_r$ 은 최소값을 갖는다.

[0064] 콤팩트 형태에 있어서, 단지  $r$  특이 값을, 즉  $U$ 의  $r$  열들 및  $V^\dagger$ 의  $r$  행들은 행렬  $A$ 를 재구성하기 위해 요구된다. 행렬들  $U$ ,  $\Sigma$  및  $V^\dagger$ 의 차원들은 원래의 형태와는 다르다. 그러나,  $\Sigma$  행렬들은 항상 이차 형태를 갖는다. 이 때,  $m > n = r$ 에 대해서는,

$$\begin{bmatrix} * * * * \\ * * * \\ * * * \\ A \\ * * * \\ * * * \\ * * * \\ * * * \end{bmatrix}_{m \times n} = \begin{bmatrix} * * * * \\ * * * \\ * * * \\ U \\ * * * \\ * * * \\ * * * \\ * * * \end{bmatrix}_{m \times n} \cdot \begin{bmatrix} \sigma_1 & 0 & \dots \\ 0 & \sigma_2 & 0 & \dots \\ 0 & 0 & \dots & 0 \\ \dots & \dots & 0 & \sigma_r \end{bmatrix}_{n \times n} \begin{bmatrix} * * * * \\ * * * \\ * * * \\ V^\dagger \\ * * * \\ * * * \end{bmatrix}_{n \times n}, \quad (20)$$

[0065]

[0066] 그리고,  $n > m = r$ 에 대해서는,

$$\begin{bmatrix} * * * * * * * \\ * * * * * * \\ A \\ * * * * * * \\ \end{bmatrix}_{m \times n} = \begin{bmatrix} * * * * \\ * * * \\ U \\ * * * \\ \end{bmatrix}_{m \times m} \cdot \begin{bmatrix} \sigma_1 & 0 & \cdot & \cdot \\ 0 & \sigma_2 & 0 & \cdot \\ 0 & 0 & \cdot & 0 \\ \cdot & \cdot & 0 & \sigma_r \\ \end{bmatrix}_{m \times m} \begin{bmatrix} * * * * * * \\ * * * * * \\ V^\dagger \\ * * * * * \\ \end{bmatrix}_{m \times n} .$$
(21)

[0068]

[0069] 따라서, SVD는 낮은 등급 근사치에 의해 매우 효율적으로 구현될 수 있으며, 위에서 언급한 골롭/반 론 텍스트 북을 참조한다. 이러한 근사치는 원래의 행렬을 정확하게 설명하지만,  $r$  등급-1 행렬들까지 포함한다. 디랙 (Dirac) 표기법에 의해, 행렬  $A$ 는  $r$  등급-1 외적에 의해 표현될 수 있다.

$$A = \sum_{i=1}^r \sigma_i |u_i\rangle\langle v_i| . \quad (22)$$

[0070]

[0071] 식(11)에서 인코더 디코더 체인을 볼 때, 행렬  $\Sigma$  과 같은 인코더에 대한 모드 행렬들뿐만 아니라 행렬  $\Psi$ 과 같은 모드 행렬들의 역행렬이 존재하거나, 다른 정교한 디코더 행렬이 고려되어야 한다. 일반적인 행렬  $A$ 에 대해서,  $A$ 의 의사 역행렬  $A^\dagger$ 은 정방 행렬  $\Sigma$ 의 반전과  $U$  및  $V^\dagger$ 의 공액 복소 교환을 수행함으로써 SVD로부터 직접 검사될 수 있으며, 그 결과는 다음과 같이 얻어진다.

$$A^\dagger = V \Sigma^{-1} U^\dagger . \quad (23)$$

[0072]

[0073] 식(22)의 벡터 기반 디스크립션에 대해서, 의사 역행렬  $A^\dagger$ 는  $|u_i\rangle$  및  $\langle v_i|$ 의 공액 교환을 수행함으로써 얻어지는 반면에, 특이 값들  $\sigma_i$ 은 반전되어야 한다. 결과로서 얻어진 의사 역행렬은 다음과 같다.

$$A^\dagger = \sum_{i=1}^r \left(\frac{1}{\sigma_i}\right) |v_i\rangle\langle u_i| . \quad (24)$$

[0074]

[0075] 서로 다른 행렬들의 SVD 기반 분해가 벡터 기반 디스크립션과 조합된다면{참조: 식(8) 및 식(10)}, 인코딩 처리에 대해 얻는다.

$$|\alpha_s\rangle = \sum_{s_i=1}^r \sigma_{s_i} |u_{s_i}\rangle\langle v_{s_i}| \cdot |x\rangle = \sum_{s_i=1}^r \sigma_{s_i} |u_{s_i}\rangle\langle v_{s_i}| x \rangle , \quad (25)$$

[0076]

[0077] 그리고, 디코더에 대해서, 의사 역행렬  $\Psi^+$ 를 고려할 때{식 (24)},

$$|y\rangle = \left(\sum_{l_i=1}^r \left(\frac{1}{\sigma_{l_i}}\right) |v_{l_i}\rangle\langle u_{l_i}| \right) |\alpha_l\rangle . \quad (26)$$

[0078]

[0079] 인코더로부터의 앰비소닉스 음장 디스크립션  $|\alpha_s\rangle$ 이 디코더에 대해  $|\alpha_l\rangle$ 와 거의 동일하고, 차원들  $r_s = r_l = r$ 인 것을 가정하면, 입력 신호  $|x\rangle$  및 출력 신호  $|y\rangle$ 에 대해서 조합된 식은 다음과 같다.

$$|y\rangle = \left(\sum_{l_i=1}^r \left(\frac{1}{\sigma_{l_i}}\right) |v_{l_i}\rangle\langle u_{l_i}| \right) \left(\sum_{s_i=1}^r \sigma_{s_i} |u_{s_i}\rangle\langle v_{s_i}| \right) |x\rangle . \quad (27)$$

## 발명의 내용

## 해결하려는 과제

[0081] 그러나, 인코더 디코더 체인의 이러한 조합된 디스크립션은 이하에 설명되는 몇 가지의 문제점을 갖는다.

### 앰비소닉스 행렬들에 미치는 영향

[0083] 고차 앰비소닉스(HOA) 모드 행렬들  $\Sigma$  및  $\Psi$  은 음원들의 위치, 또는 스피커들{식(6) 참조} 및 앰비소닉스 순서에 의해 직접 영향을 받는다. 형상이 규칙적인 경우, 즉, 음원 또는 스피커 위치들 사이의 상호 각도 거리들이 거의 동일한 경우, 식(27)이 해결될 수 있다.

[0084] 그러나, 실제 응용에 있어서, 이는 종종 사실이 아니다.  $\Sigma$  및  $\Psi$ 의 SVD를 수행하고, 대응하는 행렬  $\Sigma$ 에서 특이 값들을 조사하는 것은 의미가 있는데, 그 이유는  $\Sigma$  및  $\Psi$ 의 수치상 연산을 반영하기 때문이다.  $\Sigma$ 는 실제 특이 값들을 갖는 양의 값의 행렬이다. 그러나 그럼에도 불구하고,  $\Sigma$  특이 값들까지 존재한다 할지라도, 이 값들 사이의 수치상 관계는 음장들의 재생을 위해 매우 중요하며, 그 이유는 디코더 측에서 행렬들의 반전 또는 의사 역행렬을 구축하기 때문이다. 이 연산을 측정하기 위한 적절한 양은  $A$ 의 조건 번호이다. 조건 번호  $\kappa(A)$ 는 가장 작은 특이 값과 가장 큰 특이 값의 비율로서 정의된다.

$$\kappa(A) = \frac{\sigma_r}{\sigma_1} . \quad (28)$$

### 반전 문제점

[0087] 불량 조건 행렬들은 큰  $\kappa(A)$ 를 갖기 때문에 문제가 많다. 반전 또는 의사 역행렬의 경우에, 불량 조건 행

렬은 작은 특이 값들  $\sigma_i$  이 매우 지배적이 되는 문제로 이어진다. P.Ch. 한센에 있어서, "등급-부족 및 개별 불량-제기된 문제: 선형 반전의 수치상 관점들", 산업 및 응용 수학에 대한 협회(SIAM), 1998, 두 가지 기본적인 형태의 문제들은 특이 값들을 감소시키는 방법을 설명함으로써 구별된다(챕터 1.1. 페이지 2~3).

[0088] ● 등급-부족 문제들, 여기서, 행렬들은 크고 작은 특이 값들의 클러스터 사이의 차이(비-점진적 감소),

[0089] ● 별개의 불량-제기된 문제들, 여기서, 행렬들의 모든 특이 값들은 평균적으로 제로로 점진적으로 감소, 즉 특이 값들 스펙트럼에서 차이가 없이 감소한다.

[0090] 인코더 측에서 마이크로폰의 형상뿐만 아니라 디코더 측에서 스피커 형상에 관해서는, 주로 첫 번째로 등급 부족 문제가 발생될 것이다. 그러나, 고객 측에서 모든 가능한 스피커 위치들을 제어하기보다는 기록하는 동안 일부 마이크의 위치들을 변경하기가 쉽다. 특히 디코더 측에서 모드 행렬의 반전 또는 의사 역행렬이 수행되어야 하고, 이는 높은 모드 성분들에 대해 수치상 문제들 및 너무 과중된 값들로 이어진다(상기 언급된 한센 문헌 참조).

