



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년08월01일
(11) 등록번호 10-2426965
(24) 등록일자 2022년07월26일

- (51) 국제특허분류(Int. Cl.)
G10L 21/0364 (2013.01) G10L 19/008 (2014.01)
- (52) CPC특허분류
G10L 21/0364 (2021.08)
G10L 19/008 (2020.08)
- (21) 출원번호 10-2017-7008933
- (22) 출원일자(국제) 2015년09월30일
심사청구일자 2020년09월29일
- (85) 번역문제출일자 2017년03월31일
- (65) 공개번호 10-2017-0063667
- (43) 공개일자 2017년06월08일
- (86) 국제출원번호 PCT/EP2015/072578
- (87) 국제공개번호 WO 2016/050854
국제공개일자 2016년04월07일
- (30) 우선권주장
62/059,015 2014년10월02일 미국(US)
62/128,331 2015년03월04일 미국(US)
- (56) 선행기술조사문헌
US20100014692 A1*
US20110119061 A1*
*는 심사관에 의하여 인용된 문헌

- (73) 특허권자
돌비 인터네셔널 에이비
네덜란드 1097디피 암스테르담 엠알. 트레우블란 7
- (72) 발명자
코펜스, 예룬
스웨덴 에스-113 30 스톡홀름 야블레가탄 12에이
돌비 스웨덴 에이비 내
엑스트란드, 파르
스웨덴 에스-113 30 스톡홀름 야블레가탄 12에이
돌비 스웨덴 에이비 내
- (74) 대리인
양영준, 이민호, 백만기

전체 청구항 수 : 총 19 항

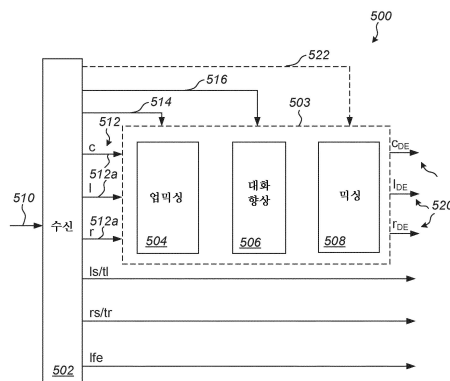
심사관 : 전세운

(54) 발명의 명칭 대화 향상을 위한 디코딩 방법 및 디코더

(57) 요약

오디오 시스템의 디코더에서 대화를 향상시키는 방법이 제공된다. 본 방법은 보다 많은 복수의 채널들의 다운믹스인 복수의 다운믹스 신호들을 수신하는 단계; 복수의 다운믹스 신호들의 서브셋으로 다운믹스되는 복수의 채널들의 서브셋과 관련하여 정의되는 대화 향상 파라미터들을 수신하는 단계; 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋을 재구성하기 위해 다운믹스 신호들의 서브셋을 파라미터적으로 업믹싱하는 단계; 적어도 하나의 대화 향상된 신호를 제공하기 위해 대화 향상 파라미터들을 사용하여 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋에 대화 향상을 적용하는 단계; 및 다운믹스 신호들의 서브셋의 대화 향상된 버전들을 제공하기 위해 적어도 하나의 대화 향상된 신호에 믹싱을 가하는 단계를 포함한다.

대표도 - 도5



명세서

청구범위

청구항 1

오디오 시스템의 디코더에서 대화(dialog)를 향상시키는 방법으로서,

인코딩된 비트스트림으로, 복수의 채널들의 다운믹스(downmix)인 복수의 다운믹스 신호들을 수신하는 단계;

인코딩된 비트스트림으로, 대화 향상 파라미터들(parameters for dialog enhancement)을 수신하는 단계 - 상기 파라미터들은 대화를 포함하는 채널들을 포함하는 상기 복수의 채널들의 서브셋과 관련하여 정의되고, 상기 복수의 채널들의 상기 서브셋은 상기 복수의 다운믹스 신호들의 서브셋으로 다운믹싱됨 -;

인코딩된 비트스트림으로, 상기 복수의 다운믹스 신호들의 상기 서브셋으로 다운믹싱되는 채널들의 파라미터적 재구성(parametric reconstruction)을 가능하게 하는 재구성 파라미터들을 수신하는 단계;

상기 대화 향상 파라미터들이 정의되는 것과 관련된 상기 복수의 채널들의 상기 서브셋을 포함하는 상기 복수의 채널들의 서브셋만을 재구성하기 위해 상기 재구성 파라미터들에 기초하여 상기 복수의 다운믹스 신호들의 상기 서브셋만을 파라미터적으로 업믹싱(upmixing)하는 단계;

적어도 하나의 대화 향상된 신호를 제공하기 위해 상기 대화 향상 파라미터들을 사용하여 상기 대화 향상 파라미터들이 정의되는 것과 관련된 상기 복수의 채널들의 상기 서브셋에 대화 향상을 적용하는 단계; 및

상기 적어도 하나의 대화 향상된 신호와 적어도 하나의 다른 신호를 믹싱하여 상기 복수의 다운믹스 신호들의 상기 서브셋의 대화 향상된 버전들을 제공하는 단계

를 포함하는, 방법.

청구항 2

제1항에 있어서, 상기 복수의 다운믹스 신호들의 상기 서브셋만을 파라미터적으로 업믹싱하는 단계에서, 상기 대화 향상 파라미터들이 정의되는 것과 관련된 상기 복수의 채널들의 상기 서브셋을 포함하는 상기 복수의 채널들의 서브셋만을 재구성하기 위해 역상관된 신호(decorrelated signal)들이 사용되지 않는, 방법.

청구항 3

제1항에 있어서, 상기 복수의 다운믹스 신호들의 상기 서브셋의 상기 대화 향상된 버전들에 대한 상기 적어도 하나의 대화 향상된 신호의 기여도(contribution)를 기술하는 믹싱 파라미터들에 따라 상기 믹싱이 행해지는, 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 복수의 다운믹스 신호들의 상기 서브셋만을 파라미터적으로 업믹싱하는 단계는 상기 대화 향상 파라미터들이 정의되는 것과 관련된 상기 복수의 채널들의 상기 서브셋만을 재구성하는 단계를 포함하고,

대화 향상을 적용하는 상기 단계는 상기 적어도 하나의 대화 향상된 신호를 제공하기 위해 상기 대화 향상 파라미터들을 사용하여 상기 대화 향상 파라미터들이 정의되는 것과 관련된 상기 복수의 채널들의 상기 서브셋으로부터 대화 컴포넌트(dialog component)를 예측하고 향상시키는 단계를 포함하며,

상기 믹싱은 상기 적어도 하나의 대화 향상된 신호를 상기 복수의 다운믹스 신호들의 상기 서브셋과 믹싱하는 것을 포함하는, 방법.

청구항 5

제1항 내지 제3항 중 어느 한 항에 있어서, 대화를 나타내는 오디오 신호를 수신하는 단계를 추가로 포함하고, 대화 향상을 적용하는 상기 단계는 대화를 나타내는 상기 오디오 신호를 추가로 사용하여 상기 대화 향상 파라미터들이 정의되는 것과 관련된 상기 복수의 채널들의 상기 서브셋에 대화 향상을 적용하는 단계를 포함하는,

방법.

청구항 6

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 적어도 하나의 대화 향상된 신호와 적어도 하나의 다른 신호를 믹싱하기 위한 믹싱 파라미터들을 수신하는 단계를 추가로 포함하는, 방법.

청구항 7

제1항 또는 제2항에 있어서, 상기 복수의 다운믹스 신호들의 상기 서브셋만을 업믹싱하는 단계, 대화 향상을 적용하는 단계, 및 믹싱하는 단계는, 각각, 상기 재구성 파라미터들, 상기 대화 향상 파라미터들, 및 상기 복수의 다운믹스 신호들의 상기 서브셋의 상기 대화 향상된 버전들에 대한 상기 적어도 하나의 대화 향상된 신호의 기여도를 기술하는 믹싱 파라미터들에 의해 정의되는 행렬 연산들로서 수행되고, 선택적으로, 상기 복수의 다운믹스 신호들의 상기 서브셋에 적용하기 전에 상기 복수의 다운믹스 신호들의 상기 서브셋만을 업믹싱하는 단계, 대화 향상을 적용하는 단계, 및 믹싱하는 단계에 대응하는 상기 행렬 연산들을, 행렬 곱셈에 의해, 단일의 행렬 연산으로 결합하는 단계를 추가로 포함하는, 방법.

청구항 8

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 대화 향상 파라미터들 및 상기 재구성 파라미터들은 주파수의 존적이고, 선택적으로, 상기 대화 향상 파라미터들은 제1 주파수 대역 세트(set of frequency bands)와 관련하여 정의되고, 상기 재구성 파라미터들은 제2 주파수 대역 세트와 관련하여 정의되며, 상기 제2 주파수 대역 세트는 상기 제1 주파수 대역 세트와 상이한, 방법.

청구항 9

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 대화 향상 파라미터들의 값들은 반복하여 수신되고, 각자의 값들이 정확히 적용되는, 제1 시간 순간 세트(set of time instants)($T_1 = \{t_{11}, t_{12}, t_{13}, \dots\}$)와 연관되며, 미리 정의된 제1 보간 패턴(I_1)이 연속적인 시간 순간들 사이에서 수행되어야 하고,

상기 재구성 파라미터들의 값들은 반복하여 수신되고, 각자의 값들이 정확히 적용되는, 제2 시간 순간 세트($T_2 = \{t_{21}, t_{22}, t_{23}, \dots\}$)와 연관되며, 미리 정의된 제2 보간 패턴(I_2)이 연속적인 시간 순간들 사이에서 수행되어야 하고,

상기 방법은

대화 향상 파라미터들 또는 재구성 파라미터들 중 어느 하나인 파라미터 유형을, 선택된 유형과 연관된 시간 순간 세트가 비선택된 유형과 연관된 세트에는 없는 시간 순간(t_p)인 적어도 하나의 예측 순간을 포함하는 방식으로 선택하는 단계;

상기 비선택된 유형의 파라미터들의 값을 상기 예측 순간(t_p)에 예측하는 단계;

적어도 상기 비선택된 유형의 상기 파라미터들의 예측된 값 및 상기 선택된 유형의 상기 파라미터들의 수신된 값에 기초하여, 상기 예측 순간(t_p)에서 적어도 상기 다운믹스 신호들의 상기 서브셋만의 업믹싱 및 그에 뒤이은 대화 향상을 나타내는 결합 처리 연산(joint processing operation)을 계산하는 단계; 및

적어도 상기 선택된 유형의 상기 파라미터들의 값 및 상기 비선택된 유형의 상기 파라미터들의 값 - 적어도 어느 하나는 수신된 값임 - 에 기초하여, 상기 선택된 유형 또는 상기 비선택된 유형과 연관된 상기 세트 내의 인접한 시간 순간(t_a)에서 상기 결합 처리 연산을 계산하는 단계를 추가로 포함하고,

상기 복수의 다운믹스 신호들의 상기 서브셋만을 업믹싱하는 상기 단계 및 대화 향상을 적용하는 상기 단계는 상기 계산된 결합 처리 연산의 보간된 값을 통해 상기 예측 순간(t_p)과 상기 인접한 시간 순간(t_a) 사이에서 수행되는, 방법.

청구항 10

제9항에 있어서, 상기 선택된 유형의 파라미터들은 상기 재구성 파라미터들인, 방법.

청구항 11

제9항에 있어서, 상기 인접한 시간 순간(t_a)에서의 상기 결합 처리 연산은 상기 선택된 유형의 상기 파라미터들의 수신된 값 및 상기 비선택된 유형의 상기 파라미터들의 수신된 값에 기초하여 계산되는, 방법.

청구항 12

제9항에 있어서,

상기 제1 및 제2 보간 패턴들에 기초하여, 미리 정의된 선택 규칙에 따라 결합 보간 패턴(I3)을 선택하는 단계를 추가로 포함하고,

상기 계산된 각자의 결합 처리 연산들의 상기 보간은 상기 결합 보간 패턴에 따르는, 방법.

청구항 13

제12항에 있어서, 상기 제1 및 제2 보간 패턴들이 상이한 경우에 대해 상기 미리 정의된 선택 규칙이 정의되고, 선택적으로, 상기 제1 보간 패턴(I1)이 선형적이고 상기 제2 보간 패턴(I2)이 구간별 일정한(piecewise constant) 것에 응답하여, 선형 보간이 상기 결합 보간 패턴으로서 선택되는, 방법.

청구항 14

제9항에 있어서, 상기 예측 순간(t_p)에 상기 비선택된 유형의 파라미터들의 값의 예측은 상기 비선택된 유형의 파라미터들에 대한 상기 보간 패턴에 따라 행해지는, 방법.

청구항 15

제9항에 있어서, 상기 결합 처리 연산은, 상기 복수의 다운믹스 신호들의 상기 서브셋에 적용되기 전에, 단일의 행렬 연산으로서 계산되고, 선택적으로,

선형 보간이 상기 결합 보간 패턴으로서 선택되고;

상기 계산된 각자의 결합 처리 연산들의 상기 보간된 값이 선형 행렬 보간에 의해 계산되는, 방법.

청구항 16

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 적어도 하나의 대화 향상된 신호와 적어도 하나의 다른 신호를 믹싱하는 것은 상기 복수의 다운믹스 신호들의 비전체 선택(selection)으로 제한되는, 방법.

청구항 17

제1항 내지 제3항 중 어느 한 항에 있어서, 다운믹스 신호들의 개수는 채널들의 개수보다 적은, 방법.

청구항 18

제1항 내지 제3항 중 어느 한 항의 방법을 수행하기 위한 명령어들을 갖는, 컴퓨터 판독가능 기록 매체에 저장된 컴퓨터 프로그램.

청구항 19

오디오 시스템에서 대화를 향상시키는 디코더로서,

제1항 내지 제3항 중 어느 한 항의 방법을 수행하도록 구성되는 하나 이상의 컴포넌트를 포함하는, 디코더.

청구항 20

삭제

청구항 21

- 삭제
- 청구항 22
- 삭제
- 청구항 23
- 삭제
- 청구항 24
- 삭제
- 청구항 25
- 삭제
- 청구항 26
- 삭제
- 청구항 27
- 삭제
- 청구항 28
- 삭제
- 청구항 29
- 삭제

발명의 설명

기술 분야

[0001] 본원에 게시되는 본 발명은 일반적으로 오디오 코딩에 관한 것이다. 상세하게는, 본 발명은 채널 기반 오디오 시스템들에서 대화를 향상시키는 방법들 및 디바이스들에 관한 것이다.

배경 기술

[0002] 대화 향상(dialog enhancement)은 다른 오디오 콘텐츠와 관련하여 대화를 향상시키는 것에 관한 것이다. 이것은, 예를 들어, 청각 장애가 있는 사람들이 영화에서의 대화를 알아들을 수 있게 하기 위해 적용될 수 있다. 채널 기반 오디오 콘텐츠에 있어서, 대화는 전형적으로 몇 개의 채널들에 존재하고 또한 다른 오디오 콘텐츠와 믹싱된다. 따라서, 대화를 향상시키는 것은 쉬운 일이 아니다.

[0003] 디코더에서 대화 향상을 수행하는 몇 가지 공지된 방법들이 있다. 이 방법들 중 일부에 따르면, 전체 채널 콘텐츠, 즉 전체 채널 구성이 먼저 디코딩되고, 이어서 수신된 대화 향상 파라미터들이 전체 채널 콘텐츠에 기초하여 대화를 예측하는 데 사용된다. 예측된 대화는 이어서 관련 채널들에서의 대화를 향상시키는 데 사용된다. 그렇지만, 이러한 디코딩 방법들은 전체 채널 구성을 디코딩할 수 있는 디코더에 의존한다.

