



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년12월23일
(11) 등록번호 10-2194515
(24) 등록일자 2020년12월17일

- (51) 국제특허분류(Int. Cl.)
HO4S 1/00 (2006.01) HO4R 5/04 (2006.01)
- (52) CPC특허분류
HO4S 1/002 (2013.01)
HO4R 5/04 (2013.01)
- (21) 출원번호 10-2020-7020317
- (22) 출원일자(국제) 2018년11월29일
심사청구일자 2020년07월14일
- (85) 번역문제출일자 2020년07월14일
- (65) 공개번호 10-2020-0089339
- (43) 공개일자 2020년07월24일
- (86) 국제출원번호 PCT/US2018/063153
- (87) 국제공개번호 WO 2019/118194
국제공개일자 2019년06월20일
- (30) 우선권주장
62/599,605 2017년12월15일 미국(US)
16/151,201 2018년10월03일 미국(US)
- (56) 선행기술조사문헌
US20080260131 A1
KR1020160030911 A
KR1020140042900 A

- (73) 특허권자
봄클라우드 360, 인코포레이티드
미국 캘리포니아주 92024 앤시니타스 #311 사우스
코스트 하이웨이 101 687
- (72) 발명자
셀데스 재커리
미국 캘리포니아주 92024 앤시니타스 #311 사우스
코스트 하이웨이 101 687 봄클라우드 360, 인코포
레이티드
- (74) 대리인
제일특허법인(유)

전체 청구항 수 : 총 22 항

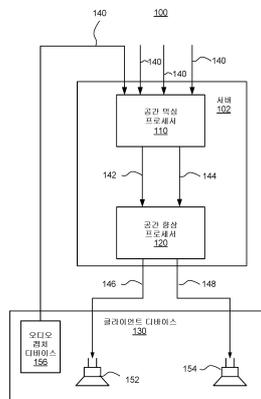
심사관 : 권영학

(54) 발명의 명칭 회의를 위한 서브밴드 공간 처리 및 크로스토크 제거 시스템

(57) 요약

실시예는 공간화된 오디오를 이용하여 클라이언트 디바이스에 회의를 제공하는 것에 관한 것이다. 입력 오디오 스트림은 클라이언트 디바이스로부터 수신된다. 각각의 클라이언트 디바이스에 대해, 음장 내에서 다른 클라이언트 디바이스들의 공간 위치를 정의하는 배치 데이터가 결정된다. 배치 데이터에 따라 다른 클라이언트 디바이스들의 입력 오디오 스트림을 믹싱 및 패닝함으로써, 클라이언트 디바이스에 대해 믹싱된 좌측 채널 및 믹싱된 우측 채널을 포함하는 믹싱된 스트림이 생성된다. 믹싱된 스트림의 믹싱된 좌측 채널 및 믹싱된 우측 채널에 서브밴드 공간 처리 및 크로스토크 처리를 적용함으로써 좌측 스피커를 위한 향상된 좌측 채널 및 우측 스피커를 위한 향상된 우측 채널을 포함하는 공간적으로 향상된 스트림이 생성된다.

대표도 - 도1



(52) CPC특허분류
H04S 2420/01 (2013.01)

명세서

청구범위

청구항 1

복수의 클라이언트 디바이스 중 하나의 클라이언트 디바이스에 회의를 제공하는 방법으로서,

상기 복수의 클라이언트 디바이스로부터 입력 오디오 스트림을 수신하는 단계와,

상기 복수의 클라이언트 디바이스 중 상기 하나의 클라이언트 디바이스에 대해, 음장(sound field) 내에서 상기 복수의 클라이언트 디바이스 중 다른 클라이언트 디바이스들에 대한 공간 위치를 정의하는 배치 데이터를 결정하는 단계와,

상기 배치 데이터에 따라 상기 다른 클라이언트 디바이스들의 입력 오디오 스트림을 믹싱 및 패닝함으로써, 상기 하나의 클라이언트 디바이스에 대해 믹싱된 좌측 채널 및 믹싱된 우측 채널을 포함하는 믹싱된 스트림을 생성하는 단계와,

상기 믹싱된 스트림의 믹싱된 좌측 채널 및 믹싱된 우측 채널에 서브밴드 공간 처리를 적용함으로써, 향상된 좌측 채널 및 향상된 우측 채널을 포함하는 공간적으로 향상된 스트림을 생성하는 단계 - 상기 서브밴드 공간 처리는,

상기 믹싱된 좌측 채널 및 상기 믹싱된 우측 채널의 중간 컴포넌트의 중간 서브밴드 컴포넌트들에 제 1 이득을 적용하여 향상된 중간 컴포넌트를 생성하는 것과,

상기 믹싱된 좌측 채널 및 상기 믹싱된 우측 채널의 측면 컴포넌트의 측면 서브밴드 컴포넌트들에 제 2 이득을 적용하여 향상된 측면 컴포넌트를 생성하는 것과,

상기 향상된 중간 컴포넌트 및 상기 향상된 측면 컴포넌트를 사용하여 상기 향상된 좌측 채널 및 상기 향상된 우측 채널을 생성하는 것을 포함함 - 와,

상기 공간적으로 향상된 스트림의 향상된 좌측 채널을 상기 하나의 클라이언트 디바이스의 좌측 스피커에 제공하고, 상기 공간적으로 향상된 스트림의 향상된 우측 채널을 상기 하나의 클라이언트 디바이스의 우측 스피커에 제공하는 단계를 포함하는

방법.

청구항 2

제 1 항에 있어서,

상기 배치 데이터를 결정하는 단계는 네트워크를 통해 상기 하나의 클라이언트 디바이스로부터 상기 배치 데이터를 수신하는 단계를 포함하는,

방법.

청구항 3

제 1 항에 있어서,

상기 복수의 클라이언트 디바이스로부터의 입력 오디오 스트림의 각각은 하나 이상의 오디오 채널을 포함하는,

방법.

청구항 4

제 1 항에 있어서,

상기 배치 데이터에 따라 상기 다른 클라이언트 디바이스들의 입력 오디오 스트림을 믹싱 및 패닝함으로써, 상기 하나의 클라이언트 디바이스에 대해 믹싱된 좌측 채널 및 믹싱된 우측 채널을 포함하는 믹싱된 스트림을 생성하는 단계는,

상기 음장 내의 상기 다른 클라이언트 디바이스들의 공간 위치에 따라 상기 다른 클라이언트 디바이스들의 입력 오디오 스트림의 각각으로부터 좌측 채널 및 우측 채널을 생성하는 단계와,

상기 다른 클라이언트 디바이스들의 입력 오디오 스트림으로부터의 좌측 채널들을 결합하여 상기 믹싱된 좌측 채널을 생성하는 단계와,

상기 다른 클라이언트 디바이스들의 입력 오디오 스트림으로부터의 우측 채널들을 결합하여 상기 믹싱된 우측 채널을 생성하는 단계를 포함하는,

방법.

청구항 5

제 1 항에 있어서,

상기 복수의 클라이언트 디바이스에 접속된 서버는 상기 믹싱된 스트림을 생성하여 상기 믹싱된 스트림을 상기 하나의 클라이언트 디바이스에 제공하고,

상기 하나의 클라이언트 디바이스는 상기 믹싱된 스트림으로부터 상기 공간적으로 향상된 스트림을 생성하는,

방법.

청구항 6

제 1 항에 있어서,

상기 복수의 클라이언트 디바이스에 접속된 서버는 상기 믹싱된 스트림 및 상기 공간적으로 향상된 스트림을 생성하여 상기 공간적으로 향상된 스트림을 상기 하나의 클라이언트 디바이스에 제공하는,

방법.

청구항 7

제 6 항에 있어서,

상기 서버는 상기 하나의 클라이언트 디바이스로부터 디바이스 설명을 수신하고, 상기 디바이스 설명에 기초하여 상기 서브밴드 공간 처리를 적용하기 위한 파라미터를 결정하는,

방법.

청구항 8

제 1 항에 있어서,

상기 하나의 클라이언트 디바이스는 상기 믹싱된 스트림 및 상기 공간적으로 향상된 스트림을 생성하는,

방법.

청구항 9

제 1 항에 있어서,

상기 공간적으로 향상된 스트림을 생성하는 단계는, 상기 믹싱된 스트림의 믹싱된 좌측 채널 및 믹싱된 우측 채널에 크로스토크 처리를 적용하는 단계를 더 포함하는, 방법.

청구항 10

비일시적 컴퓨터 판독가능 매체로서,
프로세서에 의해 실행될 때,

복수의 클라이언트 디바이스로부터 입력 오디오 스트림을 수신하고,

상기 복수의 클라이언트 디바이스 중 하나의 클라이언트 디바이스에 대해, 음장 내에서 상기 복수의 클라이언트 디바이스 중 다른 클라이언트 디바이스들에 대한 공간 위치를 정의하는 배치 데이터를 결정하고,

상기 배치 데이터에 따라 상기 다른 클라이언트 디바이스들의 입력 오디오 스트림을 믹싱 및 패닝함으로써, 상기 하나의 클라이언트 디바이스에 대해 믹싱된 좌측 채널 및 믹싱된 우측 채널을 포함하는 믹싱된 스트림을 생성하고,

상기 믹싱된 스트림의 믹싱된 좌측 채널 및 믹싱된 우측 채널에 서브밴드 공간 처리를 적용함으로써, 향상된 좌측 채널 및 향상된 우측 채널을 포함하는 공간적으로 향상된 스트림을 생성

하도록 상기 프로세서를 구성하는 프로그램 코드를 저장하되,

상기 서브밴드 공간 처리는,

상기 믹싱된 좌측 채널 및 상기 믹싱된 우측 채널의 중간 서브밴드 컴포넌트들에 제 1 이득을 적용하여 향상된 중간 컴포넌트를 생성하는 것과,

상기 믹싱된 좌측 채널 및 상기 믹싱된 우측 채널의 측면 서브밴드 컴포넌트들에 제 2 이득을 적용하여 향상된 측면 컴포넌트를 생성하는 것과,

상기 향상된 중간 컴포넌트 및 상기 향상된 측면 컴포넌트를 사용하여 상기 향상된 좌측 채널 및 상기 향상된 우측 채널을 생성하는 것을 포함하는,

비일시적 컴퓨터 판독가능 매체.

청구항 11

제 10 항에 있어서,

상기 공간적으로 향상된 스트림을 생성하도록 상기 프로세서를 구성하는 프로그램 코드는, 상기 믹싱된 스트림의 믹싱된 좌측 채널 및 믹싱된 우측 채널에 크로스토크 처리를 적용하도록 상기 프로세서를 또한 구성하는 프로그램 코드를 포함하는,

비일시적 컴퓨터 판독가능 매체.

청구항 12

제 10 항에 있어서,

실행될 때, 상기 하나의 클라이언트 디바이스로부터 디바이스 설명을 수신하고 상기 디바이스 설명에 기초하여 상기 서브밴드 공간 처리를 적용하기 위한 파라미터를 결정하도록 상기 프로세서를 또한 구성하는 프로그램 코드를 더 포함하는

비일시적 컴퓨터 판독가능 매체.

청구항 13

제 11 항에 있어서,

상기 크로스토크 처리는 크로스토크 제거 또는 크로스토크 시뮬레이션을 포함하고,

상기 공간적으로 향상된 스트림을 생성하도록 상기 프로세서를 구성하는 프로그램 코드는, 실행될 때, 상기 크로스토크 처리를 상기 믹싱된 스트림에 적용함으로써 야기되는 스펙트럼 결함을 조정하는 크로스토크 보상을 상기 믹싱된 스트림에 적용하도록 상기 프로세서를 구성하는 프로그램 코드를 포함하는,

비일시적 컴퓨터 판독가능 매체.

청구항 14

제 10 항에 있어서,

실행될 때, 상기 믹싱된 좌측 채널 및 상기 믹싱된 우측 채널을 포함하는 상기 믹싱된 스트림을 상기 하나의 클라이언트 디바이스에 제공하도록 상기 프로세서를 구성하는 프로그램 코드를 더 포함하는,

비일시적 컴퓨터 판독가능 매체.

청구항 15

제 10 항에 있어서,

실행될 때, 상기 배치 데이터를 결정하도록 상기 프로세서를 구성하는 프로그램 코드는, 네트워크를 통해 상기 하나의 클라이언트 디바이스로부터 상기 배치 데이터를 수신하도록 상기 프로세서를 구성하는 명령어를 포함하는,

비일시적 컴퓨터 판독가능 매체.