### 신호 관련된 종속성

[0092] 반전 문제를 감소시키는 것은 예를 들어, 모드 행렬의 등급을 감소시킴으로써, 즉 최소 특이 값들을 회피함으로

써, 달성될 수 있다. 그런데, 임계값은 최소 가능한 값  $\sigma_r$ 을 위해 사용되어야 한다{참조: 식(20) 및 식(21)}. 이러한 최저 특이 값에 대한 최적 값은 상기 언급한 한센 문헌에 기재되어 있다. 한센은 입력 신호의 특성에 의존하는  $\sigma_{opt} = \frac{1}{\sqrt{SNR}}$  을 제안한다(여기서,  $|x\rangle$ 에 의해 설명됨). 식(27)에서, 입력 신호가 재생에 영향을 주지만, 신호 종속성은 디코더에서 제어될 수 없다는 것을 알 수 있다.

### 비-정규 직교 기준을 갖는 문제점

[0094] HOA 인코더와 HOA 디코더 사이에 전송된 상태 벡터  $|a_s\rangle$ 는 식들(25 및 26)에 따라 서로 다른 기준으로 각각의 시스템에 설명된다. 그러나, 상태는 정규 직교 기준이 사용되는 경우 변하지 않는다.

[0095] 그 다음에, 모드 성분들은 하나에서 다른 기준으로 투영될 수 있다. 그래서, 원칙적으로, 각각의 스피커 셋업 또는 음원 디스크립션은 정규 직교 기준 시스템에 구축되어야 하는데, 그 이유는 그들 기준들 사이에서, 벡터의

표현들의 변화, 예를 들어 앰비소닉스에서, 3D 공간에서 2D 하위 공간으로의 변화가 투영될 수 있기 때문이다.

[0096] 그러나, 기준 벡터들이 거의 선형 종속되는 불량 조건 행렬들을 갖는 셋업들이 종종 존재한다. 그래서, 원칙적으로, 비-정규 직교 기준이 처리되어야 한다. 이는 하나의 하위 공간에서 다른 하위 공간으로 변화를 복잡하게 만들고, 이는 HOA 음장 디스크립션의 서로 다른 스피커 셋업들에 채택되는 경우, 또는 인코더 또는 디코더 측에서 서로 다른 HOA 차수들 및 차원들을 처리하는 것이 바람직한 경우에 필요하다.

[0097] 저밀도 스피커 세트로의 투영에 대한 전형적인 문제는, 음향 에너지가 스피커 근방에서 높다는 것과 이들 스피커들 사이의 거리가 큰 경우 낮다는 것이다. 그래서, 그에 따라 서로 다른 스피커들 사이의 위치는 에너지의 균형을 유지하는 패닝 함수를 필요로 한다.

[0098] 상술한 문제들은 본 발명의 처리에 의해 회피될 수 있고, 청구항1에 개시된 발명에 의해 해결된다. 이 방법을 이용하는 장치는 청구항2에 개시되어 있다.

[0099] 본 발명에 따라, 디코딩 처리에 대한 원래의 기준과 함께 조합된 인코딩 처리에 대한 상호간의 기준은 최저 모드 행렬 등급뿐만 아니라 절단된 특이 값 분해를 고려하여 사용된다. 이중-정규 직교 시스템이 표현되기 때문에, 인코더 및 디코더 행렬들의 곱은 적어도 최저 모드 행렬 등급에 대해 단위 행렬이 유지되는 것이 보장된다.

[0100] 이는 케트 기반 디스크립션을 이중 공간에 기초한 표현으로 변경함으로써 성취되며, 브라 공간은 상호간의 기준 벡터들을 가지며, 여기서, 모든 벡터는 케트의 수반 행렬(adjoint)이다. 모드 행렬들의 의사 역 수반 행렬을 사용하여 실현된다. '수반 행렬'은 복소 공액 교환을 의미한다.

[0101] 따라서, 의사 역 수반 행렬은 인코더 측에서 뿐만 아니라 디코더 수반 행렬에서 이미 사용된다. 처리에 있어서, 정규 직교 상호간 기준 벡터들은 기준 변화에 대해 불변이 되도록 하기 위해 사용된다. 또한, 처리의 종류는 입력 신호 의존 영향을 고려하는 것을 허용하여 정규화 과정에서  $\sigma_i$ 에 대해 노이즈 감소에 대한 최적의 임계값에 이른다.

### 과제의 해결 수단

[0102] 원칙적으로, 본 발명의 방법은 특이 값 분해를 사용하여 고차 앰비소닉스 인코딩 및 디코딩에 적합하며, 상기 방법은,

- 오디오 입력 신호를 수신하는 단계,

[0104] - 음원들의 방향 값들과 상기 오디오 입력 신호의 앰비소닉스 차수에 기초하여, 구형 고조파의 대응하는 케트 벡터들 및 대응하는 인코더 모드 행렬을 형성하는 단계,

[0105] - 상기 인코더 모드 행렬에 대해서 특이 값 분해를 실행하는 단계로서, 두 개의 대응하는 인코더 단위 행렬들 및 특이 값들과 관련된 인코더 모드 행렬 등급을 포함하는 대응하는 인코더 대각 행렬이 출력되는, 특이 값 분해를 실행하는 단계,

[0106] - 상기 오디오 입력 신호, 상기 특이 값들 및 상기 인코더 모드 행렬 등급으로부터 임계값을 결정하는 단계,

[0107] - 상기 특이 값들 중 적어도 하나와 상기 임계값을 비교하여 대응하는 최종 인코더 모드 행렬 등급을 결정하는 단계,

[0108] - 스피커들의 방향 값들 및 디코더 앰비소닉스 차수에 기초하여, 상기 방향 값들 및 대응하는 디코더 모드 행렬에 대응하는 방향들에 위치한 특정 스피커들에 대한 구형 고조파의 대응하는 케트 벡터들을 형성하는 단계,

[0109] - 상기 디코더 모드 행렬에 대해서 특이 값 분해를 실행하는 단계로서, 두 개의 대응하는 디코더 단위 행렬들, 및 특이 값들을 포함하는 하나의 대응하는 디코더 대각 행렬이 출력되고, 상기 디코더 모드 행렬의 대응하는 최종 등급이 결정되는, 특이 값 분해를 실행하는 단계,

[0110] - 상기 최종 인코더 모드 행렬 등급 및 상기 최종 디코더 모드 행렬 등급으로부터 최종 모드 행렬 등급을 결정하는 단계,

[0111] - 상기 인코더 단위 행렬들, 상기 인코더 대각 행렬, 및 상기 최종 모드 행렬 등급으로부터, 앰비소닉스 케트 벡터를 결과로서 얻는, 상기 인코더 모드 행렬의 수반 의사 역행렬을 계산하고,

- [0112] 적응된 앰비소닉스 케트 벡터를 제공하기 위해, 상기 최종 모드 행렬 등급에 따라 상기 앰비소닉스 케트 벡터의 성분들의 개수를 감소시키는 단계,
- [0113] - 상기 적응된 앰비소닉스 케트 벡터, 상기 디코더 단위 행렬들, 상기 디코더 대각 행렬 및 상기 최종 모드 행렬 등급으로부터, 모든 스피커들에 대한 출력 신호들의 케트 벡터를 결과로서 얻는 수반 디코더 모드 행렬을 계산하는 단계를 포함한다.
- [0114] 원칙적으로, 본 발명의 장치는 특이 값 분해를 사용하여 고차 앰비소닉스 인코딩 및 디코딩에 적합하며, 상기 장치는,
- [0115] - 오디오 입력 신호를 수신하도록 적응된 수단,
- [0116] - 음원들의 방향 값들과 상기 오디오 입력 신호의 앰비소닉스 차수에 기초하여, 구형 고조파의 대응하는 케트 벡터들 및 대응하는 인코더 모드 행렬을 형성하도록 적응된 수단,
- [0117] - 상기 인코더 모드 행렬에 대해서 특이 값 분해를 실행하도록 구성된 수단으로서, 두 개의 대응하는 인코더 단위 행렬들 및 특이 값들을 포함하는 대응하는 인코더 대각 행렬과, 관련된 인코더 모드 행렬 등급이 출력되는, 특이 값 분해를 실행하도록 구성된 수단,
- [0118] - 상기 오디오 입력 신호, 상기 특이 값들 및 상기 인코더 모드 행렬 등급으로부터 임계값을 결정하도록 적응된 수단,
- [0119] - 상기 특이 값들 중 적어도 하나와 상기 임계값을 비교하여 대응하는 최종 인코더 모드 행렬 등급을 결정하도록 적응된 수단,
- [0120] - 스피커들의 방향 값들 및 디코더 앰비소닉스 차수에 기초하여, 상기 방향 값들 및 대응하는 디코더 모드 행렬에 대응하는 방향들에 위치한 특정 스피커들에 대한 구형 고조파의 대응하는 케트 벡터들을 형성하도록 적응된 수단,
- [0121] - 상기 디코더 모드 행렬에 대해서 특이 값 분해를 실행하도록 적응된 수단으로서, 두 개의 대응하는 디코더 단위 행렬들, 및 특이 값들을 포함하는 하나의 대응하는 디코더 대각 행렬이 출력되고, 상기 디코더 모드 행렬의 대응하는 최종 등급이 결정되는, 특이 값 분해를 실행하도록 적응된 수단,
- [0122] - 상기 최종 인코더 모드 행렬 등급 및 상기 최종 디코더 모드 행렬 등급으로부터 최종 모드 행렬 등급을 결정하도록 적응된 수단,
- [0123] - 상기 인코더 단위 행렬들, 상기 인코더 대각 행렬, 및 상기 최종 모드 행렬 등급으로부터, 앰비소닉스 케트 벡터를 결과로서 얻는, 상기 인코더 모드 행렬의 수반 의사 역행렬을 계산하고,
- [0124] 적응된 앰비소닉스 케트 벡터를 제공하기 위해, 상기 최종 모드 행렬 등급에 따라 상기 앰비소닉스 케트 벡터의 성분들의 개수를 감소시키도록 적응된 수단,
- [0125] - 상기 적응된 앰비소닉스 케트 벡터, 상기 디코더 단위 행렬들, 상기 디코더 대각 행렬 및 상기 최종 모드 행렬 등급으로부터, 모든 스피커들에 대한 출력 신호들의 케트 벡터를 결과로서 얻는 수반 디코더 모드 행렬을 계산하도록 적응된 수단을 포함한다.
- [0126] 본 발명의 유리한 추가적인 실시예들은 각각의 종속 청구항들에 개시된다.