[0004] 그렇지만, 저 복잡도 디코더들은 전형적으로 전체 채널 구성을 디코딩하도록 설계되어 있지 않다. 그 대신에, 저 복잡도 디코더는 전체 채널 구성의 다운믹싱된 버전(downmixed version)을 나타내는 보다 적은 수의 채널들을 디코딩하고 출력할 수 있다. 그에 따라, 전체 채널 구성이 저 복잡도 디코더에서는 이용가능하지 않다. 대화 향상 파라미터들이 전체 채널 구성의 채널들과 관련하여(또는 적어도 전체 채널 구성의 채널들 중 일부와 관련하여) 정의되기 때문에, 공지된 대화 향상 방법들은 저 복잡도 디코더에 의해 직접 적용될 수 없다. 상세하게는, 이러한 이유는 대화 향상 파라미터들이 적용되는 채널들이 여전히 다른 채널들과 믹싱되어 있을 수 있기 때문이다.

[0005] 따라서 저 복잡도 디코더가 전체 채널 구성을 디코딩할 필요 없이 대화 향상을 적용할 수 있게 하도록 개선할 여지가 있다.

도면의 간단한 설명

[0006] 이하에서, 예시적인 실시예들이 더욱 상세히 그리고 첨부 도면들을 참조하여 기술될 것이다.

도 1a는 제1 다운믹싱(downmixing) 방식에 따라 5.1 다운믹스(downmix)로 다운믹싱되는 7.1+4 채널 구성의 개략도.

도 1b는 제2 다운믹싱 방식에 따라 5.1 다운믹스로 다운믹싱되는 7.1+4 채널 구성의 개략도.

도 2는 전체적으로 디코딩된 채널 구성(fully decoded channel configuration)에 대해 대화 향상을 수행하는 종래 기술의 디코더의 개략도.

도 3은 제1 모드에 따른 대화 향상의 개략도.

도 4는 제2 모드에 따른 대화 향상의 개략도.

도 5는 예시적인 실시예들에 따른 디코더의 개략도.

도 6은 예시적인 실시예들에 따른 디코더의 개략도.

도 7은 예시적인 실시예들에 따른 디코더의 개략도.

도 8은 도 2, 도 5, 도 6, 및 도 7에서의 디코더들 중 임의의 디코더에 대응하는 인코더의 개략도.

도 9는, 하위 연산(sub-operation)들 각각을 제어하는 파라미터들에 기초하여, 2개의 하위 연산 A 및 B로 이루어진 결합 처리 연산(joint processing operation) BA를 계산하는 방법들을 나타낸 도면.

도면들 모두는 개략적이고 일반적으로 본 발명을 설명하는 데 필요한 그러한 요소들만을 보여주는 반면, 다른 요소들은 생략되거나 단순히 암시될 수 있다.

발명을 실시하기 위한 구체적인 내용

[0007] 이상의 내용을 고려하여, 전체 채널 구성을 디코딩할 필요 없이 대화 향상의 적용을 가능하게 하는 디코더 및 연관된 방법들을 제공하는 것이 목적이다.

I. 개요

[0009] 제1 양태에 따르면, 예시적인 실시예들은 오디오 시스템의 디코더에서 대화를 향상시키는 방법을 제공한다. 본 방법은

[0010] 보다 많은 복수의 채널들의 다운믹스인 복수의 다운믹스 신호들을 수신하는 단계;

[0011] 대화 향상 파라미터들을 수신하는 단계 - 파라미터들은 대화를 포함하는 채널들을 포함하는 복수의 채널들의 서브셋과 관련하여 정의되고, 복수의 채널들의 서브셋은 복수의 다운믹스 신호들의 서브셋으로 다운믹싱됨 -;

[0012] 복수의 다운믹스 신호들의 서브셋으로 다운믹싱되는 채널들의 파라미터적 재구성(parametric reconstruction)을 가능하게 하는 재구성 파라미터들을 수신하는 단계;

[0013] 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋을 재구성하기 위해 재구성 파라미터들에 기초하여 복수의 다운믹스 신호들의 서브셋을 파라미터적으로 업믹싱(upmixing)하는 단계;

[0014] 적어도 하나의 대화 향상된 신호를 제공하기 위해 대화 향상 파라미터들을 사용하여 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋에 대화 향상을 적용하는 단계; 및

[0015] 복수의 다운믹스 신호들의 서브셋의 대화 향상된 버전들을 제공하기 위해 적어도 하나의 대화 향상된 신호에 믹싱을 가하는 단계를 포함한다.

[0016] 이 구성에 의해, 디코더는 대화 향상을 수행하기 위해 전체 채널 구성을 재구성할 필요가 없고, 그래서 복잡도를 감소시킨다. 그 대신에, 디코더는 대화 향상의 적용을 위해 필요하게 되는 그 채널들을 재구성한다. 이것은, 상세하게는, 수신된 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋을 포함한다. 대화 향상이 수행되었으면, 즉 적어도 하나의 대화 향상된 신호가 대화 향상 파라미터들 및 이 파라미터들이 정의되는 복수

의 채널들의 서브셋에 기초하여 결정되었을 때, 대화 향상된 신호(들)에 믹싱 절차를 가함으로써 수신된 다운믹스 신호들의 대화 향상된 버전들이 결정된다. 그 결과, 다운믹스 신호들의 대화 향상된 버전들이 오디오 시스템에 의한 차후의 재생을 위해 생성된다.

- [0017] 예시적인 실시예들에서, 업믹싱 연산은 전체적(인코딩된 채널들의 전체 세트를 재구성함)이거나 부분적(채널들의 서브셋을 재구성함)일 수 있다.
- [0018] 본원에서 사용되는 바와 같이, 다운믹스 신호는 하나 이상의 신호들/채널들의 조합인 신호를 지칭한다.
- [0019] 본원에서 사용되는 바와 같이, 파라미터적으로 업믹싱하는 것은 파라미터적 기법들에 의해 다운믹스 신호로부터 하나 이상의 신호들/채널들을 재구성하는 것을 지칭한다. 본원에 개시되는 예시적인 실시예들이 (오디오 신호들이 불변적이거나 미리 정의된 방향들, 각도들 및/또는 공간에서의 위치들과 연관되어 있다는 의미에서) 채널 기반 콘텐츠로 제한되지 않고 객체 기반 콘텐츠로도 확장된다는 점이 강조된다.
- [0020] 예시적인 실시예들에 따르면, 복수의 다운믹스 신호들의 서브셋을 파라미터적으로 업믹싱하는 단계에서, 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋을 재구성하기 위해 역상관된 신호(decorrelated signal)들이 사용되지 않는다.
- [0021] 이것은 다운믹스 신호들의 그 결과 얻어진 대화 향상된 버전들의 품질(즉, 출력에서의 품질)을 개선시키는 것과 동시에 계산 복잡도를 감소시킨다는 점에서 유리하다. 더욱 상세하게는, 업믹싱할 때 역상관된 신호들을 사용함으로써 얻어지는 장점들이 대화 향상된 신호에 가해지는 차후의 믹싱에 의해 감소된다. 따라서, 역상관된 신호들의 사용이 유리하게도 생략될 수 있고, 그로써 계산 복잡도를 절감할 수 있다. 사실상, 업믹싱에서의 역상관된 신호들의 사용은, 대화 향상과 조합하여, 품질 악화를 가져올 수 있는데, 그 이유는 그로 인해 향상된 대화에 대한 역상관기 리버브(decorrelator reverb)를 가져올 수 있기 때문이다.
- [0022] 예시적인 실시예들에 따르면, 복수의 다운믹스 신호들의 서브셋의 대화 향상된 버전들에 대한 적어도 하나의 대화 향상된 신호의 기여도를 나타내는 믹싱 파라미터들에 따라 믹싱이 행해진다. 따라서 복수의 다운믹스 신호들의 서브셋의 대화 향상된 버전들을 제공하기 위해 적어도 하나의 대화 향상된 신호를 어떻게 믹싱할지를 기술하는 어떤 믹싱 파라미터들이 있을 수 있다. 예를 들어, 믹싱 파라미터들이 복수의 다운믹스 신호들의 서브셋의 대화 향상된 버전들을 획득하기 위해 적어도 하나의 대화 향상된 신호의 얼마만큼이 복수의 다운믹스 신호들의 서브셋 내의 다운믹스 신호들 각각에 믹싱되어야만 하는지를 기술하는 가중치들의 형태로 되어 있을 수 있다. 이러한 가중치들은, 예를 들어, 적어도 하나의 대화 향상된 신호와 연관된 공간 위치들을 복수의 채널들, 그리고 따라서 대응하는 다운믹스 신호들의 서브셋과 연관된 공간 위치들과 관련하여 나타내는 렌더링 파라미터들의 형태로 되어 있을 수 있다. 다른 예들에 따르면, 믹싱 파라미터들은 적어도 하나의 대화 향상된 신호가 다운믹스 신호들의 서브셋의 대화 향상된 버전의 특징의 것에 기여(그에 포함되는 것 등)해야만 하는지 여부를 나타낼 수 있다. 예를 들어, "1"은 다운믹스 신호들의 대화 향상된 버전의 특징의 것을 형성할 때 대화 향상된 신호가 포함되어야만 한다는 것을 나타낼 수 있고, "0"은 대화 향상된 신호가 포함되어서는 안된다는 것을 나타낼 수 있다.
- [0023] 복수의 다운믹스 신호들의 서브셋의 대화 향상된 버전들을 제공하기 위해 적어도 하나의 대화 향상된 신호에 믹싱을 가하는 단계에서, 대화 향상된 신호들이 다른 신호들/채널들과 믹싱될 수 있다.
- [0024] 예시적인 실시예들에 따르면, 적어도 하나의 대화 향상된 신호가 업믹싱 단계에서 재구성되지만 대화 향상을 거치지 않은 채널들과 믹싱된다. 더욱 상세하게는, 복수의 다운믹스 신호들의 서브셋을 파라미터적으로 업믹싱하는 단계는 대화 향상 파라미터들이 정의되는 복수의 채널들 이외의 적어도 하나의 추가 채널을 재구성하는 단계를 포함할 수 있고, 여기서 믹싱은 적어도 하나의 추가 채널을 적어도 하나의 대화 향상된 신호와 함께 믹싱하는 것을 포함한다. 예를 들어, 복수의 다운믹스 신호들의 서브셋으로 다운믹싱되는 모든 채널들이 재구성되고 믹싱에 포함될 수 있다. 이러한 실시예들에서, 전형적으로 각각의 대화 향상된 신호와 채널 간에 직접적인 대응관계가 있다.
- [0025] 다른 예시적인 실시예들에 따르면, 적어도 하나의 대화 향상된 신호가 복수의 다운믹스 신호들의 서브셋과 믹싱된다. 더욱 상세하게는, 복수의 다운믹스 신호들의 서브셋을 파라미터적으로 업믹싱하는 단계는 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋만을 재구성하는 단계를 포함할 수 있고, 대화 향상을 적용하는 단계는 적어도 하나의 대화 향상된 신호를 제공하기 위해 대화 향상 파라미터들을 사용하여 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋으로부터 대화 컴포넌트를 예측하고 향상시키는 단계를 포함할 수 있고, 믹싱은 적어도 하나의 대화 향상된 신호를 복수의 다운믹스 신호들의 서브셋과 믹싱하는 것을 포함할 수 있다.

이러한 실시예들은 따라서 대화 콘텐츠를 예측하고 향상시키며 그것을 복수의 다운믹스 신호들의 서브셋에 믹싱하는 역할을 한다.

- [0026] 일반적으로 유의할 점은 채널이 비대화 콘텐츠와 믹싱되는 대화 콘텐츠를 포함할 수 있다는 것이다. 게다가, 하나의 대화에 대응하는 대화 콘텐츠가 몇 개의 채널들에 믹싱될 수 있다. 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋으로부터 대화 컴포넌트를 예측하는 것이란 일반적으로 대화를 재구성하기 위해 대화 콘텐츠가 채널들로부터 추출, 즉 분리되고 결합된다는 것을 의미한다.
- [0027] 대화 향상의 품질이 대화를 나타내는 오디오 신호를 수신하고 사용하는 것에 의해 추가로 개선될 수 있다. 예를 들어, 대화를 나타내는 오디오 신호가 저 비트레이트로 코딩되어, 개별적으로 청취될 때 잘 들리는 아티팩트들을 야기할 수 있다. 그렇지만, 파라미터적 대화 향상, 즉 대화 향상 파라미터들을 사용하여 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋에 대화 향상을 적용하는 단계와 함께 사용될 때, 그 결과 얻어진 대화 향상이, 예컨대, 오디오 품질의 면에서 개선될 수 있다. 더욱 상세하게는, 본 방법은 대화를 나타내는 오디오 신호를 수신하는 단계를 추가로 포함할 수 있고, 여기서 대화 향상을 적용하는 단계는 대화를 나타내는 오디오 신호를 추가로 사용하여 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋에 대화 향상을 적용하는 단계를 포함한다.
- [0028] 일부 실시예들에서, 믹싱 파라미터들이 디코더에서 이미 이용가능할 수 있고, 예컨대, 그들이 하드코딩되어 있을 수 있다. 적어도 하나의 대화 향상된 신호가 향상 동일한 방식으로 믹싱되는 경우에, 예컨대, 그것이 향상 동일한 재구성된 채널들과 믹싱되는 경우에, 특히 그러할 수 있다. 다른 실시예들에서, 본 방법은 적어도 하나의 대화 향상된 신호에 믹싱을 가하는 단계를 위해 믹싱 파라미터들을 수신하는 단계를 포함한다. 예를 들어, 믹싱 파라미터들은 대화 향상 파라미터들의 일부를 형성할 수 있다.
- [0029] 예시적인 실시예들에 따르면, 본 방법은 복수의 채널들 각각이 어느 다운믹스 신호로 믹싱되는지를 기술하는 다운믹싱 방식을 기술하는 믹싱 파라미터들을 수신하는 단계를 포함한다. 예를 들어, 각각의 대화 향상된 신호가, 다른 재구성된 채널들과 차례로 믹싱되는, 채널에 대응하는 경우, 각각의 채널이 올바른 다운믹스 신호에 믹싱되도록 믹싱이 다운믹싱 방식에 따라 수행된다.
- [0030] 다운믹싱 방식이 시간에 따라 변할 수 있고 - 즉, 동적일 수 있음 -, 그로써 시스템의 유연성을 증대시킬 수 있다.
- [0031] 본 방법은 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋을 식별하는 데이터를 수신하는 단계를 추가로 포함할 수 있다. 예를 들어, 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋을 식별하는 데이터는 대화 향상 파라미터들에 포함될 수 있다. 이러한 방식으로, 대화 향상이 어느 채널들에 대해 수행되어야만 하는지가 디코더에 시그널링될 수 있다. 대안적으로, 이러한 정보가 디코더에서 이용가능할 수 있고 - 예컨대, 하드코딩되어 있음 -, 이는 대화 향상 파라미터들이 향상 동일한 채널들과 관련하여 정의된다는 것을 의미한다. 상세하게는, 본 방법은 대화 향상된 신호들 중 어느 신호들이 믹싱을 거쳐야만 하는지를 나타내는 정보를 수신하는 단계를 추가로 포함할 수 있다. 예를 들어, 이 변형에 다른 방법은 특정의 모드에서 동작하는 디코딩 시스템에 의해 수행될 수 있고, 여기서 대화 향상된 신호들은 대화 향상된 신호들을 제공하기 위해 사용되었던 것과 완전히 동일한 다운믹스 신호들의 세트에 다시 믹싱되지 않는다. 이러한 방식으로, 믹싱 연산이 실제로는 복수의 다운믹스 신호들의 서브셋의 비전체 선택(selection)(하나 이상의 신호)으로 제한될 수 있다. 다른 대화 향상된 신호들이, 포맷 변환을 거친 다운믹스 신호들과 같은, 약간 상이한 다운믹스 신호들에 추가된다. 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋을 식별하는 데이터가 정의되고 다운믹싱 방식이 알려져 있으면, 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋이 다운믹싱되는 복수의 다운믹스 신호들의 서브셋을 찾아내는 것이 가능하다. 더욱 상세하게는, 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋을 식별하는 데이터가 다운믹싱 방식과 함께 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋이 다운믹싱되는 복수의 다운믹스 신호들의 서브셋을 찾아내는 데 사용될 수 있다.
- [0032] 복수의 다운믹스 신호들의 서브셋을 업믹싱하는 단계, 대화 향상을 적용하는 단계, 및 믹싱하는 단계가, 각각, 재구성 파라미터들, 대화 향상 파라미터들, 및 믹싱 파라미터들에 의해 정의되는 행렬 연산들로서 수행될 수 있다. 이것은 본 방법이 행렬 곱셈을 수행하는 것에 의해 효율적인 방식으로 구현될 수 있다는 점에서 유리하다.
- [0033] 더욱이, 본 방법은 복수의 다운믹스 신호들의 서브셋에 적용하기 전에 복수의 다운믹스 신호들의 서브셋을 업믹싱하는 단계, 대화 향상을 적용하는 단계, 및 믹싱하는 단계에 대응하는 행렬 연산들을, 행렬 곱셈에 의해, 단일의 행렬 연산으로 결합하는 단계를 포함할 수 있다. 이와 같이, 상이한 행렬 연산들이 단일의 행렬 연산으로

결합될 수 있고, 따라서 추가로 효율을 개선시키고 방법의 계산 복잡도를 감소시킬 수 있다.