청구항 16

제 10 항에 있어서,

상기 복수의 클라이언트 디바이스로부터의 입력 오디오 스트림의 각각은 하나 이상의 오디오 채널을 포함하는,

비일시적 컴퓨터 판독가능 매체.

청구항 17

제 10 항에 있어서,

상기 배치 데이터에 따라 상기 다른 클라이언트 디바이스들의 입력 오디오 스트림을 믹싱 및 패닝함으로써, 상기 하나의 클라이언트 디바이스에 대해 믹싱된 좌측 채널 및 믹싱된 우측 채널을 포함하는 믹싱된 스트림을 생성하도록 상기 프로세서를 구성하는 프로그램 코드는,

상기 음장 내의 상기 다른 클라이언트 디바이스들의 공간 위치에 따라 상기 다른 클라이언트 디바이스들의 입력 오디오 스트림의 각각으로부터 좌측 채널 및 우측 채널을 생성하고,

상기 다른 클라이언트 디바이스들의 입력 오디오 스트림으로부터의 좌측 채널들을 결합하여 상기 믹싱된 좌측 채널을 생성하고,

상기 다른 클라이언트 디바이스들의 입력 오디오 스트림으로부터의 우측 채널들을 결합하여 상기 믹싱된 우측 채널을 생성

하도록 상기 프로세서를 또한 구성하는 프로그램 코드를 포함하는, 비일시적 컴퓨터 판독가능 매체.

청구항 18

복수의 클라이언트 디바이스 중 하나의 클라이언트 디바이스에 회의를 제공하는 시스템으로서, 처리 회로를 포함하되, 상기 처리 회로는,

상기 복수의 클라이언트 디바이스로부터 입력 오디오 스트림을 수신하고,

상기 복수의 클라이언트 디바이스 중 상기 하나의 클라이언트 디바이스에 대해, 음장 내에서 상기 복수의 클라이언트 디바이스 중 다른 클라이언트 디바이스들에 대한 공간 위치를 정의하는 배치 데이터를 결정하고,

상기 배치 데이터에 따라 상기 복수의 클라이언트 디바이스 중 상기 다른 클라이언트 디바이스들의 입력 오디오 스트림을 믹싱 및 패닝함으로써, 상기 복수의 클라이언트 디바이스 중 상기 하나의 클라이언트 디바이스에 대해 믹싱된 좌측 채널 및 믹싱된 우측 채널을 포함하는 믹싱된 스트림을 생성하고,

상기 믹싱된 스트림의 믹싱된 좌측 채널 및 믹싱된 우측 채널에 서브밴드 공간 처리를 적용함으로써, 향상된 좌측 채널 및 향상된 우측 채널을 포함하는 공간적으로 향상된 스트림을 생성하도록 구성되고,

상기 서브밴드 공간 처리는,

상기 믹싱된 좌측 채널 및 상기 믹싱된 우측 채널의 중간 서브밴드 컴포넌트들에 제 1 이득을 적용하여 향상된 중간 컴포넌트를 생성하는 것과,

상기 믹싱된 좌측 채널 및 상기 믹싱된 우측 채널의 측면 서브밴드 컴포넌트들에 제 2 이득을 적용하여 향상된 측면 컴포넌트를 생성하는 것과,

상기 향상된 중간 컴포넌트 및 상기 향상된 측면 컴포넌트를 사용하여 상기 향상된 좌측 채널 및 상기 향상된 우측 채널을 생성하는 것을 포함하는,

시스템.

청구항 19

제 18 항에 있어서,

상기 공간적으로 향상된 스트림을 생성하도록 구성된 상기 처리 회로는, 상기 믹싱된 스트림의 믹싱된 좌측 채널 및 믹싱된 우측 채널에 크로스토크 처리를 적용하도록 또한 구성되는 상기 처리 회로를 포함하는,

시스템.

청구항 20

제 18 항에 있어서,

상기 복수의 클라이언트 디바이스로부터의 입력 오디오 스트림의 각각은 하나 이상의 오디오 채널을 포함하는,

시스템.

청구항 21

제 18 항에 있어서,

상기 처리 회로는 또한, 상기 믹싱된 좌측 채널 및 상기 믹싱된 우측 채널을 포함하는 상기 믹싱된 스트림을 상기 복수의 클라이언트 디바이스 중 상기 하나의 클라이언트 디바이스에 제공하도록 구성되는, 시스템.

청구항 22

제 9 항에 있어서,

상기 크로스토크 처리는 크로스토크 제거 또는 크로스토크 시뮬레이션을 포함하고,

상기 공간적으로 향상된 스트림을 생성하는 단계는, 상기 크로스토크 처리를 상기 믹싱된 스트림에 적용함으로써 야기되는 스펙트럼 결함을 조정하는 크로스토크 보상을 상기 믹싱된 스트림에 적용하는 단계를 포함하는, 방법.

발명의 설명

기술 분야

[0001] 본 명세서에 설명된 주제는 오디오 처리에 관한 것이고, 보다 구체적으로는 회의 클라이언트 디바이스를 위한 공간화된(spatialized) 오디오 처리에 관한 것이다.

배경 기술

[0002] 전자 디바이스는 여러 사용자 간에 원격 회의를 제공하는 데 사용된다. 일반적으로, 사용자에게 의한 스피치를 캡처하기 위해 사용자의 오디오 스트림이 생성되고, 사용자에게 의한 청취를 위한 사운드를 제공하기 위해 다른 사용자들의 오디오 스트림이 결합된다. 예를 들어, 결합된 스트림은 스피커에 대한 모노포닉 스트림일 수 있다. 스테레오 스피커의 경우, 좌측 및 우측 스피커에 대해 모노포닉 스트림이 재생된다. 대면 미팅과 달리, 모노포닉 스트림에서는 상이한 참가자들에 대한 음장(sound field)의 공간 분별(spatial sense)이 존재하지 않으므로, 음성 차별화 및 명료성(intelligibility)이 감소된다.

발명의 내용

[0003] 실시예는 공간화된 오디오를 이용하여 클라이언트 디바이스에 회의를 제공하는 것에 관한 것이다. 일부 실시예에서, 입력 오디오 스트림은 클라이언트 디바이스로부터 수신된다. 클라이언트 디바이스에 대해, 음장 내에서 다른 클라이언트 디바이스들의 공간 위치를 정의하는 배치 데이터가 결정된다. 배치 데이터에 따라 다른 클라이언트 디바이스들의 입력 오디오 스트림을 믹싱(mixing) 및 패닝(panning)함으로써, 클라이언트 디바이스에 대해 믹싱된 좌측 채널 및 믹싱된 우측 채널을 포함하는 믹싱된 스트림이 생성된다. 믹싱된 스트림의 믹싱된 좌측 채널 및 믹싱된 우측 채널에 서브밴드 공간 처리 및 크로스토크 처리를 적용함으로써, 향상된 좌측 채널 및 향상된 우측 채널을 포함하는 공간적으로 향상된 스트림이 생성된다. 공간적으로 향상된 스트림은 클라이언트 디바이스에 제공된다.

[0004] 일부 실시예에서, 비일시적 컴퓨터 관독가능 매체는 명령어를 저장하는데, 명령어는, 프로세서에 의해 실행될 때, 클라이언트 디바이스로부터 입력 오디오 스트림을 수신하고, 클라이언트 디바이스에 대해, 음장 내에서 다른 클라이언트 디바이스들에 대한 공간 위치를 정의하는 배치 데이터를 결정하고, 배치 데이터에 따라 다른 클라이언트 디바이스들의 입력 오디오 스트림을 믹싱 및 패닝함으로써, 클라이언트 디바이스에 대해 믹싱된 좌측 채널 및 믹싱된 우측 채널을 포함하는 믹싱된 스트림을 생성하도록 프로세서를 구성한다.

[0005] 일부 실시예에서, 시스템은 클라이언트 디바이스에 회의를 제공한다. 이 시스템은, 클라이언트 디바이스로부터 입력 오디오 스트림을 수신하고, 클라이언트 디바이스에 대해, 음장 내에서 다른 클라이언트 디바이스들에 대한 공간 위치를 정의하는 배치 데이터를 결정하고, 배치 데이터에 따라 다른 클라이언트 디바이스들의 입력 오디오 스트림을 믹싱 및 패닝함으로써, 클라이언트 디바이스에 대해 믹싱된 좌측 채널 및 믹싱된 우측 채널을 포함하는 믹싱된 스트림을 생성하도록 구성된 처리 회로를 포함한다.

[0006] 다른 양상들은 컴포넌트, 디바이스, 시스템, 개선, 방법, 프로세스, 애플리케이션, 컴퓨터 관독가능 매체, 및 전술한 것들 중 임의의 것과 관련된 다른 기술을 포함한다.

도면의 간단한 설명

- [0007] 도 1은 일부 실시예에 따른 오디오 처리 시스템의 개략적 블록도이다.
 - 도 2는 일부 실시예에 따른 공간 믹싱 프로세서의 개략적 블록도이다.
 - 도 3은 일부 실시예에 따른 크로스토크 제거를 갖는 공간 향상 프로세서의 개략적 블록도이다.
 - 도 4는 일부 실시예에 따른 서브밴드 공간 프로세서의 개략적 블록도이다.
 - 도 5는 일부 실시예에 따른 크로스토크 보상 프로세서의 개략적 블록도이다.
 - 도 6은 일부 실시예에 따른 크로스토크 제거 프로세서의 개략적 블록도이다.
 - 도 7은 일부 실시예에 따른 크로스토크 시뮬레이션을 갖는 공간 향상 프로세서의 개략적 블록도이다.
 - 도 8은 일부 실시예에 따른 크로스토크 시뮬레이션 프로세서의 개략적 블록도이다.
 - 도 9는 일부 실시예에 따른, 클라이언트 디바이스들 간의 회의를 위한 프로세스의 흐름도이다.
 - 도 10은 일부 실시예에 따른 오디오 처리 시스템의 도면이다.
 - 도 11은 일부 실시예에 따른 오디오 처리 시스템의 도면이다.
 - 도 12는 일부 실시예에 따른 오디오 처리 시스템의 도면이다.
 - 도 13은 일부 실시예에 따른 컴퓨터 시스템의 개략적 블록도이다.
- 도면들은 단지 예시의 목적으로 다양한 비제한적 실시예를 도시하고, 상세한 설명은 그에 대해 설명한다.

발명을 실시하기 위한 구체적인 내용

- [0008] 이하, 실시예들에 대한 참조가 구체적으로 이루어질 것인데, 이들의 예는 첨부된 도면에 도시된다. 다음의 상세한 설명에서는, 설명된 다양한 실시예의 철저한 이해를 제공하기 위해 다수의 특정 세부사항이 제시된다. 그러나, 설명된 실시예는 이러한 특정 세부사항 없이 실시될 수 있다. 다른 경우들에서, 공지된 방법, 절차, 컴포넌트, 회로 및 네트워크는 실시예의 양상을 불필요하게 모호하게 하지 않도록 상세히 설명되지 않았다.
- [0009] 실시예는 각각의 클라이언트 디바이스에 대한 공간화된 오디오 출력을 이용하여 클라이언트 디바이스에 회의를 제공하는 것에 관한 것이다. 예를 들어, 오디오 처리 시스템은, 실시간 공간 믹싱 및 공간 향상의 조합을 사용하여 각 클라이언트 디바이스의 렌더링 디바이스 유형에 순응적으로 최적화된 공간화된 오디오(spatialized audio)를 각 클라이언트 디바이스에 전달한다. 클라이언트 디바이스에 대한 출력 스트림을 생성하기 위해, 다른 클라이언트 디바이스들로부터의 오디오 스트림은 클라이언트 디바이스의 음장 내의 공간 위치와 각각 연관된다. 오디오 스트림은 공간 위치에 따라 믹싱 및 패닝되어 공간화된 오디오를 갖는 믹싱된 스트림을 생성한다. 다음에, 음장의 공간 분별을 향상시키기 위해 공간화된 오디오 믹싱된 스트림에 공간 향상이 적용된다. 공간 향상은 서브밴드 공간 처리 및 크로스토크 처리를 포함할 수 있다. 크로스토크 처리는 크로스토크 제거(예컨대, 라우드스피커의 경우) 또는 크로스토크 시뮬레이션(예컨대, 헤드폰의 경우)을 포함할 수 있다. 무엇보다도, 서로 다른 원격 참가자 음성에 대한 음장 내의 공간 분별은 원격 회의 음성의 차별화 및 명료성을 향상시킨다.
- [0010] 일부 실시예에서, 회의 파이프라인은 서버(또는 "브리지") 및 둘 이상의 클라이언트 디바이스를 포함한다. 클라이언트 디바이스는 다양한 사운드 렌더링 하드웨어를 포함할 수 있다. 서버는 상이한 유형의 하드웨어에 대해 디바이스-최적화된 향상된 공간 오디오를 제공한다.
- [0011] 일부 실시예에서, 들어오는 오디오 스트림의 하나 이상의 채널은 스테레오 오디오 스트림으로 믹스다운(mixdown)된다. 그런 다음, 각각의 클라이언트 디바이스의 렌더링 시스템(예컨대, 내장 랩탑 스피커, 블루투스 스피커, 헤드폰 등)에 기초하여, 디바이스-특정 공간 오디오 향상이 순응적으로 적용된다. 일부 실시예에서, 각각의 클라이언트 디바이스는 서버로부터 믹싱된 스트림을 수신하고 적절한 공간 향상을 적용한다. 일부 실시예에서, 클라이언트 디바이스는 공간 향상을 위한 파라미터를 서버에 제공할 수 있고, 서버는 공간 향상을 수행

하여 클라이언트 디바이스에 대한 공간적으로 향상된 스트림을 생성한다. 일부 실시예에서, 클라이언트 디바이스는 하나 이상의 클라이언트 오디오 스트림을 수신할 수 있고, 믹스다운 및 그 다음에 이어지는 믹싱된 스테레오 스트림에 대한 적절한 공간적 향상을 모두 적용할 수 있다.