### 발명의 효과

- [0127] 본 발명에 따라, 디코딩 처리에 대한 원래의 기준과 함께 인코딩 처리에 대한 상호간의 기준은 최저 모드 행렬 등급뿐만 아니라 절단된 특이 값 분해를 고려하여 사용된다. 이중-정규 직교 시스템이 표현되기 때문에, 인코더 및 디코더 행렬들의 곱은 최저 모드 행렬 등급에 대해 적어도 단위 행렬이 유지되는 것이 보장된다.

### 도면의 간단한 설명

- [0128] 도 1은 SVD에 기초하여 HOA 인코더 및 디코더의 블록도.  
 도 2는 선형 함수의 패닝을 포함하는 HOA 인코더 및 디코더의 블록도,  
 도 3은 행렬 패닝을 포함하는 HOA 인코더 및 디코더의 블록도.

$(\sigma_\varepsilon)$  도 4는 임계값 을 결정하기 위한 흐름도.

도 5는 감소된 모드 행렬 등급  $r_{fin_e'}$ 의 경우에 특이 값들의 재계산과  $|a'_s\rangle$ 의 계산을 설명하는 도면.

도 6은 감소된 모드 행렬 등급들  $r_{fin_e}$  및  $r_{fin_d'}$ 의 경우에 특이 값들의 재계산과 패닝을 사용하거나 사용하지 않고 스피커 신호들  $|y(\Omega_l)\rangle$ 의 계산을 설명하는 도면.

### 발명을 실시하기 위한 구체적인 내용

[0129]

본 발명의 예시적인 실시예들은 첨부한 도면들을 참조하여 설명한다.

[0130]

도 1에는 인코더 파트와 디코더 파트 모두를 사용하여 SVD에 기초한 본 발명의 HOA 처리를 위한 블록도가 도시되어 있다. 두 파트 모두는 상호간의 기준 벡터들을 생성하기 위해 SVD를 사용한다. 여기서, 공지된 모드 매칭 해법들에 대한 변경, 즉 식(27)에 관련된 변경 사항이 있다.

[0131]

HOA 인코더

[0132]

상호간의 기준 벡터들과 작업하기 위하여, 케트 기반 디스크립션이 브라 공간에 대해 변경되는데, 여기서, 모든 벡터는 케트의 헬미턴 공액 또는 수반 행렬이다. 이는 모드 행렬들의 의사 반전을 사용하여 실현된다.

[0133]

이때, 식(8)에 따라, (이중) 브라 기반 앰비소닉스 벡터는 (이중) 모드 행렬  $\Xi_d$ 로 재공식화될 수도 있다.

[0134]

$$\Xi_d : \langle a_s | = \langle x | \Xi_d = \langle x | \Xi^+ . \quad (29)$$

[0135]

인코더 측에서 결과로서 얻어진 앰비소닉스 벡터  $\langle a_s |$ 는 지금 브라 시멘틱 상태(bra semantic)이다. 그러나, 통합된 디스크립션(unified description)이 바람직한데, 예를 들어, 케트 시멘틱으로 복귀하는 것이 바람직하다.  $\Xi$ 의 의사 반전 대신에,  $\Xi_d^\dagger$  또는  $\Xi^{+\dagger}$ 의 헬미턴 공액(Hermitean conjugate)이 사용된다.

[0136]

$$|a_s\rangle = \Xi_d^\dagger |x\rangle = \Xi^{+\dagger} |x\rangle . \quad (30)$$

[0137]

식(24)에 따라,

[0138]

$$\Xi^{+\dagger} = \left( \sum_{i=1}^{r_s} \left( \frac{1}{\sigma_{s_i}} \right) |v_{s_i}\rangle \langle u_{s_i}| \right)^\dagger = \sum_{i=1}^{r_s} \left( \frac{1}{\sigma_{s_i}} \right) |u_{s_i}\rangle \langle v_{s_i}| , \quad (31)$$

[0139]

여기서, 모든 특이 값들은 실수이고,  $\sigma_{s_i}$ 의 복소 공액은 무시될 수 있다.

[0140]

이는 앰비소닉스 성분들의 다음 디스크립션으로 이어진다.

[0141]

$$|a_s\rangle = \sum_{i=1}^{r_s} \left( \frac{1}{\sigma_{s_i}} \right) |u_{s_i}\rangle \langle v_{s_i}| |x\rangle . \quad (32)$$

[0142]

음원측에 대한 벡터 기반 디스크립션은  $|a_s\rangle$ 가 역함수  $\sigma_{s_i}$ 에 의존한다는 것을 나타낸다. 이러한 것이 인코더 측에서 실행된다면, 디코더 측에서 대응하는 이중 기준 벡터들로 변경되어야 한다.

[0143]

HOA 디코더

[0144] 디코더가 원래 의사 역함수에 기초하는 경우, 스피커 신호들  $|y\rangle$  을 유도하기 위해 다음 식을 취한다.:

$$|a_l\rangle = \Psi^{+\dagger} |y\rangle , \quad (33)$$

[0145] 예를 들어, 스피커 신호들은 다음과 같다.

$$|y\rangle = (\Psi^{+\dagger})^+ \cdot |a_l\rangle = \Psi^+ \cdot |a_l\rangle . \quad (34)$$

[0146] 식(22)를 고려하면, 디코더 식은 다음 결과를 얻는다.

$$|y\rangle = (\sum_{i=1}^r \sigma_{l_i} |u_{l_i}\rangle \langle v_{l_i}|)^+ |a_l\rangle . \quad (35)$$

[0147] 따라서, 의사 역함수를 구축하는 대신에, 단지 수반 행렬 연산(Adjoint operation)('†'로 표기)이 식(35)에 남게 된다. 이는 보다 적은 산술 연산들이 디코더에서 요구된다는 것을 의미하며, 그 이유는 허수 부분들의 부호를 단지 전환하고, 이러한 전환이 단지 변경된 메모리 액세스의 문제이기 때문이다.

$$|y\rangle = (\sum_{i=1}^r \sigma_{l_i} \cdot |v_{l_i}\rangle \langle u_{l_i}|) |a_l\rangle . \quad (36)$$

[0148] 인코더 및 디코더의 앰비소닉스 표현들이 거의 동일, 즉  $|a_s\rangle = |a_l\rangle$  라고 가정하면, 식(32)에 의해, 완전한 인코더 디코더 체인은 다음 종속성을 얻는다.

$$|y\rangle = \sum_{i=1}^r \left( \frac{\sigma_{l_i}}{\sigma_{s_i}} \right) \cdot |v_{l_i}\rangle \langle u_{l_i}| |u_{s_i}\rangle \langle v_{s_i}| |x\rangle , \quad (37)$$

$$|y\rangle = \sum_{i=1}^r \left( \frac{\sigma_{l_i}}{\sigma_{s_i}} \right) \langle u_{l_i} | u_{s_i} \rangle \cdot |v_{l_i}\rangle \langle v_{s_i}| |x\rangle . \quad (38)$$

[0149] 실제 시나리오에 있어서, 식(11)로부터의 패닝 행렬  $G$  과 유한 앰비소닉스 차수가 고려되어야 한다. 후자는 음장을 설명하기 위해 사용되는 기본 벡터들의 선형 조합들의 개수를 제한한다. 또한, 기본 벡터들의 선형 독립성은 수치 라운딩 에러들 또는 측정 에러들과 같은 부가적인 에러 소스들에 의해 영향을 받는다. 실용적인 관점에서, 이는 수치 등급에 의해 무시될 수 있으며(상기 언급한 한센 문헌, 챕터 3.1 참조), 이에 따라, 모든 기본 벡터들은 특정 허용 오차 내에서 선형으로 독립적이 될 수 있다.