- [0034] 대화 향상 파라미터들 및/또는 재구성 파라미터들은 주파수 의존적일 수 있고, 따라서 파라미터들을 상이한 주파수 대역들 간에 상이하게 할 수 있다. 이러한 방식으로, 대화 향상 및 재구성이 상이한 주파수 대역들에서 최적화될 수 있고, 그로써 출력 오디오의 품질을 개선시킬 수 있다.
- [0035] 더욱 상세하게는, 대화 향상 파라미터들이 제1 주파수 대역 세트(set of frequency bands)와 관련하여 정의될 수 있고, 재구성 파라미터들은 제2 주파수 대역 세트와 관련하여 정의될 수 있으며, 제2 주파수 대역 세트는 제1 주파수 대역 세트와 상이하다. 이것은, 예컨대, 재구성 프로세스가 대화 향상 프로세스보다 더 높은 주파수 분해능으로 파라미터들을 필요로 할 때, 및/또는, 예컨대, 대화 향상 프로세스가 재구성 프로세스보다 더 작은 대역폭에 대해 수행될 때, 대화 향상 파라미터들 및 재구성 파라미터들을 비트스트림으로 전송하기 위한 비트레이트를 감소시키는 데 있어서 유리할 수 있다.
- [0036] 예시적인 실시예들에 따르면, 대화 향상 파라미터들의 (바람직하게는 이산적인) 값들이 반복하여 수신되고, 각자의 값들이 정확히 적용되는, 제1 시간 순간 세트(set of time instants)와 연관될 수 있다. 본 개시내용에서, 값이 특정 시간 순간에 "정확히" 적용되거나 알려진다는 취지의 언급은 값이, 전형적으로 값이 적용되는 시간 순간의 명시적 또는 암시적 표시와 함께 디코더에 의해 수신되었다는 것을 의미하는 것으로 의도되어 있다. 이와 달리, 특정 시간 순간에 대해 보간되거나 예측되는 값은 이 의미에서 그 시간 순간에 "정확히" 적용되지 않고, 디코더측 추정치이다. "정확히"는 값이 오디오 신호의 정확한 재구성을 달성한다는 것을 암시하지 않는다. 세트 내의 연속적인 시간 순간들 사이에서, 미리 정의된 제1 보간 패턴이 미리 정해질 수 있다. 파라미터의 값들이 알려져 있는 세트 내의 2개의 경계 시간 순간들 사이에 위치한 시간 순간에서의 파라미터의 대략적인 값을 어떻게 추정할지를 정의하는, 보간 패턴은, 예를 들어, 선형(linear) 또는 구간별 상수(piecewise constant) 보간일 수 있다. 예측 시간 순간이 경계 시간 순간들 중 하나로부터 특정 거리 떨어져 위치되는 경우, 선형 보간 패턴은 예측 시간 순간에서의 파라미터의 값이 상기 거리에 선형적으로 의존한다는 가정에 기초하는 반면, 구간별 상수 보간 패턴은 파라미터의 값이 각각의 알려진 값과 다음 값 사이에서 변하지 않는다는 것을 보장한다. 예를 들어, 주어진 예측 시간 순간에서의 파라미터의 값을 추정하기 위해 1 초과 차수의 다항식, 스플라인, 유리 함수, 가우시안 프로세스, 삼각 다항식, 웨이블릿, 또는 이들의 조합을 사용하는 패턴들을 비롯한, 다른 가능한 보간 패턴들도 있을 수 있다. 시간 순간 세트가 명시적으로 전송되거나 언급되지 않을 수 있고 그 대신에 보간 패턴, 예컨대, 오디오 처리 알고리즘의 프레임 경계들에 암시적으로 고정되어 있을 수 있는, 선형 보간 구간의 시작점 또는 끝점으로부터 추론될 수 있다. 재구성 파라미터들이 유사한 방식으로 수신될 수 있다: 재구성 파라미터들의 (바람직하게는 이산적인) 값들은 제2 시간 순간 세트와 연관될 수 있고, 제2 보간 패턴은 연속적인 시간 순간들 사이에서 수행될 수 있다.
- [0037] 본 방법은 파라미터 유형 - 유형은 대화 향상 파라미터들 또는 재구성 파라미터들 중 어느 하나임 - 을, 선택된 유형과 연관된 시간 순간 세트가 비선택된 유형과 연관된 세트에는 없는 시간 순간인 적어도 하나의 예측 순간을 포함하는 방식으로, 선택하는 단계를 추가로 포함할 수 있다. 예를 들어, 재구성 파라미터들과 연관되어 있는 시간 순간 세트가 대화 향상 파라미터들과 연관되어 있는 시간 순간 세트에는 없는 특정 시간 순간을 포함하는 경우, 선택된 유형의 파라미터들이 재구성 파라미터들이고 비선택된 유형의 파라미터들이 대화 향상 파라미터들이면 특정 시간 순간이 예측 순간일 것이다. 유사한 방식으로, 다른 상황에서, 예측 순간이 그 대신에 대화 향상 파라미터들과 연관되어 있는 시간 순간 세트에서 발견될 수 있고, 선택된 유형과 비선택된 유형이 전환될 것이다. 바람직하게는, 선택된 파라미터 유형은 연관된 파라미터 값들을 갖는 시간 순간들의 밀도가 가장 높은 유형이고; 주어진 사용 사례에서, 이것은 필요한 예측 연산들의 총량을 감소시킬 수 있다.
- [0038] 비선택된 유형의 파라미터들의 값이, 예측 순간에, 예측될 수 있다. 보간 또는 외삽과 같은, 적당한 예측 방법을 사용하여 그리고 파라미터 유형들에 대한 미리 정의된 보간 패턴을 고려하여 예측이 수행될 수 있다.
- [0039] 본 방법은, 적어도 비선택된 유형의 파라미터들의 예측된 값 및 선택된 유형의 파라미터들의 수신된 값에 기초하여, 예측 순간에서 적어도 다운믹스 신호들의 서브셋의 업믹싱 및 그에 뒤이은 대화 향상을 나타내는 결합 처리 연산을 계산하는 단계를 포함할 수 있다. 재구성 파라미터들 및 대화 향상 파라미터들의 값들에 부가하여, 계산은, 믹싱에 대한 파라미터 값들과 같은, 다른 값들에 기초할 수 있고, 결합 처리 연산은 대화 향상된 신호를 다시 다운믹스 신호에 믹싱하는 단계도 나타낼 수 있다.
- [0040] 본 방법은, 적어도 선택된 유형의 파라미터들의 (수신된 또는 예측된) 값 및 적어도 비선택된 유형의 파라미터들의 (수신된 또는 예측된) 값에 기초하여, 적어도 값들 중 어느 하나가 수신된 값이도록, 선택된 유형 또는 비선택된 유형과 연관된 세트 내의 인접한 시간 순간에서 결합 처리 연산을 계산하는 단계를 포함할 수 있다. 인

접한 시간 순간은 예측 순간보다 이르거나 늦을 수 있고, 인접한 시간 순간이 거리의 면에서 최근접 이웃일 필요가 있는 것이 필수적인 것은 아니다.

- [0041] 본 방법에서, 복수의 다운믹스 신호들의 서브셋을 업믹싱하는 단계 및 대화 향상을 적용하는 단계는 계산된 결합 처리 연산의 보간된 값에 의해 예측 순간과 인접한 시간 순간 사이에서 수행될 수 있다. 계산된 결합 처리 연산을 보간하는 것에 의해, 감소된 계산 복잡도가 달성될 수 있다. 파라미터 유형들 둘 다를 개별적으로 보간하지 않는 것에 의해, 그리고 곱(즉, 결합 처리 연산)을 형성하지 않는 것에 의해, 각각의 보간 지점에서, 인지된 청취 품질의 면에서 똑같이 유용한 결과를 달성하기 위해 보다 적은 수학적 덧셈 및 곱셈 연산들이 필요하게 될 수 있다.
- [0042] 추가의 예시적인 실시예들에 따르면, 인접한 시간 순간에서의 결합 처리 연산은 선택된 유형의 파라미터들의 수신된 값 및 비선택된 유형의 파라미터들의 예측된 값에 기초하여 계산될 수 있다. 인접한 시간 순간에서의 결합 처리 연산이 선택된 유형의 파라미터들의 예측된 값 및 비선택된 유형의 파라미터들의 수신된 값에 기초하여 계산될 수 있는, 정반대 상황이 또한 가능하다. 동일한 파라미터 유형의 값이 예측 순간에서의 수신된 값 및 인접한 시간 순간에서의 예측된 값인 상황들은, 예를 들어, 선택된 파라미터 유형과 연관되어 있는 세트 내의 시간 순간들이 비선택된 파라미터 유형과 연관되어 있는 세트 내의 시간 순간들 사이에 정확히 위치되어 있는 경우에 일어날 수 있다.
- [0043] 예시적인 실시예들에 따르면, 인접한 시간 순간에서의 결합 처리 연산은 선택된 파라미터 유형의 파라미터들의 수신된 값 및 비선택된 파라미터 유형의 파라미터들의 수신된 값에 기초하여 계산될 수 있다. 이러한 상황은, 예컨대, 프레임 경계들에 대해 또한 - 선택된 유형에 대해 - 경계들 사이의 중간에 있는 시간 순간에 대해 양 유형의 파라미터들의 정확한 값들이 수신되는 경우에 일어날 수 있다. 그러면, 인접한 시간 순간은 프레임 경계와 연관된 시간 순간이고, 예측 시간 순간은 프레임 경계들 사이의 중간에 위치되어 있다.
- [0044] 추가의 예시적인 실시예들에 따르면, 본 방법은, 제1 및 제2 보간 패턴들에 기초하여, 미리 정의된 선택 규칙에 따라 결합 보간 패턴을 선택하는 단계를 추가로 포함할 수 있고, 여기서 계산된 각자의 결합 처리 연산들의 보간은 결합 보간 패턴에 따른다. 제1 및 제2 보간 패턴들이 똑같은 경우에 대해 미리 정의된 선택 규칙이 정의될 수 있고, 제1 및 제2 보간 패턴들이 상이한 경우에 대해서도 미리 정의된 선택 규칙이 정의될 수 있다. 일 예로서, 제1 보간 패턴이 선형이고(그리고 바람직하게는, 대화 향상 연산의 파라미터들과 정량적 속성들 간에 선형 관계가 있는 경우) 제2 보간 패턴이 구간별 상수인 경우에, 결합 보간 패턴이 선형적인 것으로 선택될 수 있다.
- [0045] 예시적인 실시예들에 따르면, 예측 순간에 비선택된 유형의 파라미터들의 값을 예측하는 것은 비선택된 유형의 파라미터들에 대한 보간 패턴에 따라 행해진다. 이것은, 예측 순간에 인접한 비선택된 유형과 연관된 세트 내의 시간 순간에서, 비선택된 유형의 파라미터의 정확한 값을 사용하는 것을 수반할 수 있다.
- [0046] 예시적인 실시예들에 따르면, 결합 처리 연산은 단일의 행렬 연산으로서 계산되고 이어서 복수의 다운믹스 신호들의 서브셋에 적용된다. 바람직하게는, 업믹싱하는 단계와 대화 향상을 적용하는 단계는 재구성 파라미터들 및 대화 향상 파라미터들에 의해 정의되는 행렬 연산들로서 수행된다. 결합 보간 패턴으로서, 선형 보간 패턴이 선택될 수 있고, 계산된 각자의 결합 처리 연산들의 보간된 값이 선형 행렬 보간에 의해 계산될 수 있다. 계산 복잡도를 감소시키기 위해, 보간이 예측 순간과 인접한 시간 순간 사이에서 변하는 이러한 행렬 요소들로 제한될 수 있다.
- [0047] 예시적인 실시예들에 따르면, 수신된 다운믹스 신호들이 시간 프레임들로 세그먼트화될 수 있고, 본 방법은, 정상 상태 동작에서, 각각의 시간 프레임 내의 시간 순간에서 정확히 적용되는 각자의 파라미터 유형들의 적어도 하나의 값을 수신하는 단계를 포함할 수 있다. 본 명세서에서 사용되는 바와 같이, "정상 상태"란, 예컨대, 노래의 처음 부분과 마지막 부분의 존재를 수반하지 않는 동작 및 프레임 세분(sub-division)을 필요로 하는 내부적 과도기들을 수반하지 않는 동작을 지칭한다.
- [0048] 제2 양태에 따르면, 제1 양태의 방법을 수행하기 위한 명령어들을 갖는 컴퓨터 판독가능 매체를 포함하는 컴퓨터 프로그램 제품이 제공된다. 컴퓨터 판독가능 매체는 비일시적 컴퓨터 판독가능 매체 또는 디바이스일 수 있다.
- [0049] 제3 양태에 따르면, 오디오 시스템에서 대화를 향상시키는 디코더가 제공되고, 디코더는
- [0050] 보다 많은 복수의 채널들의 다운믹스인 복수의 다운믹스 신호들,