[0012] 예시적 오디오 처리 시스템

[0013] 도 1은 일부 실시예에 따른 오디오 처리 시스템(100)의 블록도이다. 오디오 처리 시스템(100)은 서버(102) 및 클라이언트 디바이스(130)를 포함한다. 단일 클라이언트 디바이스(130)가 도시되어 있지만, 서버(102)는 다수의 클라이언트 디바이스(130) 사이에서 회의를 제공한다. 서버(102)는 인터넷을 포함하는 네트워크를 통해 클라이언트 디바이스에 접속될 수 있다. 각각의 클라이언트 디바이스(130)는 오디오 캡처 디바이스(156)(예컨대, 마이크로폰), 좌측 스피커(152) 및 우측 스피커(154)를 포함할 수 있다. 스피커(152, 154)는 라우드스피커 또는 헤드폰일 수 있다.

[0014] 서버(102)는 공간 믹싱 프로세서(110) 및 공간 향상 프로세서(120)를 포함한다. 다수의 클라이언트 디바이스(130) 사이에서 회의를 가능하게 하기 위해, 공간 믹싱 프로세서(110)는 각각 클라이언트 디바이스(130)로부터 다수의 입력 오디오 스트림(140)을 수신하고, 입력 오디오 스트림(140)의 스테레오 오디오 스트림으로의 패닝 및 믹스다운을 적용한다. 스테레오 오디오 스트림은 믹싱된 좌측 채널(142) 및 믹싱된 우측 채널(144)을 포함한다. 입력 오디오 스트림(140)의 각각은 클라이언트 디바이스(130)의 오디오 캡처 디바이스(156)에 의해 생성된 하나 이상의 채널을 갖는 오디오 스트림일 수 있다. 일부 실시예에서, 오디오 스트림(140)은 모노포닉 오디오 스트림을 포함할 수 있다.

[0015] 회의의 각 참가자는 클라이언트 디바이스(130)와 연관된다. 각각의 클라이언트 디바이스(130)는 공간 믹싱 프로세서(110)로 전송되는 오디오 스트림(140)을 생성할 수 있다. 오디오 스트림(140)과 관련하여, 각각의 클라이언트 디바이스(130)는 디바이스 식별 데이터 및 배치 데이터를 포함하는 다양한 유형의 메타데이터를 공간 믹싱 프로세서(110)에 제공할 수 있다. 디바이스 식별 데이터는 클라이언트 디바이스(130)를 식별하는 고유 식별자이다. 배치 데이터는 클라이언트 디바이스(130)에 대해 다른 클라이언트 디바이스(130)의 다른 오디오 스트림(140)의 공간 위치를 정의한다. 공간 믹싱 프로세서(110)는 각각의 클라이언트 디바이스(130)에 대한 고유 스테레오 오디오 스트림을 생성하는데, 다른 클라이언트 디바이스(130)로부터의 오디오 신호(140)는 배치 데이터에 따라 스테레오 음장에 걸쳐 분포(예컨대, 패닝)된다.

[0016] 일부 실시예에서, 오디오 스트림(140)에 대한 패닝은 오디오 스트림(140)의 공간 분포를 정의하는 미리 정의된 배치 데이터에 기초할 수 있다. 배치 기술은 원격 클라이언트 디바이스(130)의 동일한 공간 분포, 또는 참가자 메타데이터(예컨대, 사용자 역할, 엔티티/회사 이름, 그룹 멤버십, 룸 위치, 지리적 위치, IP 주소 등)에 기초한 소스의 그룹화 및 패닝을 포함할 수 있다. 일부 실시예에서, 각각의 클라이언트 디바이스(130)에 대한 패닝은 클라이언트 디바이스(130)로부터의 제어에 기초할 수 있다. 예를 들어, 클라이언트 디바이스(130)의 사용자는 각각의 원격 참가자에 대한 원하는 패닝 위치를 지정하는 배치 데이터를 정의할 수 있다.

[0017] 공간 믹싱 프로세서(110)는 각각의 클라이언트 디바이스(130)에 대해 믹싱된 좌측 채널(142) 및 믹싱된 우측 채널(144)을 포함하는 고유 스테레오 오디오 스트림을 생성한다. 오디오 신호(140)의 공간 분포를 정의하는 배치 데이터가 프로그램적으로 정의되거나 사용자에게 의해 정의되면, 다른 클라이언트 디바이스(130)의 오디오 신호(140)는 스테레오 믹스다운으로 패닝된다. 패닝은 클라이언트 디바이스(130)에 대한 스테레오 스트림을 생성하기 위해 진폭 패닝, 지연 패닝, 바이노럴 패닝(binaural panning) 등과 같은 패닝 기술을 포함할 수 있다. 일부 실시예에서, 패닝은 최종 믹스에서 각 요소의 존재를 조정하기 위해 이득을 적용하는 것을 포함할 수 있다. 거리 큐(distance cue)를 위한 공기 흡수 시뮬레이션 또는 실내 음향 시뮬레이션과 같은 다른 인지적 동기 부여 큐가 적용될 수도 있다.

[0018] 일부 실시예에서, 클라이언트 디바이스(130)에 대한 공간 믹싱 프로세서(110)의 출력은, 예컨대 클라이언트 디바이스(130) 또는 청취 컨텍스트가 1-채널 재생을 사용하는 경우에, 단일 채널에 대한 모노포닉 감소를 포함할 수 있다.

[0019] 공간 향상 프로세서(120)는 믹싱된 좌측 채널(142) 및 믹싱된 우측 채널(144)에 공간 처리를 적용하여 향상된 좌측 채널(146) 및 향상된 우측 채널(148)을 생성할 수 있다. 공간 향상 처리는 서브밴드 공간 처리 및 크로스토크 처리를 포함할 수 있다. 크로스토크 처리는 크로스토크 제거 또는 크로스토크 시뮬레이션을 포함할 수 있다. 크로스토크 제거는 클라이언트 디바이스(130)의 스피커(152, 154)가 라우드스피커인 경우에 라우드스피커에 의해 야기된 크로스토크를 제거하는 데 사용될 수 있다. 크로스토크 시뮬레이션은 스피커(152, 154)가 헤드

폰인 경우에 라우드스피커의 효과를 시뮬레이션하는 데 사용될 수 있다. 믹싱된 좌측 채널(142) 및 믹싱된 우측 채널(144)에 크로스토크 처리가 적용되는 경우, 공간 향상 처리는 크로스토크 처리의 적용으로 인해 야기된 스펙트럼 결함을 보정하는 크로스토크 보상을 더 포함할 수 있다. 공간 향상 프로세서(120)는 향상된 좌측 채널(146)을 클라이언트 디바이스(130)의 좌측 스피커(152)에 제공하고, 향상된 우측 채널(148)을 클라이언트 디바이스(130)의 우측 스피커(154)에 제공한다. 스피커(152, 154)는 각각의 출력 채널(O_L 및 O_R)을 사운드로 변환한다.

[0020] 일부 실시예에서, 공간 향상 프로세서(120)는 서버(102)에 위치한다. 서버(102)는 각각의 클라이언트 디바이스(130)에 대해 개별적인 공간 향상 처리 인스턴스를 실행할 수 있다. 다른 실시예에서, 공간 향상 프로세서(120)는 각 클라이언트 디바이스(130)에 위치한다. 각 클라이언트 디바이스(130)에 대해, 서버(102)는 믹싱된 좌측 채널(142) 및 믹싱된 우측 채널(144)을 포함하는 믹싱된 스트림을 클라이언트 디바이스(130)의 공간 향상 프로세서(120)에 제공한다. 각 클라이언트 디바이스(130)의 공간 향상 프로세서(120)는 서버(102)로부터 수신된 믹싱된 채널을 처리하여 좌측 출력 채널(146) 및 우측 출력 채널(148)을 생성한다. 일부 실시예에서, 예컨대, 시스템(100)의 피어 투 피어 회의 구성에서 또는 서버가 믹스다운을 수행하지 않을 때, 공간 믹싱 프로세서(110)도 클라이언트 디바이스(130)에 위치한다.

[0021] 시스템(100)은 더 적거나 추가의 컴포넌트를 포함할 수 있다. 예를 들어, 오디오 스트림을 사용하는 회의는 비디오 스트림과 통합될 수 있다. 일부 실시예에서, 시스템(100)은 인공 현실 시스템이며, 여기서 각 클라이언트 디바이스(130)는 헤드-장착형 디스플레이를 포함할 수 있다. 헤드-장착형 디스플레이는 비디오에서 다른 사용자 또는 사용자들의 아바타를 렌더링할 수 있는데, 사용자는 음장 내의 믹싱된 스트림에서 클라이언트 디바이스(130)의 공간 위치에 대응하도록 위치한다. 따라서 인공 현실 환경의 몰입 품질이 개선된다.

[0022] 예시적 공간 믹싱 프로세서

[0023] 도 2는 일부 실시예에 따른 공간 믹싱 프로세서(110)의 개략적 블록도이다. 공간 믹싱 프로세서(110)는 공간 믹서(205), 바이노럴(binaural) 필터(210 및 215), 좌측 채널 결합기(220), 우측 채널 결합기(225), 및 인지 시뮬레이터(230)를 포함한다. 공간 믹서(205)는, 들어오는 오디오 스트림 내의 에너지가 최종 스테레오 믹스다운에 어떻게 분포되는지 및 이와 관련하여 최종 믹스 내의 요소가 사운드스테이지의 컨텍스트 내에서 어떻게 인지되는지에 대해 제어를 적용한다. 예를 들어, 모노 오디오 스트림의 공간적 믹싱은 일정한 전력 또는 선형 스테레오 패닝 기술을 사용하여 달성될 수 있다. 일부 실시예에서, 공간 믹서(205)는 각각의 오디오 신호(140)에 공간 믹싱을 적용하여 좌측 및 우측 채널을 생성하고, 좌측 채널들을 결합하여 좌측 채널(250)을 생성하고, 우측 채널들을 결합하여 우측 채널(252)을 생성한다.

[0024] 오디오 스트림은 또한, 바이노럴 필터링을 통해 달성될 수 있는 것처럼, 청취자의 머리 주위의 3D 공간에서 사운드를 인지 가능하게 위치시키는 방식으로 처리될 수 있다. 바이노럴 필터(210, 215) 각각은, 청취자가 입력 채널의 사운드를 인지해야 하는 타겟 소스 위치를 설명하는 HRTF(head-related transfer function)를 적용한다. 바이노럴 필터(210)는 좌측 채널(250)을 수신하고, 좌측 채널(250)과 연관된 각도 위치를 조정하는 HRTF를 적용함으로써 좌측 및 우측 출력 채널을 생성한다. 바이노럴 필터(215)는 우측 채널(252)을 수신하고, 우측 채널(252)과 연관된 각도 위치를 조정하는 HRTF를 적용함으로써 좌측 및 우측 출력을 생성한다. 좌측 채널 결합기(220)는 바이노럴 필터(210, 215)로부터 좌측 채널들을 수신하고, 이들 채널을 결합하여 좌측 채널(254)을 생성한다. 우측 채널 결합기(225)는 바이노럴 필터(210 및 215)로부터 우측 채널들을 수신하고, 이들 채널을 결합하여 우측 채널(256)을 생성한다.