[0150] 노이즈에 대해 보다 강하게 되도록, 입력의 계산된 앰비소닉스 표현과 인코더 케트에 영향을 주는 입력 신호들의 SNR이 고려된다. 그래서, 필요하다면, 즉, 반전되어야 하는 불량 조건 모드 행렬에 대해서,  $\sigma_i$  값은 인코더에서 입력 신호의 SNR에 따라 정규화된다.

### 인코더의 정규화

[0151] 정규화는 서로 다른 방식들, 예를 들어, 절단된 SVD를 통해 임계값을 사용하여 실행될 수 있다. SVD는 내림 차수로  $\sigma_i$  를 제공하는데, 여기서,  $\sigma_i$  는 저레벨 또는 최고 인덱스( $\sigma_r$ 로 표시)를 가지며, 매우 자주 전환하여 노이즈 효과 및 SNR을 초래하는 성분들을 포함한다(참조: 식(20) 및 식(21) 및 상기 언급한 한센 문헌). 따라서, 절단 SVD(TSVD)는 모든  $\sigma_i$  값들을 임계값과 비교하고, 임계값  $(\sigma_\varepsilon)$ 을 초과하는 노이즈 성분들을 무시한다. 이러한 임계값  $(\sigma_\varepsilon)$  은 고정될 수 있거나 입력 신호들의 SNR에 따라 최적으로 수정될 수 있다.

[0152] 하나의 행렬의 트레이스는 모든 대각 행렬 요소들의 합을 의미한다.

[0160] TSVD 블록(도 1 내지 도 3의 10, 20, 30)은 다음 과정을 수행한다.

[0161] ● 모드 행렬 등급  $r$  을 계산한다.

[0162] ● 임계값 이하의 노이즈 성분들을 제거하고 최종 모드 행렬 등급  $r_{fin}$  을 설정한다.

[0163] 이러한 처리는 복소 행렬들  $\Sigma$  및  $\Psi$ 를 다룬다. 그러나, 실수 값  $\sigma_i$  을 정규화하기 위해서는, 이들 행렬들을 직접 사용할 수 없다. 적당한 값은  $\Sigma$ 와 자신의 수반 행렬  $\Sigma^\dagger$  사이의 곱으로부터 얻는다. 그 결과의 행렬은 적절한 특이 값들의 이차 값들과 동일한 실제 대각 고유 값들을 갖는 이차형이다. 행렬  $\Sigma^2$ 의 트레이스에 의해 설명될 수 있는 모든 고유 값들의 합이 고정된 상태로 유지되는 경우, 시스템의 물리적 성질들은 보존된다. 이는 또한 행렬  $\Psi$ 에 적용된다.

$$\text{trace}(\Sigma^2) = \sum_{i=1}^r \sigma_i^2 \quad , \quad (39)$$

[0164]

따라서, 인코더 측(도 1 내지 도 3의 15, 25, 35)에서 블록  $\text{ONB}_S$  또는 디코더 측(도 1 내지 도 3의 19, 29, 39)에서 블록  $\text{ONB}_1$  은 특이 값들을 변경하여, 정규화 이전 및 이후의  $\text{trace}(\Sigma^2)$  가 보존된다(도 5 및 도 6 참조).

[0166] ●  $\sigma_i$  ( $i = 1 \dots r_{fin}$  에 대해서)의 나머지를 변경하여, 원래의 트레이스와 의도되어 절단된 행렬  $\Sigma_t$  을 고정 상태 ( $\text{trace}(\Sigma^2) = \text{trace}(\Sigma_t^2)$ ) 로 유지한다.

[0167]

● 다음 식을 만족하는 상수 값  $\Delta\sigma$  을 계산한다.

$$\sum_{i=1}^r \sigma_i^2 = \sum_{i=1}^{r_{fin}} (\sigma_i + \Delta\sigma)^2 \quad . \quad (40)$$

[0168]

정상적인 값과 특이 값들의 감소된 수의 차이를  $(\Delta E = \text{trace}(\Sigma) - \text{trace}(\Sigma_{r_{fin}}))$  라고 지칭하면, 그 결과의 값은 다음과 같다.

$$\Delta\sigma = \frac{1}{r_{fin}} \left( - \sum_{i=1}^{r_{fin}} \sigma_i + \sqrt{\left( \sum_{i=1}^{r_{fin}} \sigma_i \right)^2 + r_{fin} \Delta E} \right) = \frac{1}{r_{fin}} \left( - \text{trace}(\Sigma)_{r_{fin}} + \sqrt{\text{trace}(\Sigma)_{r_{fin}}^2 + r_{fin} \Delta E} \right) \quad (41)$$

[0170]

● 절단된 행렬  $\Sigma_t$  에 대해서 모든 새로운 특이 값들  $\sigma_{i,t}$  을 다시 계산한다.

[0172]

$$\sigma_{i,t} = \sigma_i + \Delta\sigma \quad . \quad (42)$$

[0173]

부가적으로, 적당한  $|a\rangle$  {식 (30) 또는 식 (33) 참조}에 대한 기준이 대응하는 SVD-관련된  $\{U^\dagger\}$  기준으로 변경되는 경우, 인코더 및 디코더에 대한 단순화가 성취될 수 있으며, 다음 식을 얻는다.

$$|a'\rangle = \sum_{i=1}^{r_{fin}} \langle u_i | \left[ \sum_{i=1}^{r_{fin}} \sigma_{i,t} |u_i\rangle \langle v_i| \right] |a\rangle = \sum_{i=1}^{r_{fin}} \sigma_{i,t} \langle v_i | a \rangle \quad (43)$$

[0174]

(비고):  $\sigma_1$  및  $|a\rangle$  가 추가적인 인코더 또는 디코더 인덱스없이 사용되는 경우, 인코더 측 또는/및 디코더 측

을 참조한다). 이러한 기준은  $|a\rangle$ 의 표준을 유지하기 위해 직교 함수이다. 즉,  $|a\rangle$  대신에 그러한 정규화는 행렬  $\Sigma$  및  $V$ 를 필요로 하지만, 행렬  $U$ 를 더 이상 필요로 하지 않는  $|a'\rangle$ 을 사용할 수 있다.

[0176] ● 등급이 확실하게 감소되는 장점이 있는  $\{U^\dagger\}$  기준에서 감소된 케트  $|a'\rangle$ 를 사용한다.

[0177] 따라서, 본 발명에 있어서, SVD는 정규 직교 기준과 개별 행렬  $\Sigma$  및  $\Psi$ 의 특이 값들을 수행할 뿐만 아니라, 그들 등급들  $r_{fin}$ 을 얻기 위하여 두 측에서 사용될 수 있다.

#### 성분 적용

[0179]  $\Sigma$ 의 음원 등급을 고려하거나 임계값 또는 최종 음원 등급에 대하여 대응하는  $\sigma_s$ 의 일부를 무시함으로써, 성분들의 개수는 감소될 수 있고, 보다 강력한 인코딩 행렬이 제공될 수 있다. 따라서, 디코더 측에서 성분들의 대응하는 개수에 따라 전송된 앰비소닉스 성분들의 개수의 적용이 실행된다. 보통은 앰비소닉스 차수  $O$ 에 의존 한다. 여기서, 인코더 행렬  $\Sigma$ 에 대한 SVD 블록으로부터 얻은 최종 모드 행렬 등급  $(r_{fin_e})$ 과 디코더 행렬  $\Psi$ 에 대한 SVD 블록으로부터 얻은 최종 모드 행렬 등급  $(r_{fin_d})$ 이 고려되어야 한다. 적용#비교 단계/단계(16)에 있어서, 성분들의 개수는 다음과 같이 적용된다.

[0180] ●  $r_{fin_e} = r_{fin_d}$  : 어떤 것도 변하지 않음- 압축 없음,

[0181] ●  $r_{fin_e} < r_{fin_d}$  : 압축, 디코더 행렬  $\Psi^\dagger$ 에서  $r_{fin_e} - r_{fin_d}$  열들을 무시  $\Rightarrow$  인코더 및 디코더 연산들 감소,

[0182] ●  $r_{fin_e} > r_{fin_d}$  : 압축, 즉 전송 이전에 앰비소닉스 상태 벡터의  $r_{fin_e} - r_{fin_d}$  성분들 취소, 즉 압축. 인코더 행렬  $\Sigma$ 에서  $r_{fin_e} - r_{fin_d}$  행들을 무시  $\Rightarrow$  인코더 및 디코더 연산들 감소.

[0183] 그 결과는 인코더 측 및 디코더 측에서 사용되는 최종 모드 행렬 등급  $r_{fin_e}$ 이  $r_{fin_d}$  및  $r_{fin_e}$  중 더 작은 하나이다.

[0184] 따라서, 인코더와 디코더 사이의 양방향 신호가 다른 측의 등급을 상호 교환하기 위해 존재한다면, 등급 차이들을 사용하여 가능한 압축을 개선하고 인코더 및 디코더에서 연산들의 개수를 감소시킬 수 있다.