- [0051] 대화 향상 파라미터들 - 파라미터들은 대화를 포함하는 채널들을 포함하는 복수의 채널들의 서브셋과 관련하여 정의되고, 복수의 채널들의 서브셋은 복수의 다운믹스 신호들의 서브셋으로 다운믹싱됨 -, 및
- [0052] 복수의 다운믹스 신호들의 서브셋으로 다운믹싱되는 채널들의 파라미터적 재구성을 가능하게 하는 재구성 파라미터들을 수신하도록 구성된 수신 컴포넌트;
- [0053] 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋을 재구성하기 위해 재구성 파라미터들에 기초하여 복수의 다운믹스 신호들의 서브셋을 업믹싱하도록 구성된 업믹싱 컴포넌트; 및
- [0054] 적어도 하나의 대화 향상된 신호를 제공하기 위해 대화 향상 파라미터들을 사용하여 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋에 대화 향상을 적용하도록 구성된 대화 향상 컴포넌트; 및
- [0055] 복수의 다운믹스 신호들의 서브셋의 대화 향상된 버전들을 제공하기 위해 적어도 하나의 대화 향상된 신호에 믹싱을 가하도록 구성된 믹싱 컴포넌트를 포함한다.
- [0056] 일반적으로, 제2 양태와 제3 양태는 제1 양태와 동일한 특징들 및 장점들을 포함할 수 있다.
- [0057] II. 예시적인 실시예들
- [0058] 도 1a 및 도 1b는 3개의 전방 채널(L, C, R), 2개의 서라운드 채널(LS, RS), 2개의 후방 채널(LB, RB), 4개의 고도 채널(elevated channel)(TFL, TFR, TBL, TBR), 및 저주파 효과 채널(LFE)을 갖는 (7.1+4 스피커 구성에 대응하는) 7.1+4 채널 구성을 개략적으로 나타내고 있다. 7.1+4 채널 구성을 인코딩하는 프로세스에서, 채널들은 전형적으로 다운믹싱 - 즉, 다운믹스 신호들이라고 지칭되는, 보다 적은 수의 신호들로 결합 - 된다. 다운믹싱 프로세스에서, 상이한 다운믹스 구성들을 형성하기 위해 채널들이 상이한 방식으로 결합될 수 있다. 도 1a는 다운믹스 신호들(l, c, r, ls, rs, lfe)을 갖는 제1 5.1 다운믹스 구성(100a)을 나타내고 있다. 도면에서의 원들은 어느 채널들이 어느 다운믹스 신호들로 다운믹싱되는지를 나타낸다. 도 1b는 다운믹스 신호들(l, c, r, tl, tr, lfe)을 갖는 제2 5.1 다운믹스 구성(100b)을 나타내고 있다. 제2 5.1 다운믹스 구성(100b)은 채널들이 상이한 방식으로 결합된다는 점에서 제1 5.1 다운믹스 구성(100a)과 상이하다. 예를 들어, 제1 다운믹스 구성(100a)에서는, L 채널과 TFL 채널이 l 다운믹스 신호로 다운믹싱되는 반면, 제2 다운믹스 구성(100b)에서는, L 채널, LS 채널, LB 채널이 l 다운믹스 신호로 다운믹싱된다. 다운믹스 구성은 때때로 어느 채널들이 어느 다운믹스 신호들로 다운믹싱되는지를 기술하는 다운믹싱 방식이라고 본원에서 지칭된다. 다운믹싱 구성, 또는 다운믹싱 방식은 오디오 코딩 시스템의 시간 프레임들 사이에서 변할 수 있다는 점에서 동적일 수 있다. 예를 들어, 제1 다운믹싱 방식(100a)은 어떤 시간 프레임들에서 사용될 수 있는 반면, 제2 다운믹싱 방식(100b)은 다른 시간 프레임들에서 사용될 수 있다. 다운믹싱 방식이 동적으로 변하는 경우에, 인코더는 채널들을 인코딩할 때 어느 다운믹싱 방식이 사용되었는지를 나타내는 데이터를 디코더로 송신할 수 있다.
- [0059] 도 2는 대화 향상을 위한 종래 기술의 디코더(200)를 나타내고 있다. 디코더는 3개의 주 컴포넌트인, 수신 컴포넌트(202), 업믹싱 또는 재구성 컴포넌트(204), 및 대화 향상(DE) 컴포넌트(206)를 포함한다. 디코더(200)는 복수의 다운믹스 신호들(212)을 수신하고, 수신된 다운믹스 신호들(212)에 기초하여 전체 채널 구성(218)을 재구성하며, 전체 채널 구성(218), 또는 적어도 그의 서브셋과 관련하여 대화 향상을 수행하고, 대화 향상된 채널들(220)의 전체 구성을 출력하는 유형이다.
- [0060] 더욱 상세하게는, 수신 컴포넌트(202)는 인코더로부터 데이터 스트림(210)(때때로 비트 스트림이라고 지칭됨)을 수신하도록 구성된다. 데이터 스트림(210)은 상이한 유형들의 데이터를 포함할 수 있고, 수신 컴포넌트(202)는 수신된 데이터 스트림(210)을 상이한 유형들의 데이터로 디코딩할 수 있다. 이 경우에, 데이터 스트림은 복수의 다운믹스 신호들(212), 재구성 파라미터들(214), 및 대화 향상 파라미터들(216)을 포함한다.
- [0061] 업믹싱 컴포넌트(204)는 이어서 복수의 다운믹스 신호들(212) 및 재구성 파라미터들(214)에 기초하여 전체 채널 구성을 재구성한다. 환언하면, 업믹싱 컴포넌트(204)는 다운믹스 신호들(212)로 다운믹싱되었던 모든 채널들(218)을 재구성한다. 예를 들어, 업믹싱 컴포넌트(204)는 재구성 파라미터들(214)에 기초하여 전체 채널 구성을 파라미터적으로 재구성할 수 있다.
- [0062] 예시된 예에서, 다운믹스 신호들(212)은 도 1a 및 도 1b의 5.1 다운믹스 구성들 중 하나의 5.1 다운믹스 구성의 다운믹스 신호들에 대응하고, 채널들(218)은 도 1a 및 도 1b의 7.1+4 채널 구성의 채널들에 대응한다. 그렇지만, 디코더(200)의 원리들이 물론 다른 채널 구성들/다운믹스 구성들에 적용될 것이다.
- [0063] 재구성된 채널들(218), 또는 적어도 재구성된 채널들(218)의 서브셋은 이어서 대화 향상 컴포넌트(206)에 의한 대화 향상을 거친다. 예를 들어, 대화 향상 컴포넌트(206)는, 대화 향상된 채널들을 출력하기 위해, 재구성된

채널들(218) 또는 적어도 재구성된 채널들(218)의 서브셋에 대해 행렬 연산을 수행할 수 있다. 이러한 행렬 연산은 전형적으로 대화 향상 파라미터들(216)에 의해 정의된다.

[0064] 예로서, 대화 향상 컴포넌트(206)는 대화 향상된 채널들(C_{DE} , L_{DE} , R_{DE})을 제공하기 위해 채널들(C , L , R)에 대화 향상을 가할 수 있는 반면, 다른 채널들은 도 2에 파선들로 표시된 바와 같이 단지 통과될 뿐이다. 이러한 상황에서, 대화 향상 파라미터들은 단지 C , L , R 채널들과 관련하여, 즉 복수의 채널들(218)의 서브셋과 관련하여 정의된다. 예를 들어, 대화 향상 파라미터들(216)은 C , L , R 채널들에 적용될 수 있는 3×3 행렬을 정의할 수 있다.

[0065]
$$\begin{bmatrix} C_{DE} \\ L_{DE} \\ R_{DE} \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \cdot \begin{bmatrix} C \\ L \\ R \end{bmatrix}$$

[0066] 대안적으로, 대화 향상에 관여되지 않은 채널들은 대응하는 대각선 위치들에 1을 갖고 대응하는 행들 및 열들에 있는 모든 다른 요소들에 0을 갖는 대화 향상 행렬에 의해 통과될 수 있다.

[0067]
$$\begin{bmatrix} C_{DE} \\ L_{DE} \\ TFL \\ R_{DE} \\ TFR \\ LS \\ TBL \\ LB \\ RS \\ TBR \\ RB \\ LFE \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & 0 & m_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ m_{21} & m_{22} & 0 & m_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ m_{31} & m_{32} & 0 & m_{33} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} C \\ L \\ TFL \\ R \\ TFR \\ LS \\ TBL \\ LB \\ RS \\ TBR \\ RB \\ LFE \end{bmatrix}$$

[0068] 대화 향상 컴포넌트(206)는 상이한 모드들에 따라 대화 향상을 수행할 수 있다. 본원에서 채널 독립적 파라미터적 향상(channel independent parametric enhancement)이라고 지칭되는, 제1 모드가 도 3에 예시되어 있다. 적어도 재구성된 채널들(218)의 서브셋, 전형적으로 대화를 포함하는 채널들, 여기서 채널들(L , R , C)과 관련하여 대화 향상이 수행된다. 대화 향상 파라미터들(216)은 향상될 채널들 각각에 대한 파라미터 세트를 포함한다. 예시된 예에서, 파라미터 세트들은 채널들(L , R , C)에, 각각, 대응하는 파라미터들(p_1 , p_2 , p_3)에 의해 주어진다. 원칙적으로, 이 모드에서 전송되는 파라미터들은, 채널에서의 시간-주파수 타일에 대해, 믹스 에너지(mix energy)에 대한 대화의 상대적 기여도(relative contribution)를 나타낸다. 게다가, 대화 향상 프로세스에 관여된 이득 인자(g)가 있다. 이득 인자(g)는 다음과 같이 표현될 수 있고:

[0069]
$$g = 10^{\frac{G}{20}} - 1$$

[0070] 여기서 G 는 dB로 표현되는 대화 향상 이득이다. 대화 향상 이득(G)은, 예를 들어, 사용자에게 의해 입력될 수 있고, 따라서 전형적으로 도 2의 데이터 스트림(210)에 포함되지 않는다.

[0071] 채널 독립적 파라미터적 향상 모드에 있을 때, 대화 향상 컴포넌트(206)는, 대화 향상된 채널들(220), 여기서 L_{DE} , R_{DE} , C_{DE} 를 생성하기 위해, 각각의 채널을 그의 대응하는 파라미터(p_i) 및 이득 인자(g)와 곱하고, 이어서 그 결과를 채널에 가산한다. 행렬 표기법을 사용하여, 이것은 다음과 같이 쓰여질 수 있고:

[0072]
$$X_e = (I + \text{diag}(p) \cdot g) \cdot X$$

[0073] 여기서 X 는 채널들(218)(L , R , C)을 행들로서 가지는 행렬이고, X_e 는 대화 향상된 채널들(220)을 행들로서 가지는 행렬이며, p 는 각각의 채널에 대한 대화 향상 파라미터들(p_1 , p_2 , p_3)에 대응하는 엔트리들을 갖는 행 벡터이고, $\text{diag}(p)$ 는 대각선 상에 p 의 엔트리들을 가지는 대각 행렬이다.

[0074] 본원에서 다채널 대화 예측(multichannel dialog prediction)이라고 지칭되는, 제2 대화 향상 모드는 도 4에

예시되어 있다. 이 모드에서, 대화 향상 컴포넌트(206)는 대화 신호(419)를 예측하기 위해 다수의 채널들(218)을 선형 결합으로 결합한다. 다수의 채널들에서의 대화의 존재의 코히런트 가산(coherent addition) 이외에, 이 접근법은 대화를 갖지 않는 다른 채널을 사용하여 대화를 포함하는 채널에서의 배경 잡음을 감산하는 것으로부터 이득을 볼 수 있다. 이를 위해, 대화 향상 파라미터들(216)은 선형 결합을 형성할 때 대응하는 채널의 계수를 정의하는 각각의 채널(218)에 대한 파라미터를 포함한다. 예시된 예에서, 대화 향상 파라미터들(216)은 L, R, C 채널들에, 각각, 대응하는 파라미터들(p_1, p_2, p_3)을 포함한다. 전형적으로, MMSE(minimum mean square error) 최적화 알고리즘들이 인코더측에서의 예측 파라미터들을 발생시키는 데 사용될 수 있다.

[0075] 대화 향상 컴포넌트(206)는 이어서, 대화 향상된 채널들(220)을 생성하기 위해, 이득 인자(g)의 적용에 의해 예측된 대화 신호(419)를 향상시키고(즉, 그에 이득을 부여하고), 향상된 대화 신호를 채널들(218)에 가산할 수 있다. 향상된 대화 신호를 올바른 공간 위치에 있는 올바른 채널들에 가산하기 위해(그렇지 않으면, 그것이 예상된 이득으로 대화를 향상시키지 않을 것임), 3개의 채널들 사이의 패닝(panning)이 렌더링 계수들, 여기서 r_1, r_2, r_3 에 의해 전송된다. 렌더링 계수들이 에너지 보존(energy preserving), 즉

[0076]
$$r_1^2 + r_2^2 + r_3^2 = 1$$

[0077] 이라는 제한 하에서, 세번째 렌더링 계수(r_3)는 $r_3 = \sqrt{1 - r_1^2 - r_2^2}$ 이도록 처음 2개의 계수들로부터 결정될 수 있다.

[0078] 행렬 표기법을 사용하여, 다채널 대화 예측 모드에 있을 때 대화 향상 컴포넌트(206)에 의해 수행되는 대화 향상은 다음과 같이 쓰여질 수 있고:

[0079]
$$X_e = (I + g \cdot H \cdot P) \cdot X$$

[0080] 또는

[0081]
$$X_e = \begin{bmatrix} 1 + g \cdot r_1 \cdot p_1 & g \cdot r_1 \cdot p_2 & g \cdot r_1 \cdot p_3 \\ g \cdot r_2 \cdot p_1 & 1 + g \cdot r_2 \cdot p_2 & g \cdot r_2 \cdot p_3 \\ g \cdot r_3 \cdot p_1 & g \cdot r_3 \cdot p_2 & 1 + g \cdot r_3 \cdot p_3 \end{bmatrix} \cdot X$$

[0082] 여기서 I는 항등 행렬이고, X는 채널들(218)(L, R, C)을 행들로서 가지는 행렬이며, X_e 는 대화 향상된 채널들(220)을 행들로서 가지는 행렬이고, P는 각각의 채널에 대한 대화 향상 파라미터들(p_1, p_2, p_3)에 대응하는 엔트리들을 갖는 행 벡터이며, H는 렌더링 계수들(r_1, r_2, r_3)을 엔트리들로서 가지는 열 벡터이고, g 는

[0083]
$$g = 10^{\frac{G}{20}} - 1$$
 을 갖는 이득 인자이다.

[0083] 본원에서 파형 파라미터적 하이브리드(waveform-parametric hybrid)라고 지칭되는 제3 모드에 따르면, 대화 향상 컴포넌트(206)는 제1 모드 및 제2 모드 중 어느 하나를 대화를 나타내는 부가의 오디오 신호(파형 신호)의 전송과 결합시킬 수 있다. 후자는 전형적으로 저 비트레이트로 코딩되어, 개별적으로 청취될 때 잘 들리는 아티팩트들을 야기한다. 채널들(218) 및 대화의 신호 속성들과, 대화 파형 신호 코딩에 할당된 비트레이트에 따라, 인코더는 또한 이득 기여도들이 (제1 또는 제2 모드로부터의) 파라미터적 기여도와 대화를 나타내는 부가의 오디오 신호 사이에 어떻게 분배되어야만 하는지를 나타내는 블렌딩 파라미터(blending parameter)(a_c)를 결정한다.

[0084] 제2 모드와 결합하여, 제3 모드의 대화 향상은 다음과 같이 쓰여질 수 있고:

[0085]
$$X_e = H \cdot g_1 \cdot d_c + (I + H \cdot g_2 \cdot P) \cdot X$$

[0086] 또는

$$X_e = \begin{bmatrix} 1 + g_2 \cdot r_1 \cdot p_1 & g_2 \cdot r_1 \cdot p_2 & g_2 \cdot r_1 \cdot p_3 & g_1 \cdot r_1 \\ g_2 \cdot r_2 \cdot p_1 & 1 + g_2 \cdot r_2 \cdot p_2 & g_2 \cdot r_2 \cdot p_3 & g_1 \cdot r_2 \\ g_2 \cdot r_3 \cdot p_1 & g_2 \cdot r_3 \cdot p_2 & 1 + g_2 \cdot r_3 \cdot p_3 & g_1 \cdot r_3 \end{bmatrix} \cdot \begin{bmatrix} X \\ d_c \end{bmatrix}$$

[0088] 여기서 d_c 는 대화를 나타내는 부가의 오디오 신호이고,

$$g_1 = \alpha_c \cdot \left(10^{\frac{G}{20}} - 1\right) \text{ 이고,}$$

$$g_2 = (1 - \alpha_c) \cdot \left(10^{\frac{G}{20}} - 1\right) \text{ 이다.}$$

[0091] 채널 독립적 향상(제1 모드)과의 결합을 위해, 각각의 채널(218)에 대해 대화를 나타내는 오디오 신호($d_{c,i}$)가

$$D_c = \begin{pmatrix} d_{c,1} \\ d_{c,2} \\ d_{c,3} \end{pmatrix}$$

수신된다. 이라고 하면, 대화 향상은 다음과 같이 쓰여질 수 있다.

$$X_e = g_1 \cdot D_c + (I + \text{diag}(p) \cdot g_2) \cdot X$$

[0093] 도 5는 예시적인 실시예들에 따른 디코더(500)를 나타내고 있다. 디코더(500)는 차후의 재생을 위해, 보다 많은 복수의 채널들의 다운믹스인, 복수의 다운믹스 신호들을 디코딩하는 유형이다. 환언하면, 디코더(500)는, 전체 채널 구성을 재구성하도록 구성되어 있지 않다는 점에서, 도 2의 디코더와 상이하다.