[0025] 처리는 또한 다른 실세계 인지 신호를 시뮬레이션하기 위해 적용될 수 있다. 인지 시뮬레이터(230)는 좌측 채널(254) 및 우측 채널(256)을 수신하고, 하나 이상의 채널에 인지 큐를 적용하여 믹싱된 좌측 채널(142) 및 믹싱된 우측 채널(144)을 생성한다. 채널은, 예를 들어 자유장 역거리 법칙(free-field inverse distance law)을 사용하여, 믹스 내에서 가변 거리 및 공기 흡수에 대한 인지를 제공하도록 스케일링 및 필터링될 수 있는데, 음압 레벨은 기준 거리(예컨대, 청취자의 가상 위치로부터 정의된 거리)에 대해 거리의 2배 마다 6 데시벨씩 감소된다. 고주파 에너지가 기준 거리에 대해 거리의 증가에 반비례하여 감소되는 공기 흡수가 시뮬레이션될 수 있다. 거리에 대한 인지를 더욱 향상시키고 실내 음향 및 환경적 컨텍스트를 생성하기 위해 리버브(reverb)가 사용될 수 있다.

[0026] 일부 실시예에서, 오디오 신호(140)는 다수의(예컨대, 좌측 및 우측) 채널을 포함한다. 공간 믹서(205)는 오디오 신호의 채널들을 모노포닉 채널로 결합함으로써 믹스다운을 적용하고, 모노포닉 채널에 일정한 전력 또는 선형 스테레오 패닝을 적용한다. 다른 예에서는, 모노포닉 신호로의 결합이 없고, 믹스다운은 오디오 신호(140)

의 각 채널을 사용하여 적용된다. 예를 들어, 공간 믹서(205)는 청취자 주위의 공간에서의 이상적 위치에 기초하여 각 채널에 상이한 바이노럴 필터링을 적용하고, 결과를 스테레오 좌측 및 우측 채널로 믹스다운할 수 있다.

[0027] 공간 믹싱 프로세서(110)는 더 적거나 추가적인 컴포넌트를 포함할 수 있다. 예를 들어, 인지 시뮬레이터(230) 또는 바이노럴 필터(210, 215)는 생략될 수 있다. 공간 믹싱 프로세서(110)는 진폭 패닝, 지연 패닝, 바이노럴 패닝 등을 포함하는 오디오 신호(140)의 다양한 유형의 믹스다운을 수행할 수 있다.

[0028] 예시적 공간 향상 프로세서

[0029] 도 3은 일 실시예에 따른 공간 향상 프로세서(300)의 개략적 블록도이다. 공간 향상 프로세서(300)는, 스피커들(152, 154)이 라우드스피커이고 크로스토크 처리가 클라이언트 디바이스(130)에 대한 크로스토크 제거인 공간 향상 프로세서(120)의 예이다. 공간 향상 프로세서(300)는 좌측 입력 채널(X_L) 및 우측 입력 채널(X_R)을 포함하는 입력 오디오 신호(X)를 수신한다. 좌측 입력 채널(X_L)은 공간 믹싱 프로세서(110)로부터의 좌측 출력 채널(142)일 수 있고, 우측 입력 채널(X_R)은 공간 믹싱 프로세서(110)로부터의 우측 출력 채널(144)일 수 있다.

[0030] 공간 향상 프로세서(300)는 입력 채널들(X_L , X_R)을 처리하여 향상된 좌측 채널(O_L) 및 향상된 우측 채널(O_R)을 포함하는 출력 오디오 신호(O)를 생성한다. 향상된 좌측 채널(O_L)은 향상된 좌측 채널(146)에 대응하고 향상된 우측 채널(O_R)은 향상된 우측 채널(148)에 대응한다. 오디오 출력 신호(O)는 크로스토크 보상 및 크로스토크 제거를 갖는 입력 오디오 신호(X)의 공간적으로 향상된 오디오 신호이다.

[0031] 공간 향상 프로세서(300)는 서브밴드 공간 프로세서(310), 크로스토크 보상 프로세서(320), 결합기(360), 및 크로스토크 제거 프로세서(370)를 포함한다. 공간 향상 프로세서(300)는 입력 오디오 입력 채널(X_L , X_R)의 크로스토크 보상 및 서브밴드 공간 처리를 수행하고, 서브밴드 공간 처리의 결과와 크로스토크 보상의 결과를 결합한 다음, 결합된 신호에 대해 크로스토크 제거를 수행한다.

[0032] 서브밴드 공간 프로세서(310)는 공간 주파수 대역 분할기(340), 공간 주파수 대역 프로세서(345) 및 공간 주파수 대역 결합기(350)를 포함한다. 공간 주파수 대역 분할기(340)는 입력 채널(X_L 및 X_R) 및 공간 주파수 대역 프로세서(345)에 결합된다. 공간 주파수 대역 분할기(340)는 좌측 입력 채널(X_L) 및 우측 입력 채널(X_R)을 수신하고, 입력 채널들을 공간(또는 "측면") 컴포넌트(Y_s) 및 비공간(또는 "중간") 컴포넌트(Y_m)로 처리한다. 예를 들어, 공간 컴포넌트(Y_s)는 좌측 입력 채널(X_L)과 우측 입력 채널(X_R) 간의 차에 기초하여 생성될 수 있다. 비공간 컴포넌트(Y_m)은 좌측 입력 채널(X_L)과 우측 입력 채널(X_R)의 합에 기초하여 생성될 수 있다. 공간 주파수 대역 분할기(340)는 공간 컴포넌트(Y_s) 및 비공간 컴포넌트(Y_m)를 공간 주파수 대역 프로세서(345)에 제공한다.

[0033] 공간 주파수 대역 프로세서(345)는 공간 주파수 대역 분할기(340) 및 공간 주파수 대역 결합기(350)에 결합된다. 공간 주파수 대역 프로세서(345)는 공간 주파수 대역 분할기(340)로부터 공간 컴포넌트(Y_s) 및 비공간 컴포넌트(Y_m)를 수신하고, 수신된 신호를 향상시킨다. 구체적으로, 공간 주파수 대역 프로세서(345)는 공간 컴포넌트(Y_s)로부터 향상된 공간 컴포넌트(E_s)를 생성하고, 비공간 컴포넌트(Y_m)로부터 향상된 비공간 컴포넌트(E_m)를 생성한다.

[0034] 예를 들어, 공간 주파수 대역 프로세서(345)는 공간 컴포넌트(Y_s)에 서브밴드 이득을 적용하여 향상된 공간 컴포넌트(E_s)를 생성하고, 비공간 컴포넌트(Y_m)에 서브밴드 이득을 적용하여 향상된 비공간 컴포넌트(E_m)를 생성한다. 일부 실시예에서, 공간 주파수 대역 프로세서(345)는 부가적으로 또는 대안적으로 공간 컴포넌트(Y_s)에 서브밴드 지연을 제공하여 향상된 공간 컴포넌트(E_s)를 생성하고, 비공간 컴포넌트(Y_m)에 서브밴드 지연을 제공하여 향상된 비공간 컴포넌트(E_m)를 생성한다. 서브밴드 이득 및/또는 지연은 공간 컴포넌트(Y_s) 및 비공간 컴포넌트(Y_m)의 상이한(예컨대, n 개) 서브밴드에 대해 상이할 수도 있고, (예컨대, 2개 이상의 서브밴드에 대해) 동일할 수도 있다. 공간 주파수 대역 프로세서(345)는 공간 컴포넌트(Y_s) 및 비공간 컴포넌트(Y_m)의 서로 다른 서브밴드에 대한 이득 및/또는 지연을 서로에 대해 조정하여 향상된 공간 컴포넌트(E_s) 및 향상된 비공간 컴포넌트

(E_m)를 생성한다. 그런 다음, 공간 주파수 대역 프로세서(345)는 향상된 공간 컴포넌트(E_s) 및 향상된 비공간 컴포넌트(E_m)를 공간 주파수 대역 결합기(350)에 제공한다.

[0035] 공간 주파수 대역 결합기(350)는 공간 주파수 대역 프로세서(345)에 결합되고, 또한 결합기(360)에도 결합된다. 공간 주파수 대역 결합기(350)는 공간 주파수 대역 프로세서(345)로부터 향상된 공간 컴포넌트(E_s) 및 향상된 비공간 컴포넌트(E_m)를 수신하고, 향상된 공간 컴포넌트(E_s)와 향상된 비공간 컴포넌트(E_m)를 공간적으로 향상된 좌측 채널(E_L) 및 공간적으로 향상된 우측 채널(E_R)로 결합한다. 예를 들어, 향상된 공간 컴포넌트(E_s)와 향상된 비공간 컴포넌트(E_m)의 합에 기초하여 공간적으로 향상된 좌측 채널(E_L)이 생성될 수 있고, 향상된 비공간 컴포넌트(E_m)와 향상된 공간 컴포넌트(E_s) 간의 차에 기초하여 공간적으로 향상된 우측 채널(E_R)이 생성될 수 있다. 공간 주파수 대역 결합기(350)는 공간적으로 향상된 좌측 채널(E_L) 및 공간적으로 향상된 우측 채널(E_R)을 결합기(360)에 제공한다. 서브밴드 공간 프로세서(310)에 관한 추가 세부사항은 도 4와 관련하여 아래에서 설명된다.

[0036] 크로스토크 보상 프로세서(320)는 크로스토크 보상을 수행하여 크로스토크 제거 시의 스펙트럼 결함 또는 아티팩트를 보상한다. 크로스토크 보상 프로세서(320)는 입력 채널(X_L 및 X_R)을 수신하고, 크로스토크 제거 프로세서(370)에 의해 수행되는 향상된 비공간 컴포넌트(E_m) 및/또는 향상된 공간 컴포넌트(E_s)의 후속 크로스토크 제거 시의 임의의 아티팩트를 보상하기 위한 처리를 수행한다. 일부 실시예에서, 크로스토크 보상 프로세서(320)는 필터를 적용하여 좌측 크로스토크 보상 채널(Z_L) 및 우측 크로스토크 보상 채널(Z_R)을 포함하는 크로스토크 보상 신호(Z)를 생성함으로써 비공간 컴포넌트(X_m) 및 공간 컴포넌트(X_s)에 대한 향상을 수행할 수 있다. 다른 실시예에서, 크로스토크 보상 프로세서(320)는 비공간 컴포넌트(X_m)에 대해서만 향상을 수행할 수 있다. 크로스토크 보상 프로세서(320)에 관한 추가 세부사항은 도 4와 관련하여 아래에서 설명된다.

[0037] 결합기(360)는 공간적으로 향상된 좌측 채널(E_L)과 좌측 크로스토크 보상 채널(Z_L)을 결합하여 향상된 좌측 보상 채널(T_L)을 생성하고, 공간적으로 향상된 우측 채널(E_R)과 우측 크로스토크 보상 채널(Z_R)을 결합하여 우측 보상 채널(T_R)을 생성한다. 결합기(360)는 크로스토크 제거 프로세서(370)에 결합되고, 향상된 좌측 보상 채널(T_L) 및 향상된 우측 보상 채널(T_R)을 크로스토크 제거 프로세서(370)에 제공한다.

[0038] 크로스토크 제거 프로세서(370)는 향상된 좌측 보상 채널(T_L) 및 향상된 우측 보상 채널(T_R)을 수신하고, 채널들(T_L , T_R)에 대해 크로스토크 제거를 수행하여 좌측 출력 채널(O_L) 및 우측 출력 채널(O_R)을 포함하는 출력 오디오 신호(O)를 생성한다. 크로스토크 제거 프로세서(370)에 관한 추가 세부사항은 도 5와 관련하여 아래에서 설명된다.

[0039] 도 4는 일부 실시예에 따른 서브밴드 공간 프로세서(310)의 개략적 블록도이다. 공간 주파수 대역 분할기(340)는, 좌측 입력 채널(X_L) 및 우측 입력 채널(X_R)을 수신하고 이들 입력을 공간 컴포넌트(Y_m) 및 비공간 컴포넌트(Y_s)로 변환하는 L/R-M/S 변환기(402)를 포함한다.