#### 패닝 함수들 고려

[0186] 패닝 함수들  $f_s, f_t$  또는 패닝 행렬  $G$ 의 사용은, 저밀도 및 불규칙한 스피커 셋업들에 대해 야기되는 에너지 분배에 관련된 문제들로 인하여, 식(11)을 참조하여, 이미 앞서 언급했다. 이러한 문제들은 앰비소닉스에서 정규적으로 사용될 수 있는 제한된 차수를 처리한다(앰비소닉스 행렬들에 대한 영향 내지 비-정규 직교 기준에 따른 문제들에 대해 참조).

[0187] 패닝 행렬  $G$ 에 대한 요구들과 후속되는 인코딩에 관하여, 일부 음원의 음장이 앰비소닉스 상태 벡터  $|a_s\rangle$ 에 의해 표현되는 양호한 상태에 있다고 가정한다. 그러나, 디코더 측에서, 상태가 어떻게 준비되었는지 정확히 알지 못한다. 즉, 시스템의 현재 상태에 대한 완전한 지식이 없다. 따라서, 상호간의 기준은 식(9) 및 식(8) 사이의 내적을 보존하기 위해 취해진다.

[0188] 인코더 측에서 미리 의사 반전을 사용하는 것은 다음과 같은 장점을 제공한다.

- [0189] ● 상호간의 기준  $(\langle x^i | x_j \rangle = \delta_j^i)$  의 사용은 인코더와 디코더 기준 사이 양방향 직교성을 만족한다.

- [0190] ● 인코딩/디코딩 체인에서 보다 작은 연산들의 개수,

- [0191] ● SNR 작용에 관한 수치적 관점 개선,

- [0192] ● 단지 선형으로만 독립된 행렬들 대신에 변경된 모드 행렬들에서 정규 직교 열들,

- [0193] ● 기준의 변경의 단순화,

- [0194] ● 등급-1 근사치 사용은, 특히 최종 등급이 낮은 경우, 메모리의 노력을 적게 할 수 있고 연산들의 개수를 감소시킬 수 있다. 일반적으로,  $M * N$  대신에  $M \times N$ 의 행렬에 대해서, 단지  $M + N$  연산들만이 요구된다.

- [0195] ● 디코더에서 의사 반전을 피할 수 있기 때문에, 디코더 측에서 적응을 단순화한다.

- [0196] ● 수치적으로 불안정한  $\sigma$ 에 의한 반전 문제들은 피할 수 있다.

[0197] 도 1에 있어서, 인코더 또는 전송자 측에서, 음원들의 서로 다른 방향 값들  $\Omega_s$  ( $s = 1, \dots, S$ ) 및 앰비소닉스 차수  $N_s$  ( $s = 1, \dots, S$ )는 구형 고조파의 대응하는 케트 벡터들  $|Y(\Omega_s)\rangle$  및 차원  $OxS$ 을 갖는 인코더 모드 행렬  $\Xi_{OxS}$ 을 형성하는 단계 또는 단(11)에 입력된다. 행렬  $\Xi_{OxS}$ 는 서로 다른 방향들  $\Omega_s$ 에 대한  $S$  음원 신호들을 포함하는 입력 신호 벡터  $(|x(\Omega_s)\rangle)$ 에 부합하여 생성된다. 따라서, 행렬  $\Xi_{OxS}$ 은 구면 고조파 케트 벡터들  $|Y(\Omega_s)\rangle$ 의 모음이다. 신호  $x(\Omega_s)$  뿐만 아니라 위치가 시간에 따라 변하기 때문에, 연산 행렬  $\Xi_{OxS}$ 은 동적으로 실행될 수 있다. 이러한 행렬은 음원에 대해 비-정규 직교 기준  $NONB$ 을 갖는다.

입력 신호  $(|x(\Omega_s)\rangle)$  및 등급 값  $r_s$  으로부터 특정된 특이 임계값  $(\sigma_\varepsilon)$ 은 단계 또는 단(12)에서 결정된다. 인코더 모드 행렬  $\Xi_{OxS}$  및 임계값  $(\sigma_\varepsilon)$ 은, 단계 또는 단(13)에서 자신의 특이 값들을 얻기 위하여 모드 행렬  $\Xi_{OxS}$ 에 대한 특이 값 분해를 실행하는 절단 특이 값 분해 TSVD 처리(10)(참조: 특이 값 분해에 관련된 단락)에 제공되고, 이에 의해, 한편으로는 단위 행렬들  $U$  및  $V^\dagger$  과  $r_s$  특이 값들  $\sigma_1 \dots \sigma_{r_s}$  을 포함하는 대각 행렬  $\Sigma$  이 출력되고, 다른 한편으로는, 관련된 인코더 모드 행렬 등급  $(r_s)$ 이 결정된다(비고:  $\sigma_i$  는  $SVD(\Xi) = U\Sigma V^+$  의 행렬  $\Sigma$ 로부터  $i$ -번째 특이 값이다).

[0198] 단계/단(12)에 있어서, 임계값  $(\sigma_\varepsilon)$ 은 인코더에서 정규화의 단락에 따라 결정된다. 임계값  $(\sigma_\varepsilon)$ 은 사용 된  $\sigma_{s_i}$  값들의 개수를 절단된 또는 최종 인코더 모드 행렬 등급  $(r_{fin_e})$  으로 제한할 수 있다. 임계값  $(\sigma_\varepsilon)$ 은 미리 결정된 값으로 설정될 수 있거나, 입력 신호:  $\sigma_{\varepsilon,opt} = \frac{1}{\sqrt{SNR}}$  의 신호-대-잡음 비율 SNR로 적응될 수 있으며, 이에 의해, 모든  $S$  음원 신호들  $(|x(\Omega_s)\rangle)$ 의 SNR은 샘플 값들의 미리 정해진 개수를 통해 측정된다.

[0199] 비교 단계 또는 단(14)에 있어서, 행렬  $\Sigma$ 로부터의 특이 값  $\sigma_r$ 은 임계값  $\sigma_\varepsilon$  과 비교되고, 이 비교로부터 절단되거나 최종 인코더 모드 행렬 등급  $r_{fin_e}$  이 계산되며, 이는 인코더에서 정규화의 단락에 따라  $\sigma_{s_i}$  값들

의 나머지를 변경한다. 최종 인코더 모드 행렬 등급  $r_{fin_e}$  은 단계 또는 단(16)에 제공된다.

[0200] 디코더 측에 대해서, 스피커들의 방향 값들  $(\Omega_l)$  ( $l = 1, \dots, L$ )로부터, 그리고, 디코더 앰비소닉스 차수  $N_l$  ( $l = 1, \dots, L$ )로부터, 방향들  $(\Omega_l)$ 에서 특정 스피커들에 대한 구형 고조파의 대응하는 케트 벡터들  $|Y(\Omega_l)\rangle$  뿐만 아니라 차원  $O \times L$  을 갖는 대응하는 디코더 모드 행렬  $\Psi_{O \times L}$ 은, 블록(17)에서 관련된 신호들의 스피커 위치들에 부합하여, 단계 또는 단(18)에서 결정된다. 인코더 행렬  $\Xi_{O \times S}$ 과 유사하게, 디코더 행렬  $\Psi_{O \times L}$ 은 모든 방향들  $(\Omega_l)$ 에 대한 구면 고조파 케트 벡터  $|y(\Omega_l)\rangle$ 들의 모음이다.  $\Psi_{O \times L}$ 의 연산은 동적으로 실행된다.

[0201] 단계 또는 단(19)에 있어서, 특이 값 분해 처리는 디코더 모드 행렬  $\Psi_{O \times L}$ 에 대해 실행되고, 그 결과의 단위 행렬들  $U$  및  $V^\dagger$  뿐만 아니라 대각 행렬  $\Sigma$ 은 블록(17)에 제공된다. 또한, 최종 디코더 모드 행렬 등급  $r_{fin_d}$ 은 계산되어 단계/단(16)에 제공된다.

[0202] 단계 또는 단(16)에 있어서, 최종 모드 행렬 등급  $(r_{fin})$ 은, 상술한 것처럼, 최종 인코더 모드 행렬 등급  $(r_{fin_e})$ 으로부터, 그리고 최종 디코더 모드 행렬 등급  $(r_{fin_d})$ 으로부터 결정된다. 최종 모드 행렬 등급  $(r_{fin})$ 은 단계/단(15) 및 단계/단(17)에 공급된다.

[0203] 인코더-측 행렬들  $U_s$ ,  $V_s^\dagger$ ,  $\Sigma_s$ , 등급 값  $(r_s)$ , 최종 모드 행렬 등급 값  $(r_{fin})$  및 모든 음원 신호들의 시간 의존 입력 신호 케트 벡터  $(|x(\Omega_s)\rangle)$ 는, 식(32)을 사용하여,  $\Xi_{O \times S}$  관련된 입력 값들로부터 그들 인코더 모드 행렬의 수반 의사 역행렬  $(\Xi^+)^{\dagger}$ 을 계산하는 단계 또는 단(15)에 제공된다. 이러한 행렬은 차원  $r_{fin_e} \times S$ 과 음원  $ONB_s$ 에 대한 정규 직교 기준을 갖는다. 복소 행렬들 및 그들의 수반 행렬들을 처리할 때, 다음의 식이 고려된다.

$$\Xi_{O \times S}^\dagger \Xi_{O \times S} = \text{trace}(\Sigma^2) = \sum_{i=1}^r \sigma_{s_i}^2$$

[0204] 단계/단(15)은 대응하는 시간-의존 앰비소닉스 케트 또는 상태 벡터  $|a'_s\rangle$ 를 출력하며, 상술한 HOA 인코더의 단락을 참조한다.