[0094] 디코더(500)는 수신 컴포넌트(502)와, 업믹싱 컴포넌트(504), 대화 향상 컴포넌트(506), 및 믹싱 컴포넌트(508)를 포함하는 대화 향상 블록(503)을 포함한다.

[0095] 도 2를 참조하여 설명된 바와 같이, 수신 컴포넌트(502)는 데이터 스트림(510)을 수신하고, 이를 그의 컴포넌트들, 이 경우에, 보다 많은 복수의 채널들의 다운믹스(도 1a 및 도 1b를 참조)인 복수의 다운믹스 신호들(512), 재구성 파라미터들(514), 및 대화 향상 파라미터들(516)로 디코딩한다. 어떤 경우에, 데이터 스트림(510)은 믹싱 파라미터들(522)을 나타내는 데이터를 추가로 포함한다. 예를 들어, 믹싱 파라미터들은 대화 향상 파라미터들의 일부를 형성할 수 있다. 다른 경우에, 믹싱 파라미터들(522)은 디코더(500)에서 이미 이용가능하고, 예컨대, 디코더(500)에 하드코딩되어 있을 수 있다. 다른 경우에, 믹싱 파라미터들(522)이 다수의 믹싱 파라미터 세트들에 대해 이용가능하고, 데이터 스트림(510) 내의 데이터는 이 다수의 믹싱 파라미터 세트들 중 어느 세트가 사용되는지에 대한 표시를 제공한다.

[0096] 대화 향상 파라미터들(516)은 전형적으로 복수의 채널들의 서브셋과 관련하여 정의된다. 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋을 식별하는 데이터는 수신된 데이터 스트림(510)에, 예를 들어, 대화 향상 파라미터들(516)의 일부로서 포함될 수 있다. 대안적으로, 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋이 디코더(500)에 하드코딩되어 있을 수 있다. 예를 들어, 도 1a를 참조하면, 대화 향상 파라미터들(516)은 1 다운믹스 신호로 다운믹스되는 L, TFL 채널들, c 다운믹스 신호에 포함되는 C 채널, 및 r 다운믹스 신호로 다운믹스되는 R, TFR 채널들과 관련하여 정의될 수 있다. 예시를 위해, 대화가 L, C, 및 R 채널들에만 존재하는 것으로 가정된다. 유의할 점은, 대화 향상 파라미터들(516)이, L, C, R 채널들과 같은, 대화를 포함하는 채널들과 관련하여 정의될 수 있지만, 또한, 이 예에서 TFL, TFR 채널들과 같은, 대화를 포함하지 않는 채널들과 관련하여 정의될 수 있다는 것이다. 그러한 방식으로, 대화를 포함하는 채널에서의 배경 잡음이, 예를 들어, 대화를 갖지 않는 다른 채널을 사용하여 감소될 수 있다.

[0097] 대화 향상 파라미터들(516)이 정의되는 채널들의 서브셋이 복수의 다운믹스 신호들(512)의 서브셋(512a)으로 다운믹스된다. 예시된 예에서, 다운믹스 신호들의 서브셋(512a)은 c, l, 및 r 다운믹스 신호들을 포함한다. 다운믹스 신호들의 이 서브셋(512a)은 대화 향상 블록(503)에 입력된다. 다운믹스 신호들의 관련 서브셋(512a)은, 예컨대, 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋 및 다운믹싱 방식에 대한 지식에 기초하여 발견될 수 있다.

- [0098] 업믹싱 컴포넌트(514)는 다운믹스 신호들의 서브셋(512a)으로 다운믹스되는 채널들의 재구성을 위한 본 기술 분야에 공지된 바와 같은 파라미터적 기법들을 사용한다. 재구성은 재구성 파라미터들(514)에 기초한다. 상세하게는, 업믹싱 컴포넌트(504)는 대화 향상 파라미터들(516)이 정의되는 복수의 채널들의 서브셋을 재구성한다. 일부 실시예들에서, 업믹싱 컴포넌트(504)는 대화 향상 파라미터들(516)이 정의되는 복수의 채널들의 서브셋만을 재구성한다. 이러한 예시적인 실시예들이 도 7을 참조하여 기술될 것이다. 다른 실시예들에서, 업믹싱 컴포넌트(504)는 대화 향상 파라미터들(516)이 정의되는 복수의 채널들의 서브셋에 부가하여 적어도 하나의 채널을 재구성한다. 이러한 예시적인 실시예들이 도 6을 참조하여 기술될 것이다.
- [0099] 재구성 파라미터들은 시변적(time variable)일 수 있을 뿐만 아니라 주파수 의존적일 수도 있다. 예를 들어, 재구성 파라미터들은 상이한 주파수 대역들에 대해 상이한 값들을 취할 수 있다. 이것은 일반적으로 재구성된 채널들의 품질을 개선시킬 것이다.
- [0100] 본 기술 분야에 공지된 바와 같이, 파라미터적 업믹싱은 일반적으로 업믹싱을 거치는 입력 신호들로부터 역상관된 신호들을 형성하는 것을 포함할 수 있고, 입력 신호들 및 역상관된 신호들에 기초하여 신호들을 파라미터적으로 재구성할 수 있다. 예를 들어, 문헌["Spatial Audio Processing: MPEG Surround and Other Applications" by Jeroen Breebaart and Christof Faller, ISBN:978-9-470-03350-0]을 참조한다. 그렇지만, 업믹싱 컴포넌트(504)는 바람직하게는 임의의 이러한 역상관된 신호들을 사용하지 않고 파라미터적 업믹싱을 수행한다. 역상관된 신호들을 사용함으로써 얻어지는 장점들은, 이 경우에, 믹싱 컴포넌트(508)에서 수행되는 차후의 다운믹싱에 의해 감소된다. 따라서, 역상관된 신호들의 사용이 유리하게도 업믹싱 컴포넌트(504)에 의해 생략될 수 있고, 그로써 계산 복잡도를 절감할 수 있다. 사실상, 업믹싱에서의 역상관된 신호들의 사용은, 대화 향상과 조합하여, 품질 악화를 가져올 수 있는데, 그 이유는 그로 인해 대화에 대한 역상관기 리버브를 가져올 수 있기 때문이다.
- [0101] 대화 향상 컴포넌트(506)는 이어서 적어도 하나의 대화 향상된 신호를 생성하기 위해 대화 향상 파라미터들(516)이 정의되는 복수의 채널들의 서브셋에 대화 향상을 적용한다. 일부 실시예들에서, 대화 향상된 신호는 대화 향상 파라미터들(516)이 정의되는 복수의 채널들의 서브셋의 대화 향상된 버전들에 대응한다. 이것이 도 6을 참조하여 이하에서 더욱 상세하게 설명될 것이다. 다른 실시예들에서, 대화 향상된 신호는 대화 향상 파라미터들(516)이 정의되는 복수의 채널들의 서브셋의 예측된 및 향상된 대화 컴포넌트에 대응한다. 이것이 도 7을 참조하여 이하에서 더욱 상세하게 설명될 것이다.
- [0102] 재구성 파라미터들과 유사하게, 대화 향상 파라미터들도 시간에서는 물론 주파수에 따라 변할 수 있다. 더욱 상세하게는, 대화 향상 파라미터들은 상이한 주파수 대역들에 대해 상이한 값들을 취할 수 있다. 재구성 파라미터들이 정의되는 주파수 대역들의 세트는 대화 향상 파라미터들이 정의되는 주파수 대역들의 세트와 상이할 수 있다.
- [0103] 믹싱 컴포넌트(508)는 이어서 다운믹스 신호들의 서브셋(512a)의 대화 향상된 버전들(520)을 제공하기 위해 적어도 하나의 대화 향상된 신호에 기초하여 믹싱을 수행한다. 예시된 예에서, 다운믹스 신호들의 서브셋(512a)의 대화 향상된 버전들(520)은 다운믹스 신호들(c, l, r)에, 각각, 대응하는 c_{DE} , l_{DE} , r_{DE} 에 의해 주어진다.
- [0104] 다운믹스 신호들의 서브셋(512a)의 대화 향상된 버전들(520)에 대한 적어도 하나의 대화 향상된 신호의 기여도를 나타내는 믹싱 파라미터들(522)에 따라 믹싱이 행해질 수 있다. 일부 실시예들에서, 도 6을 참조하고, 적어도 하나의 대화 향상된 신호가 업믹싱 컴포넌트(504)에 의해 재구성된 채널들과 함께 믹싱된다. 이러한 경우에, 믹싱 파라미터들(522)은, 각각의 채널이 대화 향상된 다운믹스 신호들(520) 중 어느 것에 믹싱되어야만 하는지를 나타내는, 다운믹싱 방식 - 도 1a 및 도 1b를 참조 - 에 대응할 수 있다. 다른 실시예들에서, 도 7을 참조하고, 적어도 하나의 대화 향상된 신호가 다운믹스 신호들의 서브셋(512a)과 함께 믹싱된다. 이러한 경우에, 믹싱 파라미터들(522)은 적어도 하나의 대화 향상된 신호가 어떻게 다운믹스 신호들의 서브셋(512a)으로 가중되어야 하는지를 나타내는 가중 인자들에 대응할 수 있다.
- [0105] 업믹싱 컴포넌트(504)에 의해 수행되는 업믹싱 연산, 대화 향상 컴포넌트(506)에 의해 수행되는 대화 향상 연산, 및 믹싱 컴포넌트(508)에 의해 수행되는 믹싱 연산은 전형적으로 각각이 행렬 연산에 의해, 즉 행렬-벡터 곱에 의해 정의될 수 있는 선형 연산들이다. 역상관된 신호들이 업믹싱 연산에서 생략되는 경우에 적어도 그러하다. 상세하게는, 업믹싱 연산과 연관된 행렬(U)은 재구성 파라미터들(514)에 의해 정의되고/그로부터 도출될 수 있다. 이와 관련하여, 유의할 점은, 업믹싱 연산에서의 역상관된 신호들의 사용이 여전히 가능하지만 역상관된 신호들의 생성이 그러면 업믹싱을 위한 행렬 연산의 일부가 아니라는 것이다. 역상관기들에 의한 업믹싱

연산은 2-스태이지 접근법으로 볼 수 있다. 제1 스테이지에서, 입력 다운믹스 신호들이 전치 역상관기 행렬(pre-decorrelator matrix)에 피드되고, 전치 역상관기 행렬의 적용 이후의 출력 신호들 각각이 역상관기에 피드된다. 제2 스테이지에서, 입력 다운믹스 신호들 및 역상관기들로부터의 출력 신호들이 업믹스 행렬에 피드되고, 여기서 입력 다운믹스 신호들에 대응하는 업믹스 행렬의 계수들은 "드라이 업믹스 행렬(dry upmix matrix)"이라고 지칭되는 것을 형성하고, 역상관기들로부터의 출력 신호들에 대응하는 계수들은 "웨트 업믹스 행렬(wet upmix matrix)"이라고 지칭되는 것을 형성한다. 각각의 하위 행렬은 업믹스 채널 구성에 매핑된다. 역상관기 신호들이 사용되지 않을 때, 업믹싱 연산과 연관된 행렬은 입력 신호들(512a)에 대한 연산만을 위해 구성되고, 역상관된 신호들에 관련된 열들(웨트 업믹스 행렬)은 행렬에 포함되지 않는다. 환언하면, 업믹스 행렬은 이 경우에 드라이 업믹스 행렬에 대응한다. 그렇지만, 앞서 살펴본 바와 같이, 역상관기 신호들의 사용은 이 경우에 전형적으로 품질 악화를 가져올 것이다.

[0106] 대화 향상 연산과 연관된 행렬(M)은 대화 향상 파라미터들(516)에 의해 정의되고/그로부터 도출될 수 있으며, 믹싱 연산과 연관된 행렬(C)은 믹싱 파라미터들(522)에 의해 정의되고/그로부터 도출될 수 있다.

[0107] 업믹싱 연산, 대화 향상 연산, 및 믹싱 연산이 모두 선형 연산이기 때문에, 대응하는 행렬들은, 행렬 곱셈에 의해, 단일의 행렬 E로 결합될 수 있다(그러면 $X_{DE}=E \cdot X$, 여기서, $E=C \cdot M \cdot U$ 임). 여기서 X는 다운믹스 신호들(512a)의 열 벡터이고, X_{DE} 는 대화 향상된 다운믹스 신호들(520)의 열 벡터이다. 이와 같이, 대화 향상 블록(503) 전체가 다운믹스 신호들의 서브셋(512a)의 대화 향상된 버전들(520)을 생성하기 위해 다운믹스 신호들의 서브셋(512a)에 적용되는 단일의 행렬 연산에 대응할 수 있다. 그에 따라, 본원에 기술되는 방법들이 아주 효율적인 방식으로 구현될 수 있다.

[0108] 도 6은 도 5의 디코더(500)의 예시적인 실시예에 대응하는 디코더(600)를 나타내고 있다. 디코더(600)는 수신 컴포넌트(602), 업믹싱 컴포넌트(604), 대화 향상 컴포넌트(606), 및 믹싱 컴포넌트(608)를 포함한다.

[0109] 도 5의 디코더(500)와 유사하게, 수신 컴포넌트(602)는 데이터 스트림(610)을 수신하고 이를 복수의 다운믹스 신호들(612), 재구성 파라미터들(614), 및 대화 향상 파라미터들(616)로 디코딩한다.

[0110] 업믹싱 컴포넌트(604)는 복수의 다운믹스 신호들(612)의 서브셋(612a)(서브셋(512a)에 대응함)을 수신한다. 서브셋(612a) 내의 다운믹스 신호들 각각에 대해, 업믹싱 컴포넌트(604)는 다운믹스 신호에 다운믹싱되었던 모든 채널들을 재구성한다($X_u=U \cdot X$). 이것은 대화 향상 파라미터들이 정의되는 채널들(618a), 및 대화 향상에 관여되도록 되어 있지 않은 채널들(618b)을 포함한다. 도 1b를 참조하면, 대화 향상 파라미터들이 정의되는 채널들(618a)은 예를 들어 L, LS, C, R, RS 채널들에 대응할 수 있을 것이고, 대화 향상에 관여되도록 되어 있지 않은 채널들(618b)은 LB, RB 채널들에 대응할 수 있다.

[0111] 대화 향상 파라미터들이 정의되는 채널들(618a)(X'_u)은 이어서 대화 향상 컴포넌트(606)에 의해 대화 향상을 거치는 반면($X_e = M \cdot X'_u$), 대화 향상에 관여되도록 되어 있지 않은 채널들(618b)(X''_u)은 대화 향상 컴포넌트(606)를 바이패스한다.

[0112] 대화 향상 컴포넌트(606)는 앞서 기술된 대화 향상의 제1, 제2, 및 제3 모드들 중 임의의 것을 적용할 수 있다. 제3 모드가 적용되는 경우에, 데이터 스트림(610)은, 앞서 설명된 바와 같이, 대화 향상 파라미터들이 정의되는 복수의 채널들의 서브셋(618a)과 함께 대화 향상에서 적용될 대화를 나타내는 오디오 신호(즉, 대화를 나타내는

코딩된 과형)를 포함할 수 있다
$$\begin{pmatrix} X_e \\ X''_u \end{pmatrix} = M \cdot \begin{pmatrix} X'_u \\ D_c \end{pmatrix}$$

[0113] 그 결과, 대화 향상 컴포넌트(606)는, 이 경우에 대화 향상 파라미터들이 정의되는 채널들의 서브셋(618a)의 대화 향상된 버전들에 대응하는, 대화 향상된 신호들(619)을 출력한다. 예로서, 대화 향상된 신호들(619)은 도 1b의 L, LS, C, R, RS 채널들의 대화 향상된 버전들에 대응할 수 있다.