[0040] 공간 주파수 대역 프로세서(345)는 비공간 컴포넌트(Y_m)를 수신하고, 한 세트의 서브밴드 필터를 적용하여 향상된 비공간 서브밴드 컴포넌트(E_m)를 생성한다. 공간 주파수 대역 프로세서(345)는 또한 공간 서브밴드 컴포넌트(Y_s)를 수신하고, 한 세트의 서브밴드 필터를 적용하여 향상된 비공간 서브밴드 컴포넌트(E_m)를 생성한다. 서브밴드 필터는 피크 필터, 노치 필터, 로우 패스 필터, 하이 패스 필터, 로우 션프 필터, 하이 션프 필터, 밴드 패스 필터, 밴드 스톱 필터 및/또는 올 패스 필터의 다양한 조합을 포함할 수 있다.

[0041] 보다 구체적으로, 공간 주파수 대역 프로세서(345)는, 비공간 컴포넌트(Y_m)의 n 개의 주파수 서브밴드 각각에 대한 서브밴드 필터 및 공간 컴포넌트(Y_s)의 n 개의 서브밴드 각각에 대한 서브밴드 필터를 포함한다. 예를 들어, $n = 4$ 개의 서브밴드에 대해, 공간 주파수 대역 프로세서(345)는, 서브밴드(1)에 대한 중간 등화(EQ) 필터(404(1)), 서브밴드(2)에 대한 중간 EQ 필터(404(2)), 서브밴드(3)에 대한 중간 EQ 필터(404(3)) 및 서브밴드(4)에 대한 중간 EQ 필터(404(4))를 포함하는, 비공간 컴포넌트(Y_m)에 대한 일련의 서브밴드 필터를 포함한다.

각각의 중간 EQ 필터(404)는 비공간 컴포넌트(Y_m)의 주파수 서브밴드 부분에 필터를 적용하여 향상된 비공간 컴포넌트(E_m)를 생성한다.

[0042] 공간 주파수 대역 프로세서(345)는, 서브밴드(1)에 대한 측면 등화(EQ) 필터(406(1)), 서브밴드(2)에 대한 측면 EQ 필터(406(2)), 서브밴드(3)에 대한 측면 EQ 필터(406(3)), 및 서브밴드(4)에 대한 측면 EQ 필터(406(4))를 포함하는, 공간 컴포넌트(Y_s)의 주파수 서브밴드에 대한 일련의 서브밴드 필터를 더 포함한다. 각각의 측면 EQ 필터(406)는 공간 컴포넌트(Y_s)의 주파수 서브밴드 부분에 필터를 적용하여 향상된 공간 컴포넌트(E_s)를 생성한다.

[0043] 비공간 컴포넌트(Y_m) 및 공간 컴포넌트(Y_s)의 n 개의 주파수 서브밴드 각각은 주파수 범위에 대응할 수 있다. 예를 들어, 주파수 서브밴드(1)는 0 내지 300Hz에 대응할 수 있고, 주파수 서브밴드(2)는 300 내지 510Hz에 대응할 수 있고, 주파수 서브밴드(3)는 510 내지 2700Hz에 대응할 수 있고, 주파수 서브밴드(4)는 2700Hz 내지 나이퀴스트(Nyquist) 주파수에 대응할 수 있다. 일부 실시예에서, n 개의 주파수 서브밴드는 통합된 임계 대역 세트이다. 임계 대역은 다양한 음악 장르로부터의 오디오 샘플 모음을 사용하여 결정될 수 있다. 24 바크 스케일(Bark scale) 임계 대역에 대한 중간-측면 컴포넌트의 장기 평균 에너지 비(long term average energy ratio)가 샘플로부터 결정된다. 그런 다음, 비슷한 장기 평균 비를 가진 연속 주파수 대역들이 함께 그룹화되어 주요 대역 세트를 형성한다. 주파수 서브밴드의 수 및 주파수 서브밴드의 범위는 조정 가능할 수 있다.

[0044] 공간 주파수 대역 결합기(350)는 중간 및 측면 컴포넌트를 수신하고, 각 컴포넌트에 이득을 적용하고, 중간 및 측면 컴포넌트를 좌측 및 우측 채널로 변환한다. 예를 들어, 공간 주파수 대역 결합기(350)는 향상된 비공간 컴포넌트(E_m) 및 향상된 공간 컴포넌트(E_s)를 수신하고, 향상된 비공간 컴포넌트(E_m) 및 향상된 공간 컴포넌트(E_s)를 공간적으로 향상된 좌측 채널(E_L) 및 공간적으로 향상된 우측 채널(E_R)로 변환하기 전에 글로벌 중간 및 측면 이득을 수행한다.

[0045] 보다 구체적으로, 공간 주파수 대역 결합기(350)는 글로벌 중간 이득부(408), 글로벌 측면 이득부(410), 및 글로벌 중간 이득부(408) 및 글로벌 측면 이득부(410)에 결합된 M/S-L/R 변환기(412)를 포함한다. 글로벌 중간 이득부(408)는 향상된 비공간 컴포넌트(E_m)를 수신하여 이득을 적용하고, 글로벌 측면 이득부(410)는 향상된 비공간 컴포넌트(E_s)를 수신하여 이득을 적용한다. M/S-L/R 변환기(412)는 글로벌 중간 이득부(408)로부터의 향상된 비공간 컴포넌트(E_m) 및 글로벌 측면 이득부(410)로부터의 향상된 공간 컴포넌트(E_s)를 수신하고, 이들 입력을 공간적으로 향상된 좌측 채널(E_L) 및 공간적으로 향상된 우측 채널(E_R)로 변환한다.

[0046] 예시적 크로스토크 보상 프로세서

[0047] 도 5는 일부 실시예에 따른 크로스토크 보상 프로세서(320)의 개략적 블록도이다. 크로스토크 보상 프로세서(320)는 L/R-M/S 변환기(502), 중간 컴포넌트 프로세서(520), 측면 컴포넌트 프로세서(530) 및 M/S-L/R 변환기(514)를 포함한다.

[0048] 크로스토크 보상 프로세서(320)는 좌측 및 우측 입력 채널(X_L , X_R)을 수신하고, 크로스토크 보상 처리를 수행하여 좌측 크로스토크 보상 채널(Z_L) 및 우측 크로스토크 보상 채널(Z_R)을 생성한다. 채널(Z_L , Z_R)은 크로스토크 제거 또는 크로스토크 시뮬레이션과 같은 크로스토크 처리 시의 아티팩트 또는 스펙트럼 결함을 보상하는 데 사용될 수 있다. L/R-M/S 변환기(502)는 좌측 입력 채널(X_L) 및 우측 입력 채널(X_R)을 수신하고, 입력 채널(X_L , X_R)의 비공간 컴포넌트(X_m) 및 공간 컴포넌트(X_s)를 생성한다. 좌측 및 우측 채널은 좌측 및 우측 채널의 비공간 컴포넌트를 생성하기 위해 합산될 수 있고, 좌측 및 우측 채널의 공간 컴포넌트를 생성하기 위해 감산될 수 있다.

[0049] 중간 컴포넌트 프로세서(520)는 m 개의 중간 필터(540(a), 540(b) 내지 540(m))와 같은 복수의 필터(540)를 포함한다. 여기서, m 개의 중간 필터(540) 각각은 비공간 컴포넌트(X_m)의 m 개의 주파수 대역 중 하나를 처리한다. 중간 컴포넌트 프로세서(520)는 비공간적 컴포넌트(X_m)를 처리함으로써 중간 크로스토크 보상 채널(Z_m)을 생성한다. 일부 실시예에서, 중간 필터(540)는 시뮬레이션을 통한 크로스토크 처리와 함께 비공간 컴포넌트(X_m)의 주파수 응답 플롯을 사용하도록 구성된다. 또한, 주파수 응답 플롯을 분석함으로써, 크로스토크 처리의 아티팩트로서 발생하는, 주파수 응답 플롯에서의 미리 결정된 임계값(예컨대, 10dB)을 초과하는 피크(peak) 또는 트로프

(trough)와 같은 임의의 스펙트럼 결함이 추정될 수 있다. 이러한 아티팩트는 주로 크로스토크 처리 시의 지연되고 어찌면 반전된 대측 신호(contralateral signal)와 대응하는 동측 신호(ipsilateral signal)의 합산으로 인해 초래되어 최종 렌더링 결과에 콤 필터와 유사한 주파수 응답을 유효하게 도입한다. 추정된 피크 또는 트로프를 보상하는 중간 컴포넌트 프로세서(520)에 의해 중간 크로스토크 보상 채널(Z_m)이 생성될 수 있는데, 여기서 m 개의 주파수 대역 각각은 피크 또는 트로프에 대응한다. 구체적으로, 스펙트럼의 특정 영역에서 에너지의 가변 증폭 및/또는 감쇠를 야기하는 피크 및 트로프는, 크로스토크 처리에 적용된 특정 지연, 필터링 주파수, 및 이득에 기초하여, 주파수 응답에서 위아래로 이동한다. 중간 필터(540) 각각은 하나 이상의 피크 및 트로프를 조정하도록 구성될 수 있다.

[0050] 측면 컴포넌트 프로세서(530)는 m 개의 측면 필터(550(a), 550(b) 내지 550(m))와 같은 복수의 필터(550)를 포함한다. 측면 컴포넌트 프로세서(530)는 공간 컴포넌트(X_s)를 처리하여 측면 크로스토크 보상 채널(Z_s)을 생성한다. 일부 실시예에서, 크로스토크 처리를 갖는 공간 컴포넌트(X_s)의 주파수 응답 플롯은 시뮬레이션을 통해 획득될 수 있다. 주파수 응답 플롯을 분석함으로써, 크로스토크 처리의 아티팩트로서 발생하는, 주파수 응답 플롯에서의 미리 결정된 임계값(예컨대, 10dB)을 초과하는 피크 또는 트로프와 같은 임의의 스펙트럼 결함이 추정될 수 있다. 추정된 피크 또는 트로프를 보상하는 측면 컴포넌트 프로세서(530)에 의해 측면 크로스토크 보상 채널(Z_s)이 생성될 수 있다. 구체적으로, 스펙트럼의 특정 영역에서 에너지의 가변 증폭 및/또는 감쇠를 야기하는 피크 및 트로프는, 크로스토크 처리에 적용된 특정 지연, 필터링 주파수, 및 이득에 기초하여, 주파수 응답에서 위아래로 이동한다. 측면 필터(550) 각각은 하나 이상의 피크 및 트로프를 조정하도록 구성될 수 있다. 일부 실시예에서, 중간 컴포넌트 프로세서(520) 및 측면 컴포넌트 프로세서(530)는 상이한 수의 필터를 포함할 수 있다.

[0051] 일부 실시예에서, 중간 필터(540) 및 측면 필터(550)는 수학적 식 1에 의해 정의된 전달 함수를 갖는 바이쿼드(biquad) 필터를 포함할 수 있다.

수학적 식 1

[0052]
$$H(z) = \frac{b_0 + b_1z^{-1} + b_2z^{-2}}{a_0 + a_1z^{-1} + a_2z^{-2}}$$

[0053] 여기서, z 는 복소 변수(complex variable)이고, a_0, a_1, a_2, b_0, b_1 및 b_2 는 디지털 필터 계수이다. 이러한 필터를 구현하는 한 가지 방법은 수학적 식 2에 의해 정의된 다이렉트 폼 I 토폴로지(direct form I topology)이다.

수학적 식 2

[0054]
$$Y[n] = \frac{b_0}{a_0}X[n-1] + \frac{b_1}{a_0}X[n-1] + \frac{b_2}{a_0}X[n-2] - \frac{a_1}{a_0}Y[n-1] - \frac{a_2}{a_0}Y[n-2]$$

[0055] 여기서, X 는 입력 벡터이고 Y 는 출력이다. 최대 단어 길이 및 포화 동작에 따라 다른 토폴로지가 사용될 수 있다.

[0056] 그런 다음, 바이쿼드(biquad)를 사용하여 실수치(real-valued) 입력 및 출력을 갖는 2차 필터를 구현할 수 있다. 불연속 시간 필터를 설계하기 위해서는, 연속 시간 필터를 설계한 다음, 이를 이중 선형 변환(bilinear transform)을 통해 불연속 시간으로 변환한다. 또한, 중심 주파수 및 대역폭에서 초래되는 시프트는 주파수 와핑(frequency warping)을 사용하여 보상될 수 있다.

[0057] 예를 들어, 피킹 필터(peaking filter)는 수학적 식 3에 의해 정의된 S-평면 전달 함수를 가질 수 있다.