[0206] 단계 또는 단(16)에 있어서,  $|a'_s\rangle$ 의 성분들의 개수는, 성분 적응의 단락에서 설명한 것처럼, 전송된 정보의 양을 가능한 감소시키기 위하여, 최종 모드 행렬 등급  $(r_{fin})$ 을 사용하여 감소되며, 그 결과로서 적응 이후에 시간-의존 앰비소닉스 케트 또는 상태 벡터  $|a'_l\rangle$ 가 생성된다.

[0207] 앰비소닉스 케트 또는 상태 벡터  $|a'_l\rangle$ 로부터, 디코더 측의 행렬들  $U_l^\dagger$ ,  $V_l$ ,  $\Sigma_l$  및 모드 행렬  $\Psi_{O \times L}$ 로부터 유도된 등급 값  $r_l$ 으로부터, 그리고, 단계/단(16)으로부터의 최종 모드 행렬 등급 값  $(r_{fin})$ 으로부터, 차원  $L \times r_{fin_d}$  및 스피커들  $ONB_l$ 에 대한 정규 직교 기준을 갖는 수반 디코더 모드 행렬  $(\Psi)^\dagger$ 이 계산되어, 그 결과로서 모든 스피커들의 시간-의존 출력 신호들의 케트 벡터  $(|y(\Omega_l)\rangle)$ 가 생성된다(상술한 HOA 디코더의 단락 참조). 디코딩은 특정 스피커 위치들에 의존하는 정규 모드 행렬의 공액 교환에 의해 실행된다.

- [0208] 추가적인 렌더링에 대해서는 특정 패닝 행렬이 사용되어야 한다.
- [0209] 디코더는 단계/단(18, 19 및 17)에 의해 표현된다. 인코더는 다른 단계들/단들에 의해 표현된다.
- [0210] 도 1의 단계들/단들(11 내지 19)은 원칙적으로 도 2의 단계들/단들(21 내지 29)과 도 3의 단계들/단들(31 내지 39)에 각각 대응한다.
- [0211] 도 2에 있어서, 단계 또는 단(211)에서 계산된 인코더 측에 대한 패닝 함수  $f_s$  와, 단계 또는 단(281)에서 계산된 디코더 측에 대한 패닝 함수  $f_t$  (281)는 선형 함수의 패닝을 위해 또한 사용된다. 패닝 함수  $f_s$  는 단계/단(21)에 대한 추가적인 입력 신호이고, 패닝 함수  $f_t$  는 단계/단(28)에 대한 추가적인 입력 신호이다. 이러한 패닝 함수들을 사용하는 이유는 패닝 함수들을 고려하는 상기 단락에서 설명되었다.
- [0212] 도 1과 비교하여, 도 3에서, 패닝 행렬  $G$ 는 단계/단(37)의 출력에서 모든 스피커들 중 시간-의존 출력 신호들의 예비 케트 벡터에 대해 패닝 처리(371)를 제어한다. 이는 모든 스피커의 시간-의존 출력 신호의 적응된 케트 벡터  $|y(\Omega_t)\rangle$  의 결과를 얻는다.
- [0213] 도 4는 인코더 모드 행렬  $\Xi_{Oxs}$  의 특이 값 분해 SVD 처리(40)에 기초하여 임계값  $(\sigma_\varepsilon)$  을 결정하기 위한 처리를 보다 상세히 도시한다. 상기 SVD 처리는 행렬  $\Sigma \{ \sigma_1 \}$ 에서  $\sigma_{r_s}$  까지 수행하는 대각선의 모든 특이 값을 내림 차순으로 포함, 식(20) 및 식(21) 참조} 및 행렬  $\Sigma$ 의 등급  $(r_s)$  을 전달한다.
- [0214] 고정된 임계값이 사용되는 경우{블록(41)},  $i = 1$ 에 의해 루프를 시작하여  $i = (r_s)$  까지 수행할 수 있는 변수( $i$ )에 의해 제어되는 루프 내에서{블록 (42 및 43)}, 이를  $\sigma_i$  값들 사이의 수량 값 차이가 있는지를 검사한다 {블록(45)}. 이러한 차이는 특이 값  $\sigma_{i+1}$  의 수량 값이 그 이전 것의 특이 값  $\sigma_1$ 의 수량 값보다 상당히 작은 경우, 예를 들어, 1/10보다 작은 경우에 발생하는 것으로 가정한다. 그러한 차이가 검출될 때, 루프는 정지하고 임계값  $(\sigma_\varepsilon)$  은 현재 특이 값  $\sigma_i$  으로 설정된다{블록(46)}.  $i = r_s$  {블록(44)}인 경우, 최저 특이 값  $\sigma_i = \sigma_r$  에 도달하고, 루프는 종료하고,  $(\sigma_\varepsilon)$  는  $\sigma_r$  로 설정된다{블록(46)}.
- [0215] 고정된 임계값이 사용되지 않는 경우{블록(41)}, 모든 S 음원 신호들  $X = [|x(\Omega_s, t = 0)\rangle, \dots, |x(\Omega_s, t = T)\rangle]$  ( $=$  행렬  $S \times T$ )에 대한 T 샘플의 블록은 조사된다{블록(47)}. X에 대한 신호-대- 노이즈 비율(SNR)이 계산되고{블록(48)}, 임계값  $(\sigma_\varepsilon)$  은  $\sigma_\varepsilon = \frac{1}{\sqrt{SNR}}$  로 설정된다[{블록(49)}].
- [0216] 도 5는 단계/단(15, 25, 35) 내에서 감소된 모드 행렬 등급  $r_{fin}$ , 의 경우에 특이 값들의 재계산과,  $|a'_s\rangle$  의 계산을 도시한다. 도 1 내지 도 3에서 블록(10, 20, 30)으로부터 인코더 대각 행렬  $\Sigma_s$  은, 값  $(r_s)$  을 사용하여 전체 에너지  $trace(\Sigma^2) = \sum_{i=1}^{r_s} \sigma_{s_i}^2$  를 계산하는 단계 또는 단(51)에, 값  $r_{fine}$  을 사용하여 감소된 전체 에너지  $trace(\Sigma_{r_{fine}}^2) = \sum_{i=1}^{r_{fine}} \sigma_{s_i}^2$  를 계산하는 단계 또는 단(52)에, 그리고, 단계 또

는 단(54)에 제공된다. 전체 에너지 값과 감소된 전체 에너지 값 사이의 차이  $\Delta E$ , 값  $trace(\Sigma_{r_{fin_e}})$  및 값  $r_{fin_e}$  은 다음 식을 계산하는 단계 또는 단(53)에 제공된다.

$$\Delta\sigma = \frac{1}{r_{fin_e}} \left( -trace(\Sigma_{r_{fin_e}}) + \sqrt{[trace(\Sigma_{r_{fin_e}})]^2 + r_{fin_e} \Delta E} \right)$$

[0217]

[0218] 값  $\Delta\sigma$  은,  $trace(\Sigma^2) = \sum_{i=1}^r \sigma_{l_i}^2$  에 의해 설명되는 에너지가 유지되어 그 결과가 물리적으로 이해할 수 있도록 보장하기 위해 필요하다.

[0219]

인코더 또는 디코더 측에서, 행렬 감소로 인하여 에너지가 감소되는 경우, 그러한 에너지의 손실은 모든 나머지 행렬 요소들에 동일한 방식으로 분배되는 값  $\Delta\sigma$ 에 의해 보상된다. 즉,  $\sum_{i=1}^r (\sigma_i + \Delta\sigma)^2 = \sum_{i=1}^r (\sigma_i)^2$

[0220]

단계 또는 단(54)은  $\Sigma_s$ ,  $\Delta\sigma$  및  $r_{fin_e}$ 로부터  $\Sigma_t^+ = \sum_{i=1}^{r_{fin_e}} \frac{1}{(\sigma_{s_i} + \Delta\sigma)} I$  을 계산한다.

[0221]

입력 신호 벡터  $(|x(\Omega_s)\rangle)$ 에 행렬  $V_s^\dagger$  이 곱해진다. 이 결과에  $\Sigma_t^+$  을 곱한다. 후자의 곱셈 결과는 케트 벡터  $|a'_s\rangle$  이다.