[0114] 믹싱 컴포넌트(608)는 이어서, 다운믹스 신호들의 서브셋(612a)의 대화 향상된 버전들(620)을 생성하기 위해,

$$X_{DE} = C \cdot \begin{pmatrix} X_e \\ X''_u \end{pmatrix}$$
 대화 향상된 신호들(619)을 대화 향상에 관여되었던 채널들(618b)과 함께 믹싱한다.

믹싱 컴포넌트(608)는, 도 1b에 예시된 다운믹싱 방식과 같은, 현재 다운믹싱 방식에 따라 믹싱을 행한다. 이 경우에, 믹싱 파라미터들(622)은 따라서 각각의 채널(619, 618b)이 어느 다운믹스 신호(620)에 믹싱되어야 하는지를

기술하는 다운믹싱 방식에 대응한다. 다운믹싱 방식은 정적일 수 있고 따라서 디코더(600)에 알려져 있을 수 있거나 - 이는 동일한 다운믹싱 방식이 항상 적용된다는 것을 의미함 -, 다운믹싱 방식이 동적일 수 있다 - 이는 프레임마다 달라질 수 있거나, 디코더에 알려져 있는 몇 개의 방식들 중 하나일 수 있다는 것을 의미함 -. 후자의 경우에, 다운믹싱 방식에 관한 표시가 데이터 스트림(610)에 포함된다.

[0115] 도 6에서, 디코더는 임의적인 리셔플 컴포넌트(reshuffle component)(630)를 갖추고 있다. 리셔플 컴포넌트(630)는 상이한 다운믹싱 방식들 간에 변환하는 데, 예컨대, 방식(100b)을 방식(100a)으로 변환하는 데 사용될 수 있다. 유의할 점은, 리셔플 컴포넌트(630)가 전형적으로 c 및 lfe 신호들을 변하지 않은 채로 놓아둔다 - 즉, 이 신호들과 관련하여 통과 컴포넌트(pass-through component)로서 기능함 - 는 것이다. 리셔플 컴포넌트(630)는, 예를 들어, 재구성 파라미터들(614) 및 대화 향상 파라미터들(616)과 같은 다양한 파라미터들을 수신하고 그에 기초하여 동작(도시되지 않음)할 수 있다.

[0116] 도 7은 도 5의 디코더(500)의 예시적인 실시예에 대응하는 디코더(700)를 나타내고 있다. 디코더(700)는 수신 컴포넌트(702), 업믹싱 컴포넌트(704), 대화 향상 컴포넌트(706), 및 믹싱 컴포넌트(708)를 포함한다.

[0117] 도 5의 디코더(500)와 유사하게, 수신 컴포넌트(702)는 데이터 스트림(710)을 수신하고 이를 복수의 다운믹스 신호들(712), 재구성 파라미터들(714), 및 대화 향상 파라미터들(716)로 디코딩한다.

[0118] 업믹싱 컴포넌트(704)는 복수의 다운믹스 신호들(712)의 서브셋(712a)(서브셋(512a)에 대응함)을 수신한다. 도 6과 관련하여 기술된 실시예와 달리, 업믹싱 컴포넌트(704)는 대화 향상 파라미터들(716)이 정의되는 복수의 채널들의 서브셋(718a)만을 재구성한다($X'_u = U' \cdot X$). 도 1b를 참조하면, 대화 향상 파라미터들이 정의되는 채널들(718a)은 예를 들어 C, L, LS, R, RS 채널들에 대응할 수 있을 것이다.

[0119] 대화 향상 컴포넌트(706)는 이어서 대화 향상 파라미터들이 정의되는 채널들(718a)에 대해 대화 향상을 수행한다($X_d = M_d \cdot X'_u$). 이 경우에, 대화 향상 컴포넌트(706)는, 제2 대화 향상 모드에 따라, 채널들(718a)의 선형 결합을 형성하는 것에 의해 채널들(718a)에 기초하여 대화 컴포넌트를 계속하여 예측한다. 도 7에서 p_1 내지 p_5 에 의해 표시된, 선형 결합을 형성할 때 사용되는 계수들이 대화 향상 파라미터들(716)에 포함된다. 예측된 대화 컴포넌트가 이어서 대화 향상된 신호(719)를 생성하기 위해 이득 인자(g)를 곱하는 것에 의해 향상된다. 이득 인자(g)는 다음과 같이 표현될 수 있고:

[0120]
$$g = 10^{\frac{G}{20}} - 1$$

[0121] 여기서 G는 dB로 표현되는 대화 향상 이득이다. 대화 향상 이득(G)은, 예를 들어, 사용자에게 의해 입력될 수 있고, 따라서 전형적으로 데이터 스트림(710)에 포함되지 않는다. 유의할 점은, 몇 개의 대화 컴포넌트들이 있는 경우에, 이상의 예측 및 향상 절차가 대화 컴포넌트당 한 번씩 적용될 수 있다는 것이다.

[0122] 예측된 대화 향상된 신호(719)(즉, 예측된 및 향상된 대화 컴포넌트들)가 이어서 다운믹스 신호들의 서브셋(712a)의 대화 향상된 버전들(720)을 생성하기 위해 다운믹스 신호들의 서브셋(712a)에 믹싱된다

$(X_{DE} = C \cdot \begin{bmatrix} X_d \\ X \end{bmatrix})$. 다운믹스 신호들의 서브셋의 대화 향상된 버전들(720)에 대한 대화 향상된 신호(719)의 기여도를 나타내는 믹싱 파라미터들(722)에 따라 믹싱이 행해진다. 믹싱 파라미터들은 전형적으로 데이터 스트림(710)에 포함된다. 이 경우에, 믹싱 파라미터들(722)은 적어도 하나의 대화 향상된 신호(719)가 어떻게 다운믹스 신호들의 서브셋(712a)으로 가중되어야 하는지를 나타내는 가중 인자들(r_1, r_2, r_3)에 대응한다.

[0123]
$$X_{DE} = X + \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \cdot X_d = \begin{bmatrix} r_1 & 1 & 0 & 0 \\ r_2 & 0 & 1 & 0 \\ r_3 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \bar{X}_d \\ X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

[0124] 더욱 상세하게는, 가중 인자들은, 대화 향상된 신호(719)가 올바른 공간 위치들에서 다운믹스 신호들(712a)에 가산되도록, 다운믹스 신호들의 서브셋(712a)과 관련하여 적어도 하나의 대화 향상된 신호(719)의 패닝을 기술하는 렌더링 계수들에 대응할 수 있다.

- [0125] 데이터 스트림(710) 내의 렌더링 계수들(믹싱 파라미터들(722))은 업믹싱된 채널들(718a)에 대응할 수 있다. 예시된 예에서, 5개의 업믹싱된 채널들(718a)이 있고, 따라서 5개의 대응하는 렌더링 계수들(말하자면, rc_1, rc_2, \dots, rc_5)이 있을 수 있다. (다운믹스 신호들(712a)에 대응하는) r_1, r_2, r_3 의 값들은 이어서, 다운믹싱 방식과 결합하여, rc_1, rc_2, \dots, rc_5 로부터 계산될 수 있다. 채널들(718a) 중 다수가 동일한 다운믹스 신호(712a)에 대응할 때, 대화 렌더링 계수들이 합산될 수 있다. 예를 들어, 예시된 예에서, $r_1=rc_1, r_2=rc_2+rc_3$, 및 $r_3=rc_4+rc_5$ 이 성립한다. 채널들의 다운믹싱이 다운믹싱 계수들을 사용하여 행해진 경우에 이것은 또한 가중합(weighted summation)일 수 있다.
- [0126] 유의할 점은, 또한 이 경우에 대화 향상 컴포넌트(706)가 대화를 나타내는 부가적으로 수신된 오디오 신호를 사용할 수 있다는 것이다. 이러한 경우에, 예측된 대화 향상된 신호(719)는 믹싱 컴포넌트(708)에 입력되기 전에 대화를 나타내는 오디오 신호와 함께 가중될 수 있다
- $$(X_d = (1 - \alpha_c) \cdot M_d \cdot X'_u + \alpha_c \cdot g \cdot D_c)$$
- . 적절한 가중이 대화 향상 파라미터들(716)에 포함된 블렌딩 파라미터(α_c)에 의해 주어진다. 블렌딩 파라미터(α_c)는 이득 기여도들이 (앞서 기술된 바와 같은) 예측된 대화 컴포넌트(719)와 대화(D_c)를 나타내는 부가의 오디오 신호 사이에서 어떻게 분배되어야만 하는지를 나타낸다. 이것은 제2 대화 향상 모드와 결합될 때 제3 대화 향상 모드와 관련하여 기술된 것과 유사하다.
- [0127] 도 7에서, 디코더는 임의적인 리서플 컴포넌트(730)를 갖추고 있다. 리서플 컴포넌트(730)는 상이한 다운믹싱 방식들 간에 변환하는 데, 예컨대, 방식(100b)을 방식(100a)으로 변환하는 데 사용될 수 있다. 유의할 점은, 리서플 컴포넌트(730)가 전형적으로 c 및 lfe 신호들을 변하지 않은 채로 놓아둔다 - 즉, 이 신호들과 관련하여 통과 컴포넌트로서 기능함 - 는 것이다. 리서플 컴포넌트(730)는, 예를 들어, 재구성 파라미터들(714) 및 대화 향상 파라미터들(716)과 같은 다양한 파라미터들을 수신하고 그에 기초하여 동작(도시되지 않음)할 수 있다.
- [0128] 이상의 내용은 주로 7.1+4 채널 구성 및 5.1 다운믹스와 관련하여 설명되었다. 그렇지만, 본원에 기술되는 디코더들 및 디코딩 방법들의 원리들이 다른 채널 및 다운믹스 구성들에 똑같이 잘 적용된다는 것을 잘 알 것이다.
- [0129] 도 8은 디코더로 전송하기 위한 데이터 스트림(810)을 생성하기 위해 복수의 채널들(818) - 그 중 일부가 대화를 포함함 - 을 인코딩하는 데 사용될 수 있는 인코더(800)의 예시이다. 인코더(800)는 디코더들(200, 500, 600, 700) 중 임의의 것과 함께 사용될 수 있다. 인코더(800)는 다운믹싱 컴포넌트(805), 대화 향상 인코딩 컴포넌트(806), 파라미터적 인코딩 컴포넌트(804), 및 전송 컴포넌트(802)를 포함한다.
- [0130] 인코더(800)는 복수의 채널들(818), 예컨대, 도 1a 및 도 1b에 도시된 채널 구성들(100a, 100b)의 채널들을 수신한다.
- [0131] 다운믹싱 컴포넌트(805)는 복수의 채널들(818)을 복수의 다운믹스 신호들(812)로 다운믹싱하고, 복수의 다운믹스 신호들(812)은 이어서 데이터 스트림(810)에 포함시키기 위해 전송 컴포넌트(802)에 피드된다. 복수의 채널들(818)은, 예컨대, 도 1a에 또는 도 1b에 예시된 것과 같은, 다운믹싱 방식에 따라 다운믹싱될 수 있다.
- [0132] 복수의 채널들(818) 및 다운믹스 신호들(812)은 파라미터적 인코딩 컴포넌트(804)에 입력된다. 그의 입력 신호들에 기초하여, 파라미터적 인코딩 컴포넌트(804)는 다운믹스 신호들(812)로부터 채널들(818)을 재구성하는 것을 가능하게 하는 재구성 파라미터들(814)을 계산한다. 재구성 파라미터들(814)은, 예컨대, 본 기술 분야에 공지된 바와 같은 MMSE(minimum mean square error) 최적화 알고리즘들을 사용하여 계산될 수 있다. 재구성 파라미터들(814)은 이어서 데이터 스트림(810)에 포함시키기 위해 전송 컴포넌트(802)에 피드된다.
- [0133] 대화 향상 인코딩 컴포넌트(806)는 하나 이상의 대화 신호들(813) 및 복수의 채널들(818) 중 하나 이상에 기초하여 대화 향상 파라미터들(816)을 계산한다. 대화 신호들(813)은 순수한 대화를 나타낸다. 주목할 만한 점은, 대화가 채널들(818) 중 하나 이상에 이미 믹싱되어 있다는 것이다. 따라서 대화 신호들(813)에 대응하는 하나 이상의 대화 컴포넌트들이 채널들(818)에 있을 수 있다. 전형적으로, 대화 향상 인코딩 컴포넌트(806)는 MMSE(minimum mean square error) 최적화 알고리즘들을 사용하여 대화 향상 파라미터들(816)을 계산한다. 이러한 알고리즘들은 복수의 채널들(818) 중 일부로부터 대화 신호들(813)을 예측하는 것을 가능하게 하는 파라미터들을 제공할 수 있다. 대화 향상 파라미터들(816)이 이와 같이 복수의 채널들(818)의 서브셋, 즉 채널들 - 이들로부터 대화 신호들(813)이 예측될 수 있음 - 과 관련하여 정의될 수 있다. 대화 예측에 대한 파라미터들(816)은 데이터 스트림(810)에 포함시키기 위해 전송 컴포넌트(802)에 피드된다.