수학식 3

$$H(s) = \frac{s^2 + s \left(\frac{A}{Q} \right) + 1}{s^2 + s \left(\frac{A}{Q} \right) + 1}$$

[0058]

[0059] 여기서, s 는 복소 변수이고, A 는 피크의 진폭이고, Q 는 필터 "품질"이며, 디지털 필터 계수는 다음과 같이 정의된다.

$$b_0 = 1 + \alpha A$$

[0060]

$$b_1 = -2 * \cos(\omega_0)$$

[0061]

$$b_2 = 1 - \alpha A$$

[0062]

$$\alpha_0 = 1 + \frac{\alpha}{A}$$

[0063]

$$\alpha_1 = -2\cos(\omega_0)$$

[0064]

$$\alpha_2 = 1 + \frac{\alpha}{A}$$

[0065]

[0066] 여기서, ω_0 는 필터의 중심 주파수(라디안)이며, $\alpha = \frac{\sin(\omega_0)}{2Q}$ 이다.

[0067]

[0067] 또한, 필터 품질 Q 는 수학식 4에 의해 정의될 수 있다.

수학식 4

$$Q = \frac{f_c}{\Delta f}$$

[0068]

[0069] 여기서, Δf 는 대역폭이고 f_c 는 중심 주파수이다.

[0070]

[0070] 중간 필터들(540)은 직렬인 것으로 도시되어 있고, 측면 필터들(550)은 직렬인 것으로 도시되어 있다. 일부 실시예에서, 중간 필터(540)는 중간 컴포넌트(X_m)에 병렬로 적용되고, 측면 필터는 측면 컴포넌트(X_s)에 병렬로 적용된다.

[0071] M/S-L/R 변환기(514)는 중간 크로스토크 보상 채널(Z_m) 및 측면 크로스토크 보상 채널(Z_s)을 수신하고, 좌측 크로스토크 보상 채널(Z_L) 및 우측 크로스토크 보상 채널(Z_R)을 생성한다. 중간 및 측면 컴포넌트는 중간 및 측면 컴포넌트의 좌측 채널을 생성하기 위해 합산될 수 있고, 중간 및 측면 컴포넌트는 중간 및 측면 컴포넌트의 우측 채널을 생성하기 위해 감산될 수 있다.

[0072] 예시적 크로스토크 제거 프로세서

[0073]

[0073] 도 6은 일부 실시예에 따른 크로스토크 제거 프로세서(370)의 개략적 블록도이다. 크로스토크 제거 프로세서(370)는 향상된 좌측 보상 채널(T_L) 및 향상된 우측 보상 채널(T_R)을 수신하고, 채널들(T_L , T_R)에 대해 크로스토크 제거를 수행하여 좌측 출력 채널(O_L) 및 우측 출력 채널(O_R)을 생성한다.

[0074] 크로스토크 제거 프로세서(370)는 내-외(in-out) 대역 분할기(610), 인버터(620 및 622), 대역 추정기(630 및 640), 결합기(650 및 652), 및 내-외 대역 결합기(660)를 포함한다. 이 컴포넌트들은 함께 동작하여, 입력 채널들(T_L , T_R)을 대역 내 컴포넌트 및 대역 외 컴포넌트로 분할하고, 대역 내 컴포넌트에 대해 크로스토크 제거를

수행하여 출력 채널(O_L , O_R)을 생성한다.

[0075] 입력 오디오 신호(T)를 상이한 주파수 대역 컴포넌트들로 분할하고 선택적 컴포넌트(예컨대, 대역 내 컴포넌트)에 대해 크로스토크 제거를 수행함으로써, 다른 주파수 대역에서의 악화를 방지하면서 특정 주파수 대역에 대해 크로스토크 제거가 수행될 수 있다. 입력 오디오 신호(T)를 상이한 주파수 대역으로 분할하지 않고 크로스토크 제거가 수행되면, 그러한 크로스토크 제거 후의 오디오 신호는 저주파수(예컨대, 350Hz 미만), 고주파수(예컨대, 12000Hz 초과), 또는 둘 다에서 비공간 및 공간 컴포넌트의 상당한 감쇠 또는 증폭을 나타낼 수 있다. 영향력이 강한 공간 큐(impactful spatial cues)의 대부분이 존재하는 대역 내(예컨대, 250Hz와 14000Hz 사이)에 대해 크로스토크 제거를 선택적으로 수행함으로써, 믹싱의 스펙트럼에 걸쳐, 특히 비공간 컴포넌트에서, 균형 잡힌 전반적 에너지가 유지될 수 있다.

[0076] 내-외 대역 분할기(610)는 입력 채널(T_L , T_R)을 대역 내 채널($T_{L,In}$, $T_{R,In}$) 및 대역 외 채널($T_{L,Out}$, $T_{R,Out}$)로 각각 분리한다. 특히, 내-외 대역 분할기(610)는 향상된 좌측 보상 채널(T_L)을 좌측 대역 내 채널($T_{L,In}$) 및 좌측 대역 외 채널($T_{L,Out}$)로 분할한다. 유사하게, 내-외 대역 분할기(610)는 향상된 우측 보상 채널(T_R)을 우측 대역 내 채널($T_{R,In}$) 및 우측 대역 외 채널($T_{R,Out}$)로 분리한다. 각각의 대역 내 채널은, 예를 들어 250Hz 내지 14kHz를 포함하는 주파수 범위에 대응하는 각각의 입력 채널의 일부를 포함할 수 있다. 주파수 대역의 범위는, 예를 들어 스피커 파라미터에 따라 조정될 수 있다.

[0077] 인버터(620)와 대측 추정기(630)는 함께 동작하여, 좌측 대역 내 채널($T_{L,In}$)로 인한 대측 사운드 컴포넌트를 보상하는 좌측 대측 제거 컴포넌트(S_L)를 생성한다. 유사하게, 인버터(622)와 대측 추정기(640)는 함께 동작하여, 우측 대역 내 채널($T_{R,In}$)로 인한 대측 사운드 컴포넌트를 보상하는 우측 대측 제거 컴포넌트(S_R)를 생성한다.

[0078] 일 접근법에서, 인버터(620)는 대역 내 채널($T_{L,In}$)을 수신하고, 수신된 대역 내 채널($T_{L,In}$)의 극성을 반전시켜 반전된 대역 내 채널($T_{L,In}'$)을 생성한다. 대측 추정기(630)는 반전된 대역 내 채널($T_{L,In}'$)을 수신하고, 필터링을 통해 반전된 대역 내 채널($T_{L,In}'$)에서 대측 사운드 컴포넌트에 대응하는 부분을 추출한다. 반전된 대역 내 채널($T_{L,In}'$)에서 필터링이 수행되기 때문에, 대측 추정기(630)에 의해 추출된 부분은 대역 내 채널(T_L)에서 대측 사운드 컴포넌트에 기여하는 부분의 역수가 된다. 따라서, 대측 추정기(630)에 의해 추출된 부분은 좌측 대측 제거 컴포넌트(S_L)가 되는데, 이는 대역 내 채널($T_{L,In}$)로 인한 대측 사운드 컴포넌트를 감소시키기 위해 상대(counterpart) 대역 내 채널($T_{R,In}$)에 추가될 수 있다. 일부 실시예에서, 인버터(620) 및 대측 추정기(630)는 상이한 순서로 구현된다.

[0079] 인버터(622) 및 대측 추정기(640)는 대역 내 채널($T_{R,In}$)에 대해 유사한 동작을 수행하여 우측 대측 제거 컴포넌트(S_R)를 생성한다. 따라서, 간결성을 위해 그에 대한 상세한 설명은 생략한다.

[0080] 일 예시적 구현에서, 대측 추정기(630)는, 필터(632), 증폭기(634), 및 지연 유닛(636)을 포함한다. 필터(632)는 반전된 입력 채널($T_{L,In}'$)을 수신하고, 필터링 기능을 통해 반전된 대역 내 채널($T_{L,In}'$)에서 대측 사운드 요소에 대응하는 부분을 추출한다. 예시적 필터 구현은 5000 내지 10000Hz 사이에서 선택된 중심 주파수 및 0.5 내지 1.0 사이에서 선택된 Q를 갖는 노치(Notch) 또는 하이 셸프(Highshelf) 필터이다. 데시벨 이득(G_{dB})은 수학식 5로부터 도출될 수 있다.

수학식 5

[0081] $G_{dB} = -3.0 - \log_{1.333}(D)$

[0082] 여기서, D는, 예를 들어 48kHz의 샘플링 속도에서 샘플의 지연 유닛(636 및 646)에 의한 지연량이다. 대안적인 구현은 5000과 10000Hz 사이에서 선택된 코너 주파수 및 0.5와 1.0 사이에서 선택된 Q를 갖는 로우 패스 필터이다. 또한, 증폭기(634)는 추출된 부분을 대응하는 이득 계수($G_{L,In}$)에 의해 증폭시키고, 지연 유닛(636)은 증폭

기(634)로부터 증폭된 출력을 지연 함수(D)에 따라 지연시켜 좌측 대측 제거 컴포넌트(S_L)를 생성한다. 대측 추정기(640)는 필터(642), 증폭기(644), 및 지연 유닛(646)을 포함하여, 반전된 대역 내 채널($T_{R,In}$)에 대해 유사한 동작을 수행하여 우측 대측 제거 컴포넌트(S_R)를 생성한다. 일 예에서, 대측 추정기(630, 640)는 아래의 수학적식에 따라 좌측 및 우측 대측 제거 컴포넌트(S_L, S_R)를 생성한다.

수학적식 6

[0083] $S_L = D[G_{L,In} * F[T_{L,In}]]$

수학적식 7

[0084] $S_R = D[G_{R,In} * F[T_{R,In}]]$

[0085] 여기서 F[]는 필터 함수이고 D[]는 지연 함수이다.

[0086] 크로스토크 제거의 구성은 스피커 파라미터에 의해 결정될 수 있다. 일 예에서, 필터 중심 주파수, 지연량, 증폭기 이득, 및 필터 이득은 청취자에 대해 두 개의 스피커(280) 사이에서 형성된 각도에 따라 결정될 수 있다. 일부 실시예에서, 스피커 각도들 사이의 값은 다른 값을 보간하는 데 사용된다.

[0087] 결합기(650)는 우측 대측 제거 컴포넌트(S_R)와 좌측 대역 내 채널($T_{L,In}$)을 결합하여 좌측 대역 내 크로스토크 채널(U_L)을 생성하고, 결합기(652)는 좌측 대측 제거 컴포넌트(S_L)와 우측 대역 내 채널($T_{R,In}$)을 결합하여 우측 대역 내 크로스토크 채널(U_R)을 생성한다. 내-외 대역 결합기(660)는 좌측 대역 내 크로스토크 채널(U_L)과 대역 외 채널($T_{L,Out}$)을 결합하여 좌측 출력 채널(O_L)을 생성하고, 우측 대역 내 크로스토크 채널(U_R)과 대역 외 채널($T_{R,Out}$)을 결합하여 우측 출력 채널(O_R)을 생성한다.

[0088] 따라서, 좌측 출력 채널(O_L)은 대역 내 채널($T_{R,In}$)에서 대측 사운드에 기여하는 부분의 역수에 대응하는 우측 대측 제거 컴포넌트(S_R)를 포함하고, 우측 출력 채널(O_R)은 대역 내 채널($T_{L,In}$)에서 대측 사운드에 기여하는 부분의 역수에 대응하는 좌측 대측 제거 컴포넌트(S_L)를 포함한다. 이 구성에서, 우측 귀에 도착한, 우측 출력 채널(O_R)에 따라 우측 라우드스피커에 의해 출력된 동측 사운드 컴포넌트의 파면은, 좌측 출력 채널(O_L)에 따라 좌측 라우드스피커에 의해 출력된 대측 사운드 컴포넌트의 파면을 상쇄시킬 수 있다. 유사하게, 좌측 귀에 도착한, 좌측 출력 채널(O_L)에 따라 좌측 라우드스피커에 의해 출력된 동측 사운드 컴포넌트의 파면은, 우측 출력 채널(O_R)에 따라 우측 라우드스피커에 의해 출력된 대측 사운드 컴포넌트의 파면을 상쇄시킬 수 있다. 이와 같이, 대측 사운드 컴포넌트는 공간 검출 가능성을 향상시키기 위해 감소될 수 있다.