[0222]

도 6은 단계/단(17, 27, 37)에서 감소된 모드 행렬 등급  $(r_{fin})$  의 경우에 특이 값들의 재계산과, 패닝을 사용하거나 사용하지 않고, 스피커 신호들  $|y(\Omega_l)\rangle$  의 계산을 도시한다. 도 1 내지 도 3에서 블록(19, 29, 39)으로부터 디코더 대각 행렬  $\Sigma_l$  은 값  $r_l$  을 사용하여 전체 에너지  $trace(\Sigma^2) = \sum_{i=1}^{r_l} \sigma_{s_i}^2$  를 계산하는 단계 또는 단(61)에, 값  $r_{fin_d}$  을 사용하여 감소된 전체 에너지  $trace(\Sigma_{r_{fin_d}}^2) = \sum_{i=1}^{r_{fin_d}} \sigma_{s_i}^2$  를 계산하는 단계 또는 단(62)에, 그리고 단계 또는 단(64)에 제공된다. 전체 에너지 값과 감소된 전체 에너지 값 사이의 차이  $\Delta E$ , 값  $trace(\Sigma_{r_{fin_d}})$  및 값  $r_{fin_d}$  은 다음을 계산하는 단계 또는 단(63)에 제공된다.

[0223]

$$\Delta\sigma = \frac{1}{r_{fin_d}} (-trace(\Sigma_{r_{fin_d}}) + \sqrt{[trace(\Sigma_{r_{fin_d}})]^2 + r_{fin_d} \Delta E}) .$$

[0224]

단계 또는 단(64)은  $\Sigma_l$ ,  $\Delta\sigma$  및  $r_{fin_d}$ 로부터  $\Sigma_t = \sum_{i=1}^{r_{fin_d}} \frac{1}{(\sigma_{l_i} + \Delta\sigma)} I$  을 계산한다.

[0225]

케트 벡터  $|a'_s\rangle$ 에 행렬  $\Sigma_t$  이 곱해진다. 이 결과에 행렬  $V$  를 곱한다. 후자의 곱셈 결과는 모든 스피커들의 시간-의존 출력 신호들의 케트 벡터  $(|y(\Omega_l)\rangle)$  이다.

[0226]

본 발명의 처리는 하나의 프로세서 또는 전자 회로에 의해 수행될 수 있거나, 본 발명의 서로 다른 부분에 동

작 및/또는 병렬로 동작하는 전자 회로들 또는 여러 프로세서들에 의해 수행될 수 있다.