- [0134] 결과적으로, 데이터 스트림(810)은 따라서 복수의 다운믹스 신호들(812), 재구성 파라미터들(814), 및 대화 향상 파라미터들(816)을 적어도 포함한다.
- [0135] 디코더의 정상 동작 동안, 상이한 유형의 파라미터들(대화 향상 파라미터들, 또는 재구성 파라미터들 등)의 값들이 특정 레이트들로 디코더에 의해 반복하여 수신된다. 상이한 파라미터 값들이 수신되는 레이트들이 디코더로부터의 출력이 계산되어야만 하는 레이트보다 더 낮은 경우, 파라미터들의 값들이 보간될 필요가 있을 수 있다. 일반 파라미터(p)의 값이, 시점들 t_1 및 t_2 에서, 각각, $p(t_1)$ 및 $p(t_2)$ 인 것으로 알려져 있으면, 중간 시점 $t_1 \leq t < t_2$ 에서의 파라미터의 값 $p(t)$ 는 상이한 보간 방식들을 사용하여 계산될 수 있다. 본원에서 선형 보간 패턴이라고 지칭되는, 이러한 방식의 일 예는 선형 보간을 사용하여 중간 값을 계산할 수 있다, 예컨대, $p(t) = p(t_1) + [p(t_2) - p(t_1)](t - t_1)/(t_2 - t_1)$ 이다. 본원에서 구간별 상수 보간 패턴이라고 지칭되는, 다른 패턴은 그 대신에 시간 구간 전체 동안 파라미터 값을 기지의 값들 중 하나, 예컨대, $p(t) = p(t_1)$ 또는 $p(t) = p(t_2)$ 로, 또는 예를 들어, 평균 값 $p(t) = [p(t_1) + p(t_2)]/2$ 와 같은 기지의 값들의 조합에 고정된 채로 유지하는 것을 포함할 수 있다. 특정 시간 구간 동안 특정 파라미터 유형에 대해 어떤 보간 방식이 사용되어야 하는지에 관한 정보가, 파라미터들 자체와 함께 또는 수신된 신호에 포함된 부가 정보로서와 같이, 상이한 방식들로 디코더에 내장되거나 디코더에 제공될 수 있다.
- [0136] 예시적인 예에서, 디코더는 제1 및 제2 파라미터 유형에 대한 파라미터 값들을 수신한다. 각각의 파라미터 유형의 수신된 값들은 제1 시간 순간 세트($T1=\{t11, t12, t13, \dots\}$)와 제2 시간 순간 세트($T2=\{t21, t22, t23, \dots\}$)에서, 각각, 정확히 적용가능하고, 디코더는 또한 대응하는 세트에 존재하지 않는 시간 순간에서 값이 추정될 필요가 있는 경우에 각각의 파라미터 유형의 값들이 어떻게 보간되어야 하는지에 관한 정보에 액세스할 수 있다. 파라미터 값들은 신호들에 대한 수학적 연산들 - 이 연산들은, 예를 들어, 행렬들로서 표현될 수 있음 - 의 정량적 속성들을 제어한다. 이하의 예에서, 제1 파라미터 유형에 의해 제어되는 연산이 제1 행렬 A에 의해 표현되고, 제2 파라미터 유형에 의해 제어되는 연산이 제2 행렬 B에 의해 표현되며, 용어들 "연산"과 "행렬"이 예에서 서로 바꾸어 사용될 수 있는 것으로 가정된다. 디코더로부터의 출력 값이 계산될 필요가 있는 시간 순간에서, 양 연산들의 합성(composition)에 대응하는 결합 처리 연산이 계산되어야 한다. 행렬 A가 (재구성 파라미터들에 의해 제어되는) 업믹싱의 연산이고 행렬 B가 (대화 향상 파라미터들에 의해 제어되는) 대화 향상을 적용하는 연산인 것으로 추가로 가정되는 경우, 결과적으로, 업믹싱과 그에 뒤이은 대화 향상의 결합 처리 연산이 행렬 곱 BA에 의해 표현된다.
- [0137] 결합 처리 연산을 계산하는 방법들은 도 9a 내지 도 9e에 예시되어 있고, 여기서 시간이 수평 축을 따라 있고 축 틱 표시(axis tick-mark)들은 결합 처리 연산이 계산되어야 하는 시간 순간들(출력 시간 순간들)을 나타낸다. 도면들에서, 삼각형들은 (업믹싱의 연산을 나타내는) 행렬 A에 대응하고, 원들은 (대화 향상을 적용하는 연산을 나타내는) 행렬 B에 대응하며, 정사각형들은 (업믹싱 및 그에 뒤이은 대화 향상의 결합 연산을 나타내는) 결합 연산 행렬 BA에 대응한다. 채워진 삼각형들 및 원들은 각자의 행렬이 대응하는 시간 순간에서 정확히 알려져 있다는 것(즉, 행렬이 나타내는 연산을 제어하는, 파라미터들이 정확히 알려져 있다는 것)을 나타내는 반면, 비어있는 삼각형들 및 원들은 각자의 행렬의 값이 (예컨대, 앞서 개략적으로 기술된 보간 패턴들 중 임의의 것을 사용하여) 예측되거나 보간된다는 것을 나타낸다. 채워진 정사각형은 결합 연산 행렬 BA가, 대응하는 시간 순간에, 예컨대, 행렬들 A와 B의 행렬 곱에 의해 계산되었다는 것을 나타내고, 비어있는 정사각형은 BA의 값이 이전의 시간 순간으로부터 보간되었다는 것을 나타낸다. 게다가, 파선 화살표들은 어느 시간 순간들 사이에서 보간이 수행되는지를 나타낸다. 마지막으로, 시간 순간들을 연결시키는 실선 수평 라인은 행렬의 값이 그 구간에서 구간별 상수인 것으로 가정된다는 것을 나타낸다.
- [0138] 본 발명을 사용하지 않는, 결합 처리 연산 BA를 계산하는 방법이 도 9a에 예시되어 있다. 연산들 A와 B에 대한 수신된 값들이, 각각, 시간 순간들 $t11, t21$ 및 $t12, t22$ 에서 정확히 적용되고, 각각의 출력 시간 순간에서 결합 처리 연산 행렬을 계산하기 위해, 본 방법은 각각의 행렬을 개별적으로 보간한다. 각각의 시간상 순방향 스텝(each forward step in time)을 계산하기 위해, 결합 처리 연산을 나타내는 행렬이 A와 B의 예측된 값들의 곱으로서 계산된다. 여기서, 각각의 행렬이 선형 보간 패턴을 사용하여 보간되어야 하는 것으로 가정된다. 행렬 A가 N' 개의 행들 및 N 개의 열들을 가지며, 행렬 B가 M 개의 행들 및 N' 개의 열들을 가지는 경우, 각각의 시간상 순방향 스텝은 (결합 처리 행렬 BA를 계산하는 데 필요한 행렬 곱셈을 수행하기 위해) 파라미터 대역당 $O(MN'N)$ 개의 곱셈 연산들을 필요로 할 것이다. 높은 밀도의 출력 시간 순간들 및/또는 많은 수의 파라미터 대역들은 따라서 (덧셈 연산과 비교하여 곱셈 연산의 비교적 높은 계산 복잡도로 인해) 계산 자원들을 많이 요구할 위험이 있다. 계산 복잡도를 감소시키기 위해, 도 9b에 예시된 대안의 방법이 사용될 수 있다. 파라미터

값들이 변하는 시간 순간들에서만(즉, 수신된 값들이 정확히 적용가능한 경우, t_{11} , t_{21} 및 t_{12} , t_{22} 에서) 결합 처리 연산을 계산하는 것(예컨대, 행렬 곱셈을 수행하는 것)에 의해, 행렬들 A와 B를 개별적으로 보간하는 대신에 결합 처리 연산 행렬 BA가 직접 보간될 수 있다. 그렇게 함으로써, 연산들이 행렬들에 의해 표현되는 경우, (정확한 파라미터 값들이 변하는 시간 순간들 사이의) 각각의 시간상 순방향 스텝은 파라미터 대역당 (행렬 덧셈을 위해) $O(NM)$ 개의 연산들을 필요로 할 것이고, 감소된 계산 복잡도는 계산 자원들을 덜 요구할 것이다. 또한, 행렬들 A와 B가 $N' > N \times M / (N + M)$ 이도록 되어 있는 경우, 결합 처리 연산 BA를 나타내는 행렬은 결합되는 개개의 행렬들 A와 B에서 발견되는 것보다 더 적은 요소들을 가질 것이다. 그렇지만, 행렬 BA를 직접 보간하는 방법은 A와 B 둘 다가 동일한 시간 순간들에서 알려져 있을 것을 필요로 할 것이다. A가 정의되는 시간 순간들이 B가 정의되는 시간 순간들과 (적어도 부분적으로) 상이할 때, 개선된 보간 방법이 필요하게 된다. 본 발명의 예시적인 실시예들에 따른, 이러한 개선된 방법은 도 9c 내지 도 9e에 예시되어 있다. 도 9a 내지 도 9e의 논의와 관련하여, 간략함을 위해, 결합 처리 연산 행렬 BA가 개개의 행렬들 A와 B - 이들 각각은 (수신된 또는 예측된/보간된) 파라미터 값들에 기초하여 발생되었음 - 의 곱으로서 계산되는 것으로 가정된다. 다른 상황에서, 2개의 행렬 인자들로서의 표현을 통해 전달함이 없이, 행렬 BA에 의해 표현되는 연산을 파라미터 값들로부터 직접 계산하는 것이 똑같이 또는 더 유리할 수 있다. 도 9c 내지 도 9e를 참조하여 예시된 기법들 중 임의의 것과 결합하여, 이 접근법들 각각이 본 발명의 범주 내에 속한다.

[0139] 도 9c에서, 행렬 A에 대응하는 파라미터에 대한 시간 순간들의 세트 T1이 세트 T2(행렬 B에 대응하는 파라미터에 대한 시간 순간들)에 존재하지 않는 시간 값 t_{12} 를 포함하는 상황이 예시되어 있다. 양 행렬들은 선형 보간 패턴을 사용하여 보간되어야 하고, 본 방법은 행렬 B의 값이 (예컨대, 보간을 사용하여) 예측되어야만 하는 예측 순간 $t_p=t_{12}$ 를 식별한다. 값이 구해진 후에, t_p 에서의 결합 처리 연산 행렬 BA의 값이 A와 B를 곱하는 것에 의해 계산될 수 있다. 계속하기 위해, 본 방법은 인접한 시간 순간 $t_a=t_{11}$ 에서 BA의 값을 계산하고, 이어서 t_a 와 t_p 사이에서 BA를 보간한다. 본 방법은 또한, 원하는 경우, 다른 인접한 시간 순간 $t_a=t_{13}$ 에서 BA의 값을 계산할 수 있고, t_p 와 t_a 사이에서 BA를 보간할 수 있다. ($t_p=t_{12}$ 에서의) 부가의 행렬 곱셈이 필요하게 되더라도, 본 방법은, 예컨대, 도 9a에서의 방법과 비교하여 계산 복잡도를 여전히 감소시키면서, 결합 처리 연산 행렬 BA를 직접 보간하는 것을 가능하게 한다. 앞서 언급된 바와 같이, 결합 처리 연산은 대안적으로 2개의 행렬들의 명시적 곱 - 이는 차례로 각자의 파라미터 값들에 의존함 - 으로서가 아니라 (수신된 또는 예측된/보간된) 파라미터 값들로부터 직접 계산될 수 있다.

[0140] 이전의 경우에, A에 대응하는 파라미터 유형만이 B에 대응하는 파라미터 유형의 순간들 중에 포함되지 않은 시간 순간들을 가졌다. 도 9d에서, 시간 순간 t_{12} 가 세트 T2에 없고 시간 순간 t_{22} 가 세트 T1에 없는 상이한 상황이 예시되어 있다. BA의 값이 t_{12} 와 t_{22} 사이의 중간 시간 순간 t' 에서 계산되어야 하는 경우, 본 방법은 $t_p = t_{12}$ 에서의 B의 값과 $t_a = t_{22}$ 에서의 A의 값 둘 다를 예측할 수 있다. 양 시각들에서 결합 처리 연산 행렬 BA를 계산한 후에, BA가 t' 에서의 그의 값을 구하기 위해 보간될 수 있다. 일반적으로, 본 방법은 파라미터 값들이 변하는 시간 순간들에서(즉, 수신된 값들이 정확히 적용가능한 세트들 T1 및 T2 내의 시간 순간들에서) 행렬 곱셈들을 수행할 뿐이다. 그 사이에서, 결합 처리 연산의 보간은 행렬 덧셈들 - 그들의 곱셈 대응물보다 더 적은 계산 복잡도를 가짐 - 을 필요로 할 뿐이다.

[0141] 이상의 예들에서, 모든 보간 패턴들은 선형적인 것으로 가정되었다. 파라미터들이 처음에 상이한 방식들을 사용하여 보간되어야 할 때의 보간 방법이 또한 도 9e에 예시되어 있다. 도면에서, 행렬 A에 대응하는 파라미터의 값들은 값들이 급격히 변하는 시간 순간 t_{12} 까지 구간별 상수인 것으로 유지된다. 파라미터 값들이 프레임 단위로 수신되면, 각각의 프레임은 수신된 값이 정확히 적용되는 시간 순간을 나타내는 시그널링을 담고 있을 수 있다. 이 예에서, B에 대응하는 파라미터는 t_{21} 및 t_{22} 에서 정확히 적용가능한 수신된 값들을 가질 뿐이고, 본 방법은 먼저 t_{12} 직전의 시간 순간 t_p 에서 B의 값을 예측할 수 있다. t_p , 및 $t_a = t_{11}$ 에서 결합 처리 연산 행렬 BA를 계산한 후에, 행렬 BA가 t_a 와 t_p 사이에서 보간될 수 있다. 본 방법은 이어서 새로운 예측 순간 $t_p = t_{12}$ 에서 B의 값을 예측하고, t_p 및 $t_a = t_{22}$ 에서 BA의 값들을 계산하며, t_p 와 t_a 사이에서 BA를 직접 보간할 수 있다. 한번 더 말하지만, 결합 처리 연산 BA가 구간에 걸쳐 보간되었고, 그의 값이 모든 출력 시간 순간들에서 구해졌다. A와 B가 개별적으로 보간되었고 BA가 각각의 출력 시간 순간에서 A와 B를 곱하는 것에 의해 계산되었던, 도 9a에 예시된 바와 같은, 이전의 상황과 비교하여, 감소된 수의 행렬 곱셈들이 필요하고 계산 복잡도가 저해된다.

[0142] 등가물들, 확장들, 대안들 및 기타

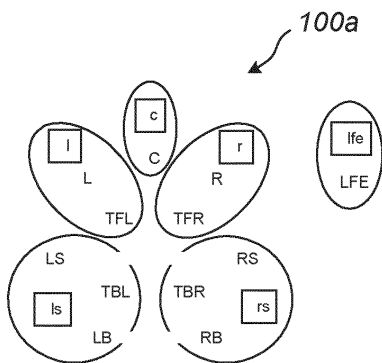
[0143] 본 개시내용의 추가의 실시예들이 이상의 설명을 살펴본 후에 본 기술 분야의 통상의 기술자에게 명백하게 될 것이다. 비록 본 설명 및 도면들이 실시예들 및 예들을 개시하고 있지만, 본 개시내용이 이 특정 예들로 제한되지 않는다. 첨부된 청구항들에 의해 정의되는 본 개시내용의 범주를 벗어남이 없이 수많은 수정들 및 변형들이 행해질 수 있다. 청구항들에서 나오는 어떤 참조 부호들도 청구항들의 범주를 제한하는 것으로 이해되어서는 안된다.

[0144] 그에 부가하여, 도면들, 개시 내용, 및 첨부된 청구항들을 살펴보는 것으로부터, 본 개시내용을 실시할 때 개시된 실시예들에 대한 변형들이 본 기술 분야의 통상의 기술자에 의해 이해되고 실시될 수 있다. 청구항들에서, 단어 "포함하는(comprising)"은 다른 요소들 또는 단계들을 배제하지 않으며, 단수 관형사 "한" 또는 "어떤"은 복수를 배제하지 않는다. 특정 대책들이 서로 다른 종속 청구항들에서 인용되고 있다는 단순한 사실이 이 대책들의 조합이 유리하게 사용될 수 없다는 것을 나타내지 않는다.

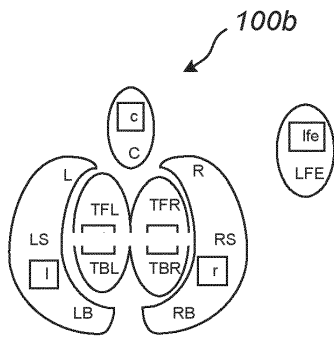
[0145] 앞서 개시된 시스템들 및 방법들은 소프트웨어, 펌웨어, 하드웨어 또는 이들의 조합으로서 구현될 수 있다. 하드웨어 구현에서, 이상의 설명에서 언급된 기능 유닛들 사이의 작업들의 분할이 꼭 물리적 유닛들로의 분할에 대응하지는 않고; 그와 달리, 하나의 물리적 컴포넌트가 다수의 기능들을 가질 수 있고, 하나의 작업이 몇 개의 물리적 컴포넌트들에 의해 협력하여 수행될 수 있다. 특정 컴포넌트들 또는 모든 컴포넌트들이 디지털 신호 프로세서 또는 마이크로프로세서에 의해 실행되는 소프트웨어로서 구현될 수 있거나, 하드웨어로서 또는 ASIC(application-specific integrated circuit)으로서 구현될 수 있다. 이러한 소프트웨어는, 컴퓨터 저장 매체(또는 비일시적 매체) 및 통신 매체(또는 일시적 매체)를 포함할 수 있는, 컴퓨터 판독가능 매체 상에 분산되어 있을 수 있다. 본 기술 분야의 통상의 기술자에게 널리 공지된 바와 같이, 용어 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어들, 데이터 구조들, 프로그램 모듈들 또는 다른 데이터와 같은 정보의 저장을 위해 임의의 방법 또는 기술로 구현되는 휘발성 및 비휘발성, 이동식 및 비이동식 매체 둘 다를 포함한다. 컴퓨터 저장 매체는 RAM, ROM, EEPROM, 플래시 메모리 또는 다른 메모리 기술, CD-ROM, DVD(digital versatile disk) 또는 다른 광학 디스크 저장소, 자기 카세트, 자기 테이프, 자기 디스크 저장소 또는 다른 자기 저장 디바이스, 또는 원하는 정보를 저장하는 데 사용될 수 있고 컴퓨터에 의해 액세스될 수 있는 임의의 다른 매체를 포함하지만, 이들로 제한되지 않는다. 게다가, 통신 매체가 전형적으로 반송파 또는 다른 전송 메커니즘과 같은 변조된 데이터 신호로 컴퓨터 판독가능 명령어들, 데이터 구조들, 프로그램 모듈들 또는 다른 데이터를 구현하고 임의의 정보 전달 매체를 포함한다는 것이 통상의 기술자에게 널리 공지되어 있다.