[0089] 예시적 공간 향상 프로세서

[0090] 도 7은 일부 실시예에 따른 공간 향상 프로세서(700)의 개략적 블록도이다. 공간 향상 프로세서(700)는, 크로스토크 처리가 헤드폰을 사용하는 클라이언트 디바이스(130)에 대한 크로스토크 시뮬레이션인 공간 향상 프로세서(120)의 예이다. 공간 향상 프로세서(700)는 입력 오디오 신호(X)에 대한 크로스토크 시뮬레이션을 수행하여, 좌측 헤드폰에 대한 좌측 출력 채널(O_L) 및 우측 헤드폰에 대한 우측 출력 채널(O_R)을 포함하는 출력 오디오 신호(O)를 생성한다. 좌측 입력 채널(X_L)은 공간 믹싱 프로세서(110)로부터의 좌측 출력 채널(142)일 수 있고, 우측 입력 채널(X_R)은 공간 믹싱 프로세서(110)로부터의 우측 출력 채널(144)일 수 있다. 향상된 좌측 채널(O_L)은 향상된 좌측 채널(146)에 대응하고, 향상된 우측 채널(O_R)은 향상된 우측 채널(148)에 대응한다.

[0091] 공간 향상 프로세서(700)는 서브밴드 공간 프로세서(310), 크로스토크 보상 프로세서(720), 크로스토크 시뮬레이션 프로세서(780), 및 결합기(760)를 포함한다. 크로스토크 보상 프로세서(720)는 입력 채널들(X_L, X_R)을 수신하고, 크로스토크 시뮬레이션 프로세서(780)에 의해 생성된 좌측 크로스토크 시뮬레이션 신호(W_L) 및 우측 크

로스토크 시물레이션 신호(W_R)를 포함하는 크로스토크 시물레이션 신호(W)와 서브밴드 공간 프로세서(310)에 의해 생성된 향상된 좌측 채널(E_L) 및 향상된 우측 채널(E_R)을 포함하는 향상된 신호(E)의 후속 결합에서의 아티팩트를 보상하는 처리를 수행한다. 크로스토크 보상 프로세서(720)는, 좌측 크로스토크 보상 채널(Z_L) 및 우측 크로스토크 보상 채널(Z_R)을 포함하는 크로스토크 보상 신호(Z)를 생성한다. 크로스토크 시물레이션 프로세서(780)는 좌측 크로스토크 시물레이션 채널(W_L) 및 우측 크로스토크 시물레이션 채널(W_R)을 생성한다. 서브밴드 공간 프로세서(310)는 향상된 좌측 채널(E_L) 및 향상된 우측 채널(E_R)을 생성한다. 크로스토크 시물레이션 프로세서(780)에 관한 추가 세부사항은 도 8과 관련하여 아래에서 설명된다.

[0092] 결합기(760)는, 향상된 좌측 채널(E_L), 향상된 우측 채널(E_R), 좌측 크로스토크 시물레이션 채널(W_L), 우측 크로스토크 시물레이션 채널(W_R), 좌측 크로스토크 보상 채널(Z_L), 및 우측 크로스토크 보상 채널(Z_R)을 수신한다. 결합기(760)는 향상된 좌측 채널(E_L), 우측 크로스토크 시물레이션 채널(W_R) 및 좌측 크로스토크 보상 채널(Z_L)을 결합함으로써 좌측 출력 채널(O_L)을 생성한다. 결합기(760)는 향상된 좌측 채널(E_L), 우측 크로스토크 시물레이션 채널(W_R) 및 좌측 크로스토크 보상 채널(Z_L)을 결합함으로써 우측 출력 채널(O_R)을 생성한다.

[0093] 예시적 크로스토크 시물레이션 프로세서

[0094] 도 8은 일부 실시예에 따른 크로스토크 시물레이션 프로세서(780)의 개략적 블록도이다. 크로스토크 시물레이션 프로세서(780)는, 스테레오 헤드폰으로 출력하기 위한 대측 사운드 컴포넌트를 생성하여, 헤드폰에서 라우드 스피커와 같은 청취 경험을 제공한다. 좌측 입력 채널(X_L)은 공간 믹싱 프로세서(110)로부터의 좌측 출력 채널(142)일 수 있고, 우측 입력 채널(X_R)은 공간 믹싱 프로세서(110)로부터의 우측 출력 채널(144)일 수 있다.

[0095] 크로스토크 시물레이션 프로세서(780)는, 좌측 입력 채널(X_L)을 처리하기 위해, 좌측 헤드 새도우 로우 패스 필터(802), 좌측 헤드 새도우 하이 패스 필터(824), 좌측 크로스토크 지연부(804), 및 좌측 헤드 새도우 이득부(810)를 포함한다. 크로스토크 시물레이션 프로세서(780)는, 우측 입력 채널(X_R)을 처리하기 위해, 우측 헤드 새도우 로우 패스 필터(806), 우측 헤드 새도우 하이 패스 필터(826), 우측 크로스토크 지연부(808), 및 우측 헤드 새도우 이득부(812)를 더 포함한다. 좌측 헤드 새도우 로우 패스 필터(802) 및 좌측 헤드 새도우 하이 패스 필터(824)는, 청취자의 머리를 통과한 이후의 신호의 주파수 응답을 모델링하는 변조를 좌측 입력 채널(X_L)에 적용한다. 좌측 헤드 새도우 하이 패스 필터(824)의 출력은 시간 지연을 적용하는 좌측 크로스토크 지연부(804)에 제공된다. 시간 지연은 동측 사운드 컴포넌트에 대한 대측 사운드 컴포넌트에 의해 횡단되는 트랜스오럴 거리(transaural distance)를 나타낸다. 좌측 헤드 새도우 이득부(810)는 좌측 크로스토크 지연부(804)의 출력에 이득을 적용하여 좌측 크로스토크 시물레이션 채널(W_L)을 생성한다.

[0096] 우측 입력 채널(X_R)에 대해서도 유사하게, 우측 헤드 새도우 로우 패스 필터(806) 및 우측 헤드 새도우 하이 패스 필터(826)는, 청취자의 머리의 주파수 응답을 모델링하는 변조를 우측 입력 채널(X_R)에 적용한다. 우측 헤드 새도우 하이 패스 필터(826)의 출력은 시간 지연을 적용하는 우측 크로스토크 지연부(808)에 제공된다. 우측 헤드 새도우 이득부(812)는 우측 크로스토크 지연부(808)의 출력에 이득을 적용하여 우측 크로스토크 시물레이션 채널(W_R)을 생성한다.

[0097] 좌측 및 우측 채널 각각에 대한 헤드 새도우 로우 패스 필터, 헤드 새도우 하이 패스 필터, 크로스토크 지연, 및 헤드 새도우 이득의 적용은 상이한 순서로 수행될 수 있다.

[0098] 예시적 회의 처리

[0099] 도 9는 일부 실시예에 따른, 클라이언트 디바이스들 간의 회의를 위한 프로세스(900)의 흐름도이다. 프로세스(900)는, 클라이언트 디바이스(130)와 회의하는 다른 클라이언트 디바이스(130)로부터의 입력 오디오 스트림을 포함하는 공간적으로 향상된 스테레오폰닉(stereophonic) 오디오 스트림을 클라이언트 디바이스(130)에 제공하기 위해 수행된다. 프로세스(900)는 더 적거나 추가적인 단계를 포함할 수 있고, 단계들은 상이한 순서로 수행될 수 있다. 프로세스(900)는, 일부 실시예에 따라, 오디오 처리 시스템(100)의 도면을 각각 포함하는 도 10 및 11을 참조하여 설명된다.

[0100] (예컨대, 공간 믹싱 프로세서(110)를 갖는) 서버(102)는 다수의 클라이언트 디바이스(130)로부터 입력 오디오

스트림을 수신한다(905). 일 예에서, 오디오 스트림은 모노포닉(monophonic) 오디오 스트림을 포함할 수 있다. 다른 예에서, 하나 이상의 입력 오디오 스트림은 다수의 오디오 채널(예컨대, 스테레오, 5.1 채널 서라운드 사운드, 1차 앰비소닉(1st order ambisonics) 등)을 포함할 수 있다. 입력 오디오 스트림은 네트워크를 통해 다수의 클라이언트 디바이스(130)로부터 전송될 수 있다. 일부 실시예에서, 각각의 클라이언트 디바이스(130)는 입력 오디오 스트림과 관련하여 메타데이터를 전송한다. 메타데이터는 클라이언트 디바이스를 식별하는 디바이스 식별 데이터, 다른 클라이언트 디바이스의 다른 오디오 스트림의 원하는 공간 위치를 정의하는 배치 데이터를 포함할 수 있다.

[0101] (예컨대, 공간 믹싱 프로세서(110)를 갖는) 서버(102)는, 클라이언트 디바이스(130)에 대해, 음장 내에서 다른 클라이언트 디바이스(130)의 공간 위치를 정의하는 배치 데이터를 결정한다(910). 배치 데이터는 다양한 방식으로 정의될 수 있다. 예를 들어, 배치 데이터는 클라이언트 디바이스로부터 메타데이터로서 수신될 수 있거나, 서버(102)에 의해 미리 정의될 수 있다. 일부 실시예에서, 클라이언트 디바이스(130)는 사용자 패닝 제어없이 자동으로 패닝된다. 선택적으로, 사용자는 자신의 선호에 따라 음성을 패닝할 수 있다.

[0102] (예컨대, 공간 믹싱 프로세서(110)를 갖는) 서버(102)는 배치 데이터에 따라 입력 오디오 스트림을 믹싱 및 패닝함으로써 믹싱된 스트림을 생성한다(915). 믹싱된 스트림은 믹싱된 좌측 채널 및 믹싱된 우측 채널을 포함하는 스테레오 오디오 신호일 수 있다. 예를 들어, 공간 믹싱 프로세서(110)는 음장 내에서 오디오 스트림(140)에 대한 공간 위치를 결정하고, 오디오 스트림(140)과 연관된 공간 위치에 따라 오디오 스트림(140)으로부터 좌측 채널 및 우측 채널을 생성할 수 있다. 좌측 및 우측 채널은 다른 클라이언트 디바이스들(130)의 각 오디오 스트림(140)에 대해 생성된다. 다수의 오디오 스트림(140)의 좌측 채널은 믹싱된 좌측 채널(142)에 결합되고, 다수의 오디오 스트림(140)의 우측 채널은 믹싱된 우측 채널(144)에 결합된다.

[0103] 도 10을 참조하면, 다수의 클라이언트 디바이스(130a, 130b, 130c 및 130d)가 회의 중이고 서버(110)에 접속되어 있다. 클라이언트 디바이스(130a)에 대해, 클라이언트 디바이스(130a) 주위에 정의된 음장(1002)이 도시되어 있다. 다른 클라이언트 디바이스들(130b, 130c 및 130d)은 배치 데이터에 따라 음장(1002) 내에 위치된다. 이격된 클라이언트 디바이스들(130)의 균일한 공간 분포가 음장(1002) 내에 도시되어 있지만, 공간 분포가 반드시 균일한 것은 아니다. 예를 들어, 클라이언트 디바이스들(130)은 사용자 역할, 엔티티/회사 이름, 그룹 멤버십, 룸 위치, 지리적 위치, IP 주소 등을 정의하는 배치 데이터에 기초하여 음장(1002) 내에 분포될 수 있다.

[0104] (예컨대, 공간 향상 프로세서(120)를 갖는) 서버(102)는 믹싱된 스트림에 서브밴드 공간 처리 및 크로스토크 처리를 적용함으로써 공간적으로 향상된 스트림을 생성한다(920). 공간적으로 향상된 스트림은 향상된 좌측 채널 및 향상된 우측 채널을 포함하는 스테레오 오디오 신호일 수 있다. 공간 향상 프로세서(120)는 다양한 유형의 공간 향상을 수행할 수 있다. 서브밴드 공간 처리를 위해, 공간 향상 프로세서(120)는, 믹싱된 스트림의 믹싱된 좌측 채널 및 믹싱된 우측 채널의 중간 및 측면 서브밴드 컴포넌트를 이득 조정함으로써 좌측 채널 및 우측 채널을 생성한다. 크로스토크 처리는, 예컨대, 클라이언트 디바이스(130)가 라우드스피커를 사용하는지 또는 헤드폰을 사용하는지에 의존하여, 크로스토크 제거 또는 크로스토크 시뮬레이션을 포함할 수 있다. 일부 실시예에서는, 상이한 클라이언트 디바이스(130)에 대해 상이한 유형의 공간 향상이 수행될 수 있다.