### 부호의 설명

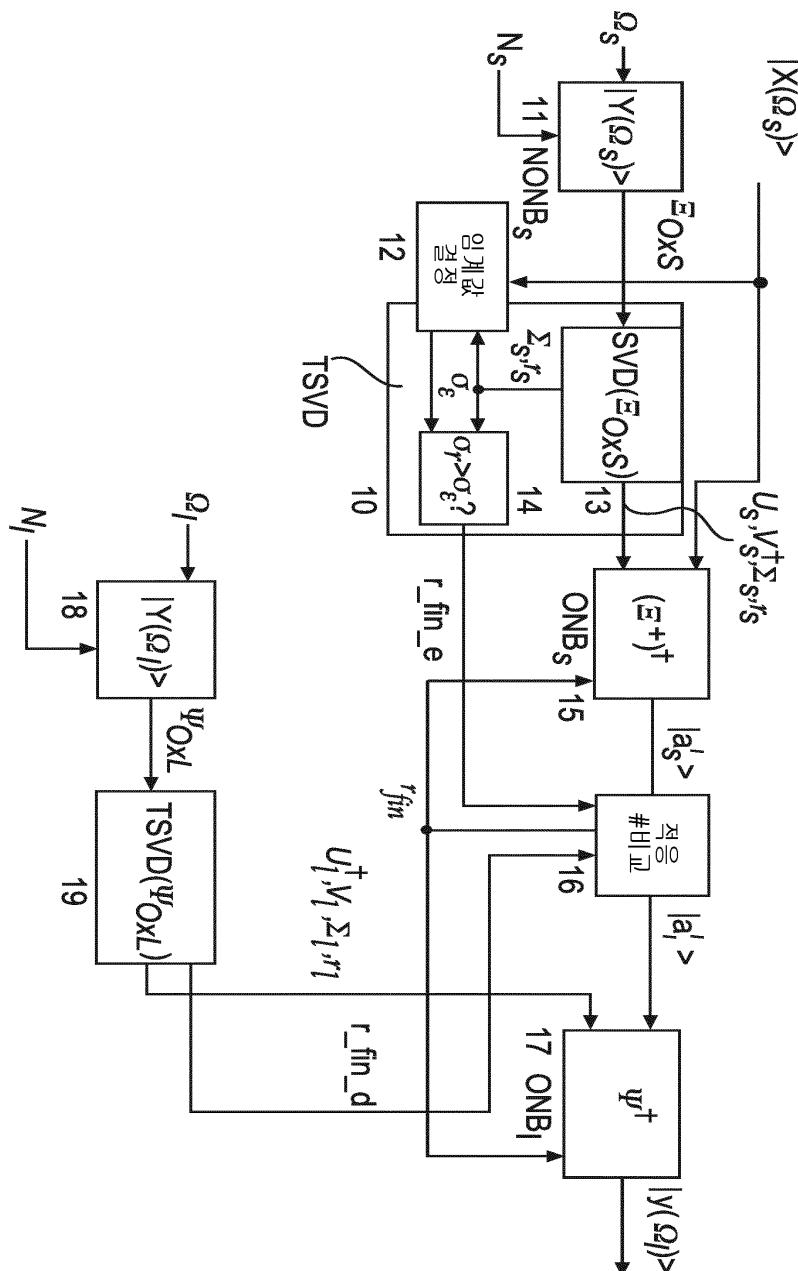
12, 22, 32: 임계값 결정

16, 26, 36: 적응#비교

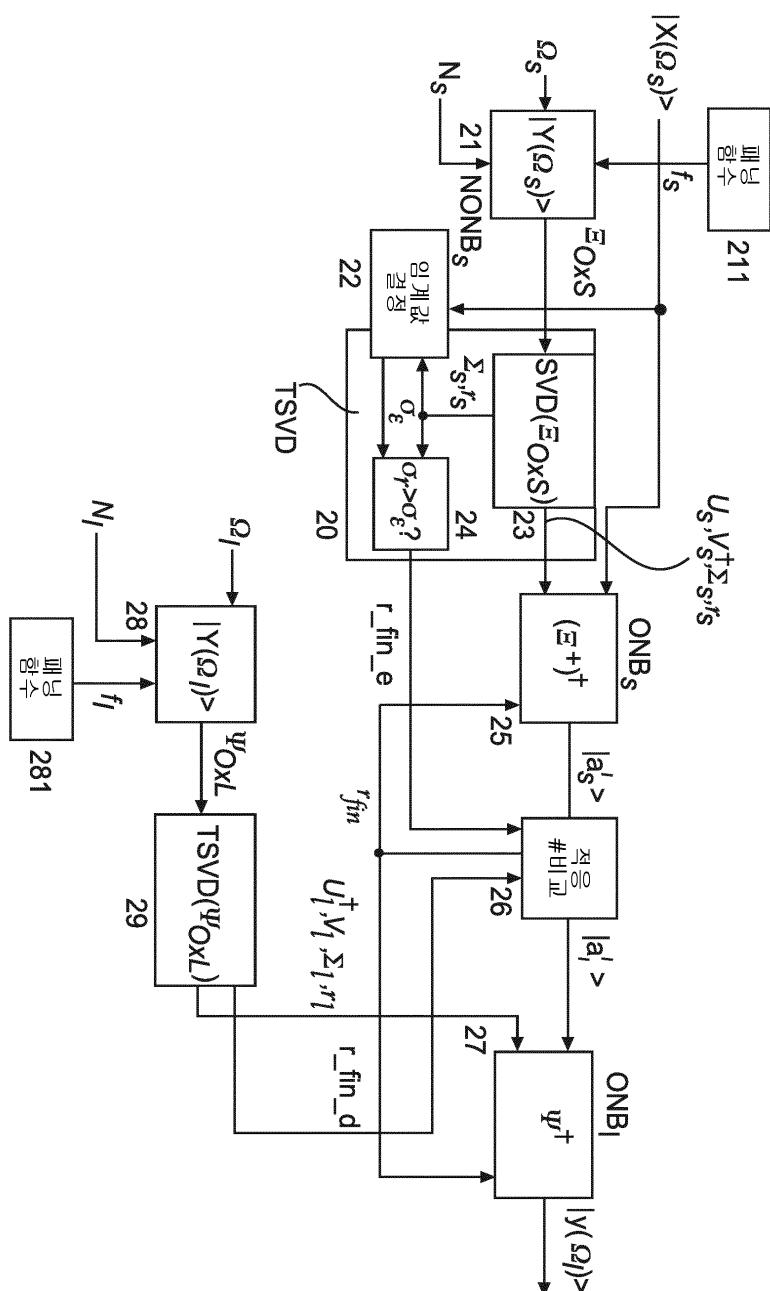
211, 281, 371: 패닝 함수

### 도면

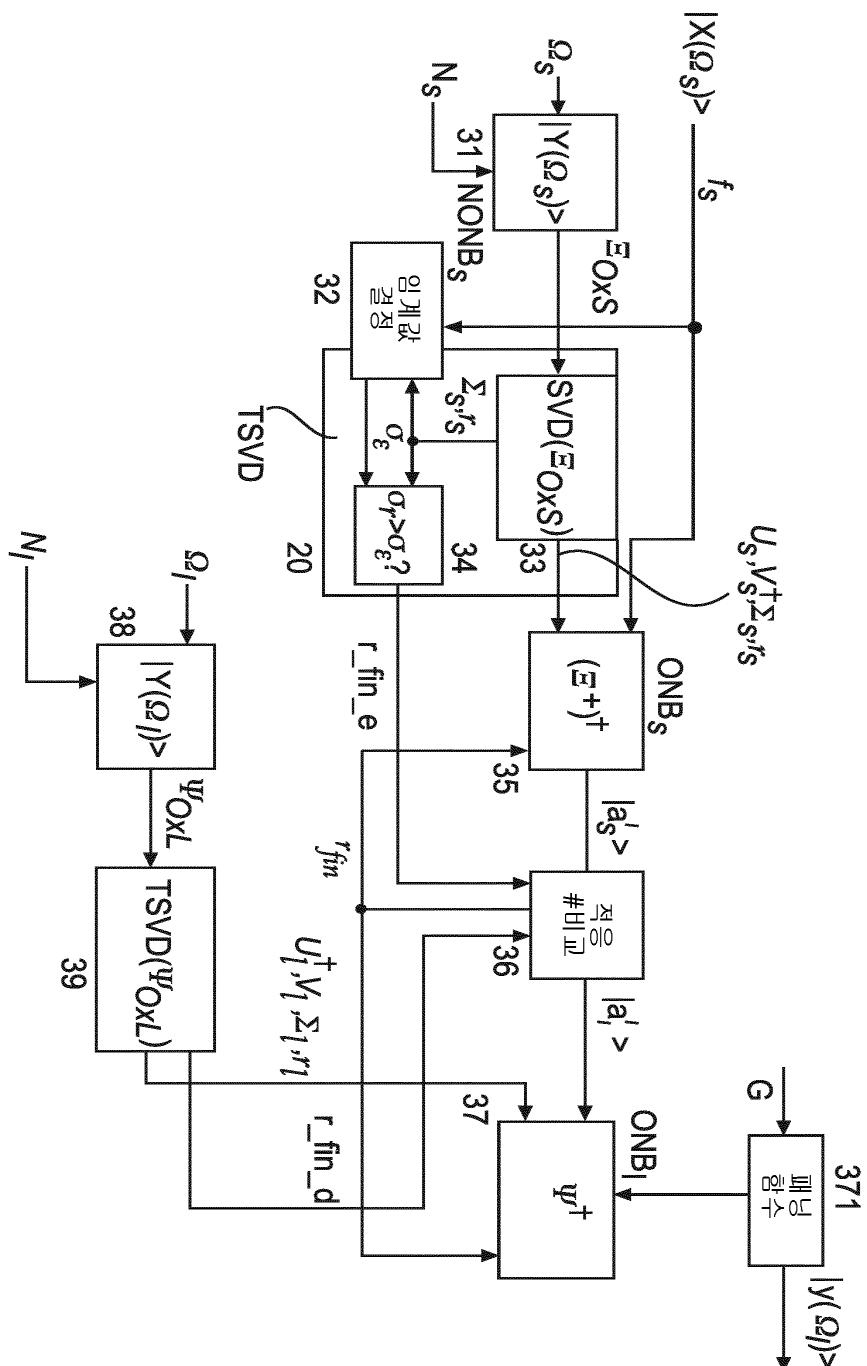
#### 도면1



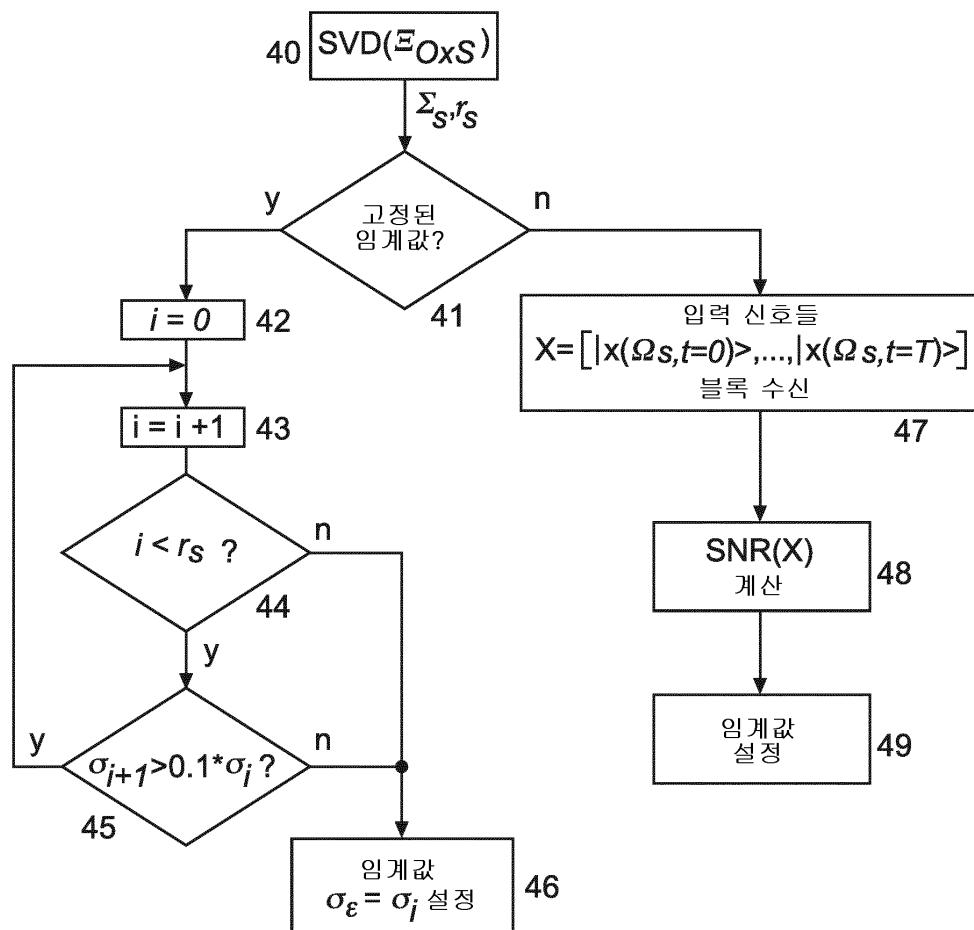
도면2



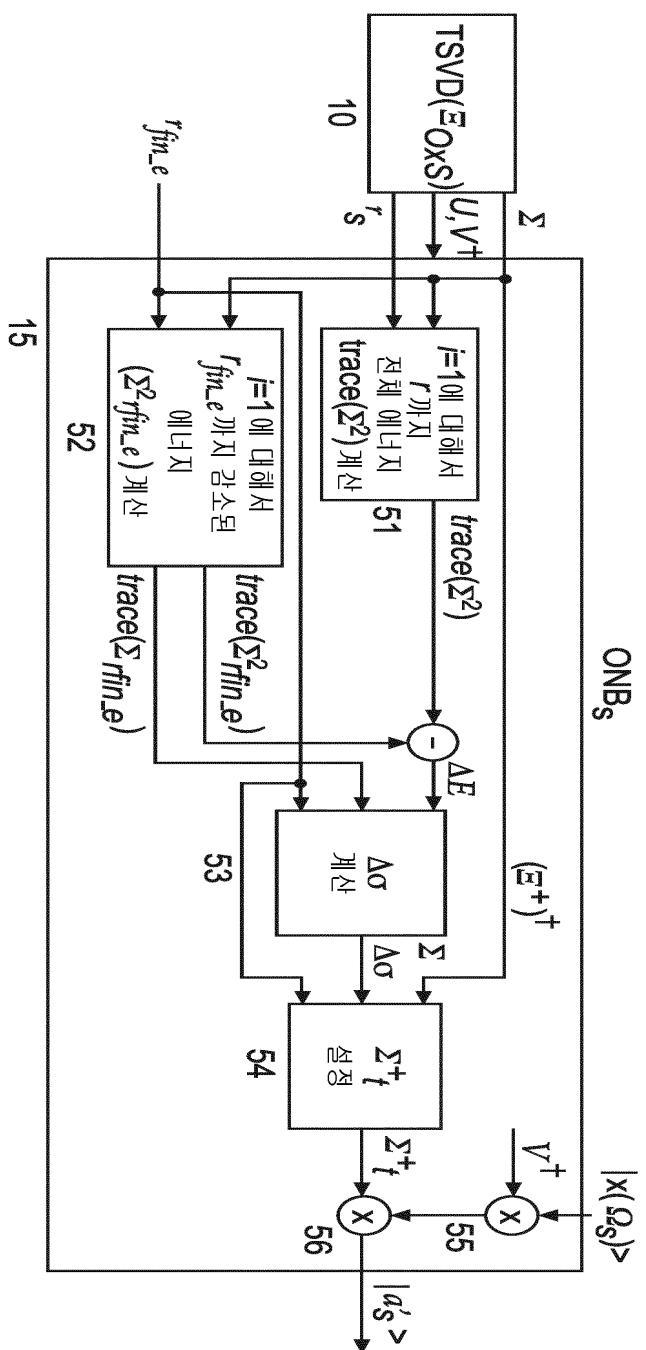
도면3



## 도면4



## 도면5



도면6