도면

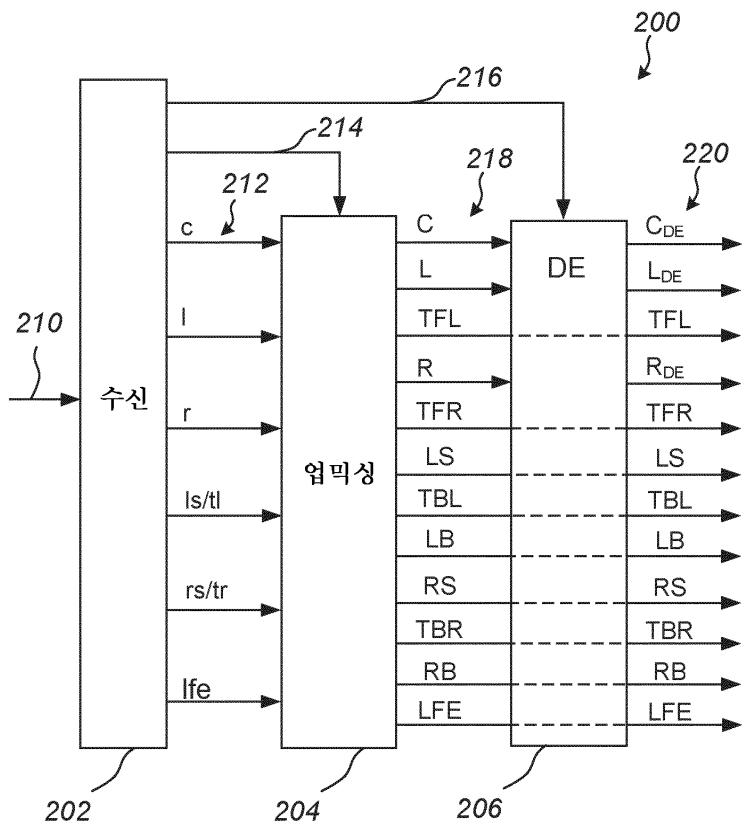
도면1a



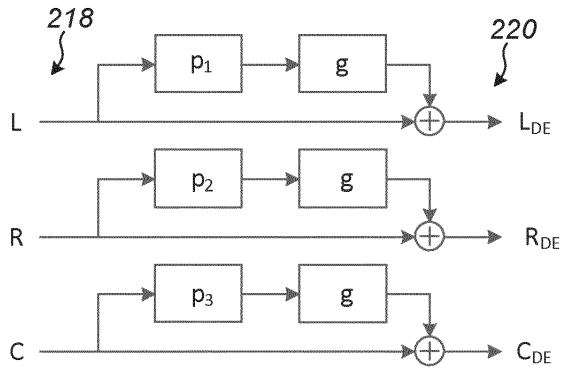
도면1b



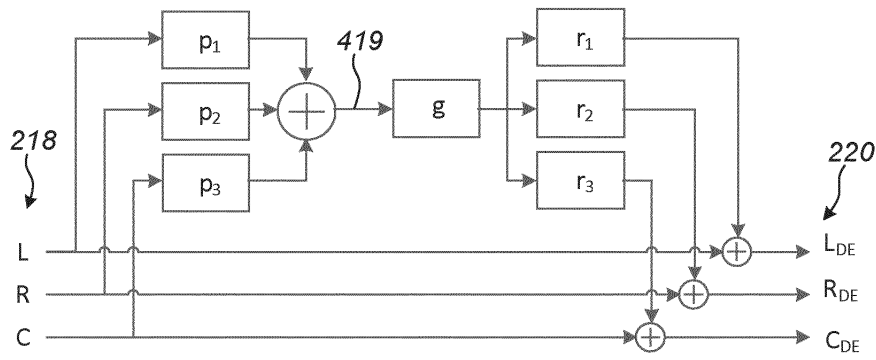
도면2



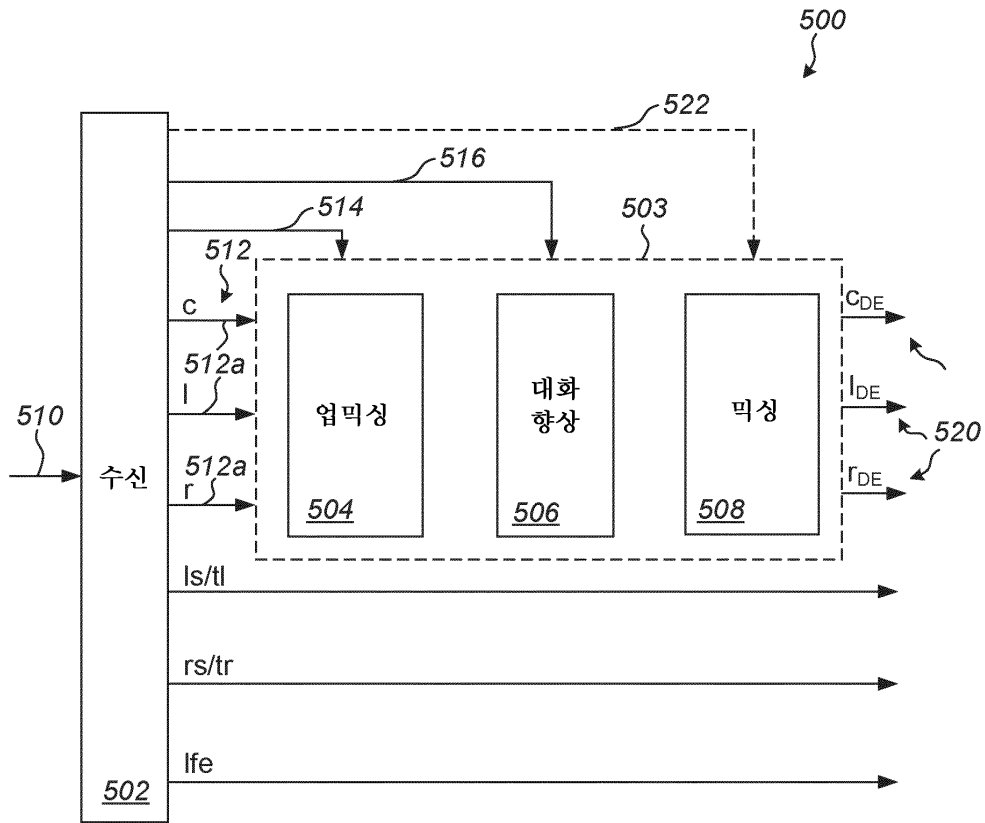
도면3



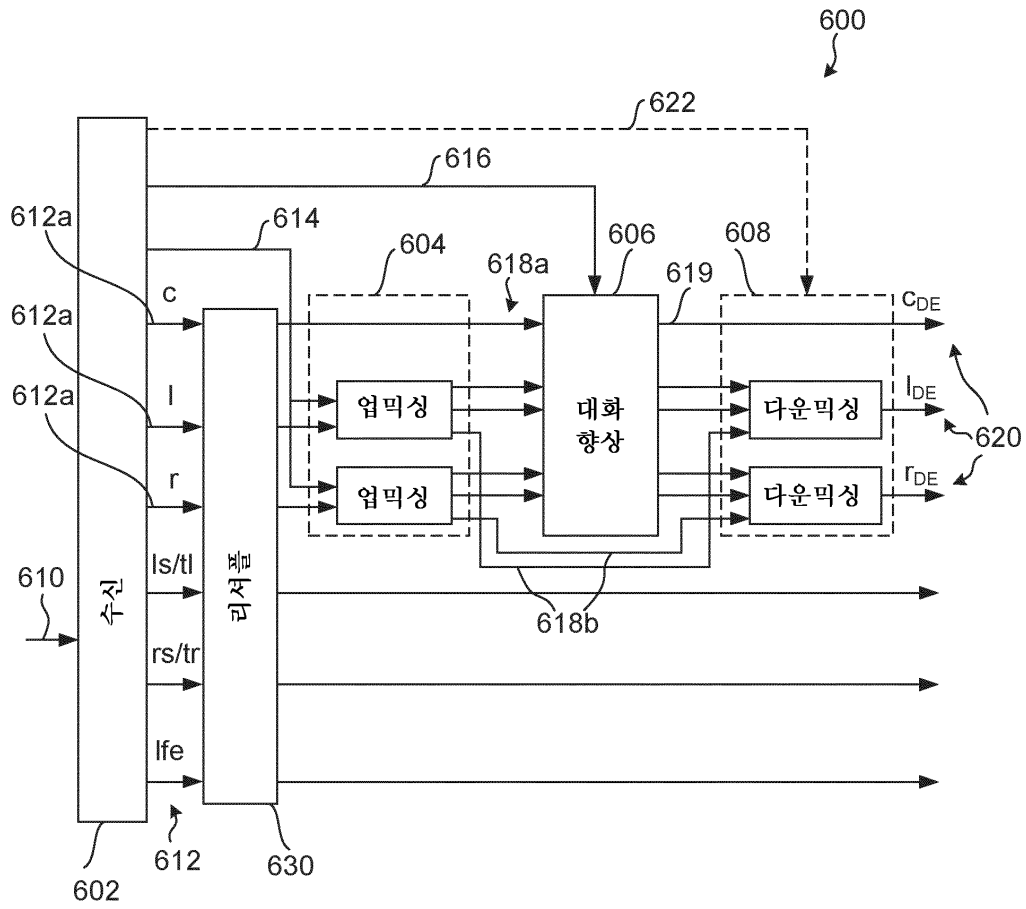
도면4



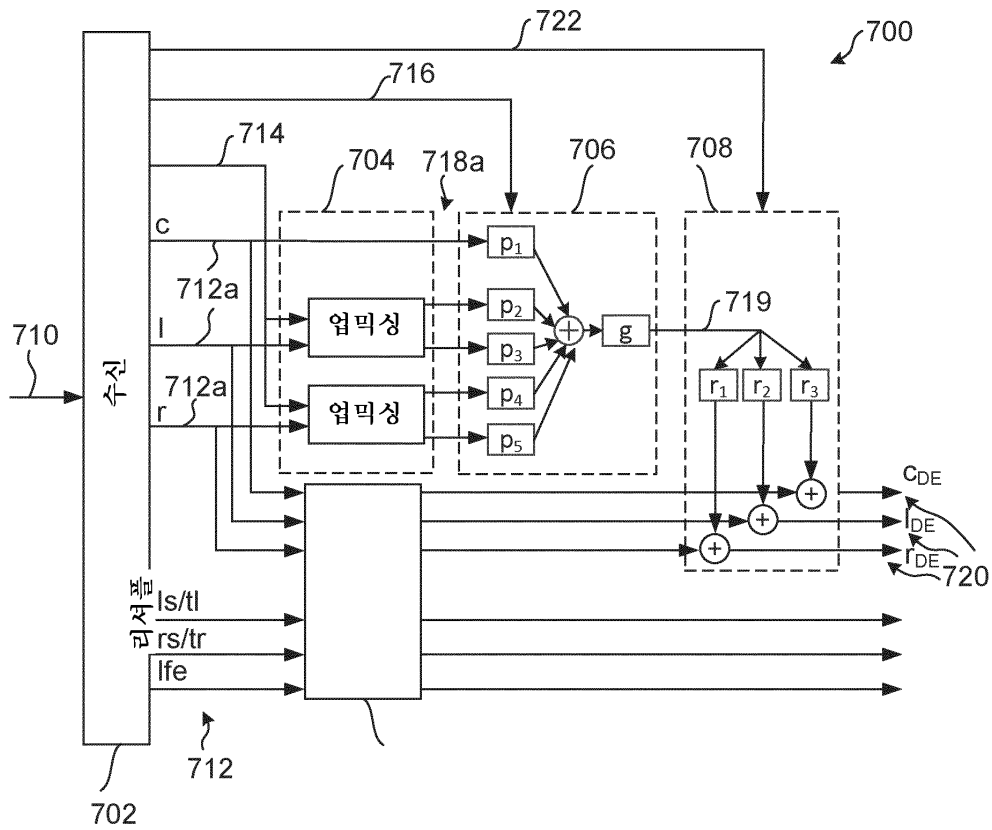
도면5



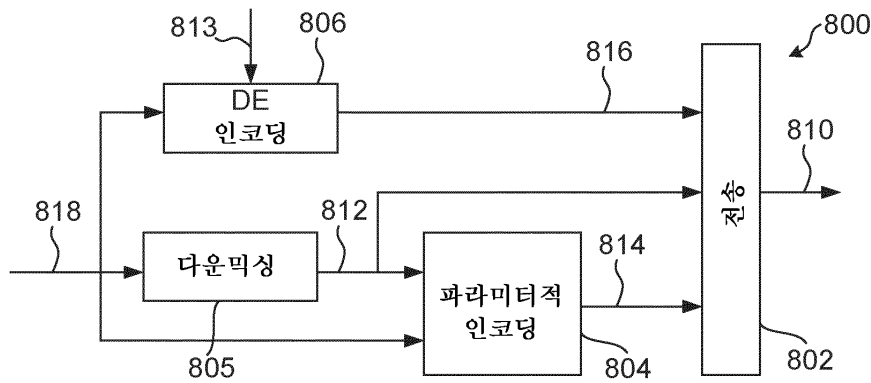
도면6



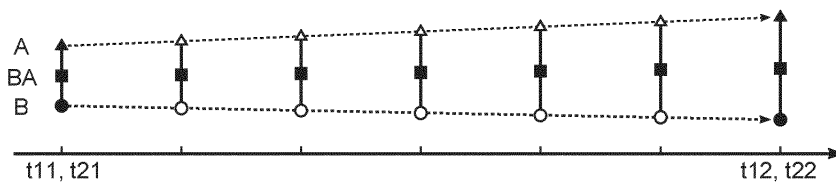
도면7



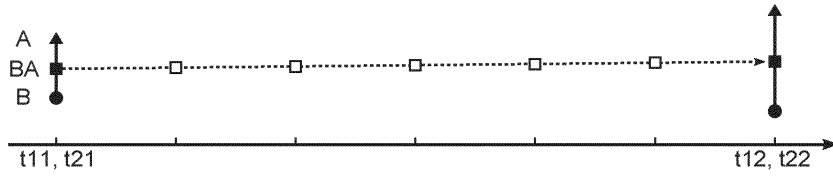
도면8



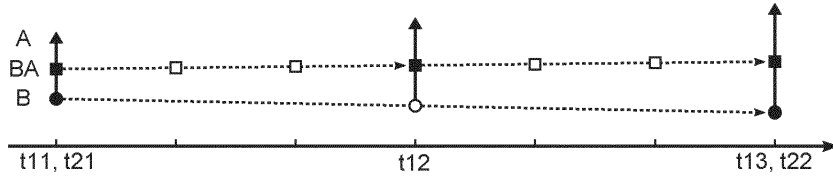
도면9a



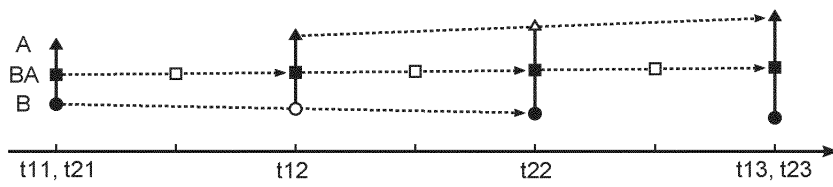
도면9b



도면9c



도면9d



도면9e