[0105] 일부 실시예에서, 클라이언트 디바이스(130)에 대해 믹싱된 스트림에 적용된 공간 향상 처리는, 클라이언트 디바이스(130)로부터의 디바이스 정보에 기초할 수 있다. 디바이스 정보는, 스피커 파라미터(예컨대, 크기, 주파수 응답, 위치, 방향 등)와 같은 공간 향상 프로세서(120)의 컴포넌트에 대한 파라미터, 또는 필터 값, 이득 값, 시간 지연 값, 공간 향상 효과의 세기 등과 같은 공간 향상 프로세서의 동작 파라미터를 포함할 수 있다. 일부 실시예에서, 공간 향상 프로세서(120)에 대한 동작 파라미터는 스피커 파라미터로부터 도출된다. 상이한 유형의 클라이언트 디바이스(130)에 대해 상이한 유형의 공간 향상이 사용될 수 있다. 예를 들어, 클라이언트 디바이스(130)는 스피커의 유형(예컨대, 라우드스피커 또는 헤드폰) 및 스피커의 파라미터(예컨대, 주파수 응답, 상대적 크기, 상대적 방향 등)에 기초하여 달라질 수 있다.

[0106] 일부 실시예에서, 서버(102)는 각각의 클라이언트 디바이스로부터 디바이스 설명을 수신하고, 디바이스 설명에 기초하여 공간 향상 프로세서(120)의 파라미터를 결정한다. 디바이스 설명은, 랩탑, 사운드바, 헤드폰, 라우드스피커, 데스크탑 스피커 등과 같은 클라이언트 디바이스의 유형을 식별할 수 있다. 상이한 유형의 클라이언트 디바이스는 상이한 파라미터와 연관될 수 있고, 파라미터는 디바이스 유형에 기초하여 선택된다.

[0107] 서버(102)는 공간적으로 향상된 스트림을 클라이언트 디바이스(130)에 제공한다(925). 예를 들어, 서버(102)는 향상된 좌측 채널 및 향상된 우측 채널을 포함하는 공간적으로 향상된 스트림을 네트워크를 통해 클라이언트 디바이스(130)에 전송한다. 공간적으로 향상된 스트림에 따라 사운드를 재생하기 위해, 향상된 좌측 채널은 좌측

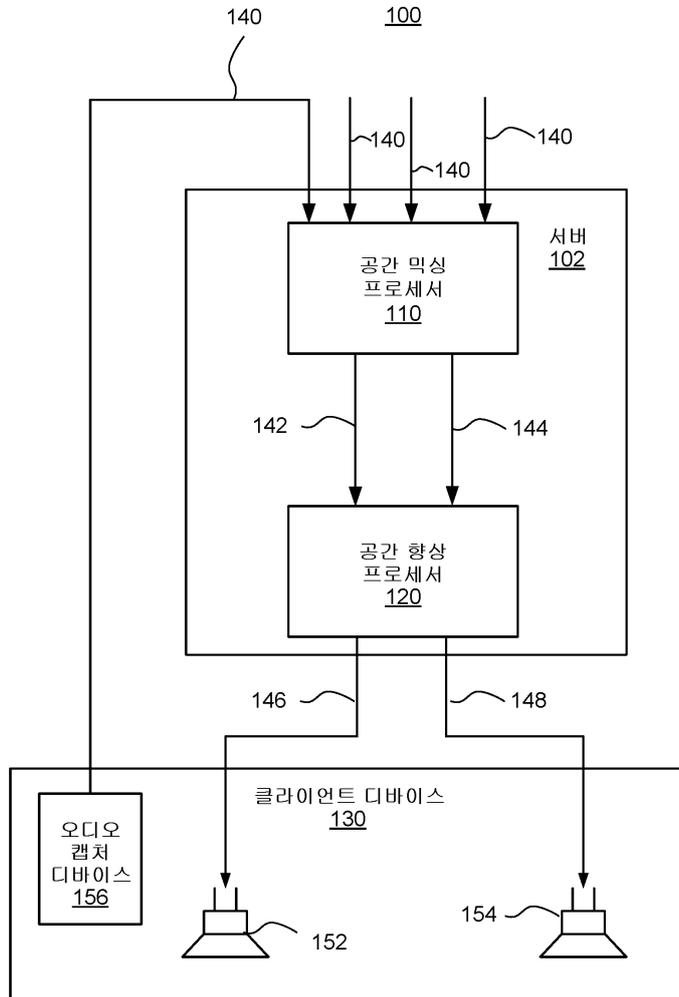
스피커(152)에 제공되고 향상된 우측 채널은 우측 스피커(154)에 제공된다.

- [0108] 일부 실시예에서, 서버(102)(예컨대, 공간 믹싱 프로세서(110))는 공간적으로 향상된 스트림(또는 믹싱된 스트림)과 관련하여 패닝 메타데이터를 각각의 클라이언트 디바이스(130)에 제공한다. 패닝 메타데이터는 오디오 스트림 내에 참가자를 식별하는 정보를 포함할 수 있다. 참가자는 사용자 역할, 엔터티/회사 이름, 그룹 멤버십, 룸 위치, 지리적 위치, IP 주소에 의해 식별될 수 있다.
- [0109] 일부 실시예에서, 서버(102)는 각 클라이언트 디바이스(130)에 의한 공간 향상 처리를 위해 믹싱된 스트림을 클라이언트 디바이스(130)에 제공한다. 클라이언트 디바이스(130)는 믹싱된 스트림을 사용하여 향상된 좌측 및 우측 채널을 생성하는 공간 향상 프로세서(120)를 포함한다. 도 11을 참조하면, 각각의 클라이언트 디바이스(130a 내지 130d)는 공간 믹싱 프로세서(110)로부터의 믹싱된 스트림을 처리하는 공간 향상 프로세서(120)를 포함한다. 공간 향상 프로세서(120)의 컴포넌트에 대한 파라미터는 클라이언트 디바이스(120)에서 공간 향상 프로세서(120)에 의해 사용되며, 파라미터 또는 디바이스 정보는 서버(102)에 전송될 필요가 없다. 전송한 바와 같이, 상이한 유형의 클라이언트 디바이스(130) 및 스피커 구성에 대해 상이한 유형의 공간 향상 처리가 사용될 수 있다.
- [0110] 일부 실시예에서, 서버(102)는 오디오 신호의 믹싱되지 않은 스트림(140)을 클라이언트 디바이스(130)에 제공하는데, 클라이언트 디바이스(130)는 공간 향상이 수반되는 스테레오 믹스다운을 처리한다. 도 12를 참조하면, 각각의 클라이언트 디바이스(130a 내지 130d)는, 믹스다운을 수행하여 믹싱된 스트림을 생성하는 공간 믹싱 프로세서(110), 및 공간 믹싱 프로세서(110)로부터의 믹싱된 스트림을 처리하는 공간 향상 프로세서(120)를 포함한다. 서버(102)는, 클라이언트 디바이스들(130a 내지 130d)을 접속하고 각각의 클라이언트 디바이스(130)에 다른 클라이언트 디바이스(130)로부터의 입력 신호(140)를 제공하는 중개자를 제공한다. 피어-투-피어 구성과 같은 일부 실시예에서, 각각의 클라이언트 디바이스(130)는 서버(102) 없이 서로 접속된다. 여기서, 각각의 클라이언트 디바이스(130)는 다른 클라이언트 디바이스(130)에 오디오 신호(140)를 제공한다.
- [0111] 단계들(910 내지 925)은 회의에 참여하는 각각의 클라이언트 디바이스(130)에 대해 병렬로 수행될 수 있다. 하나 이상의 다른 클라이언트 디바이스(130)의 입력 오디오 스트림으로부터 각각의 클라이언트 디바이스(130)에 대해 상이한 믹싱된 스트림이 생성될 수 있다. 서버(102)는 각각의 클라이언트 디바이스(130)에 대해 각각 상이한 믹싱된 스트림을 처리하는 공간 향상 프로세서(120)의 복수의 인스턴스를 포함할 수 있다. 또한, 단계들(910-925)은 회의에 참여하는 각각의 클라이언트 디바이스(130)에 오디오 스트림 입력 및 출력을 제공하기 위해 반복될 수 있다.
- [0112] 예시적 컴퓨터
- [0113] 도 13은 일 실시예에 따른 컴퓨터(1300)의 개략적 블록도이다. 컴퓨터(1300)는 오디오 시스템을 구현하는 회로의 예이다. 칩셋(1304)에 결합된 적어도 하나의 프로세서(1302)가 도시되어 있다. 칩셋(1304)은 메모리 제어기 허브(1320) 및 입력/출력(I/O) 제어기 허브(1322)를 포함한다. 메모리(1306) 및 그래픽 어댑터(1312)가 메모리 제어기 허브(1320)에 결합되고, 디스플레이 디바이스(1318)가 그래픽 어댑터(1312)에 결합된다. 저장 디바이스(1308), 키보드(1310), 포인팅 디바이스(1314), 및 네트워크 어댑터(1316)가, I/O 제어기 허브(1322)에 결합된다. 컴퓨터(1300)는 다양한 유형의 입력 또는 출력 디바이스를 포함할 수 있다. 컴퓨터(1300)의 다른 실시예는 상이한 아키텍처를 갖는다. 예를 들어, 메모리(1306)는 일부 실시예에서 프로세서(1302)에 직접 결합된다.
- [0114] 저장 디바이스(1308)는, 하드 드라이브, CD-ROM(compact disk read-only memory), DVD 또는 솔리드 스테이트 메모리 디바이스와 같은 하나 이상의 비일시적 컴퓨터 판독가능 저장 매체를 포함한다. 메모리(1306)는 프로세서(1302)에 의해 사용되는 프로그램 코드(하나 이상의 명령어로 구성됨) 및 데이터를 보유한다. 프로그램 코드는 도 1 내지 도 11을 사용하여 설명된 처리 양상에 대응할 수 있다.
- [0115] 포인팅 디바이스(1314)는 컴퓨터 시스템(1300)에 데이터를 입력하기 위해 키보드(1310)와 조합하여 사용된다. 그래픽 어댑터(1312)는 디스플레이 디바이스(1318)에 이미지 및 다른 정보를 디스플레이한다. 일부 실시예에서, 디스플레이 디바이스(1318)는 사용자 입력 및 선택을 수신하기 위한 터치 스크린 기능을 포함한다. 네트워크 어댑터(1316)는 컴퓨터 시스템(1300)을 네트워크에 결합시킨다. 컴퓨터(1300)의 일부 실시예는 도 13에 도시된 것과 상이한 및/또는 다른 컴포넌트를 갖는다.
- [0116] 특정 실시예 및 애플리케이션이 도시되고 설명되었지만, 본 발명은 여기에 개시된 정확한 구성 및 컴포넌트로 제한되지 않으며, 본 개시의 사상 및 범위를 벗어나지 않으면서, 본 명세서에 개시된 방법 및 장치의 구성, 동

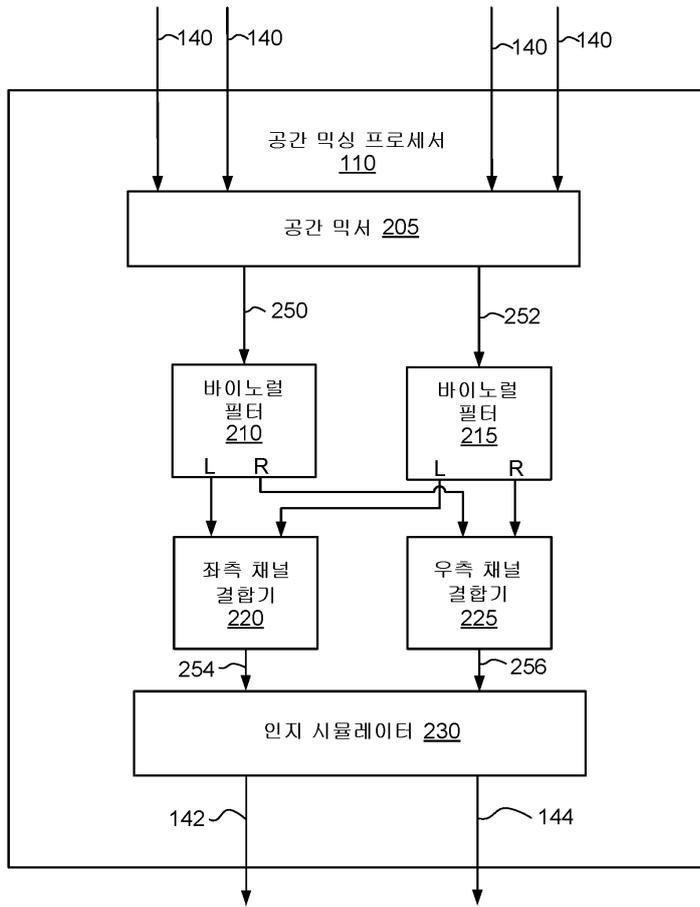
작 및 세부사항에서 당업자에게 명백한 다양한 수정, 변경 및 변형이 이루어질 수 있음이 이해되어야 한다.

도면

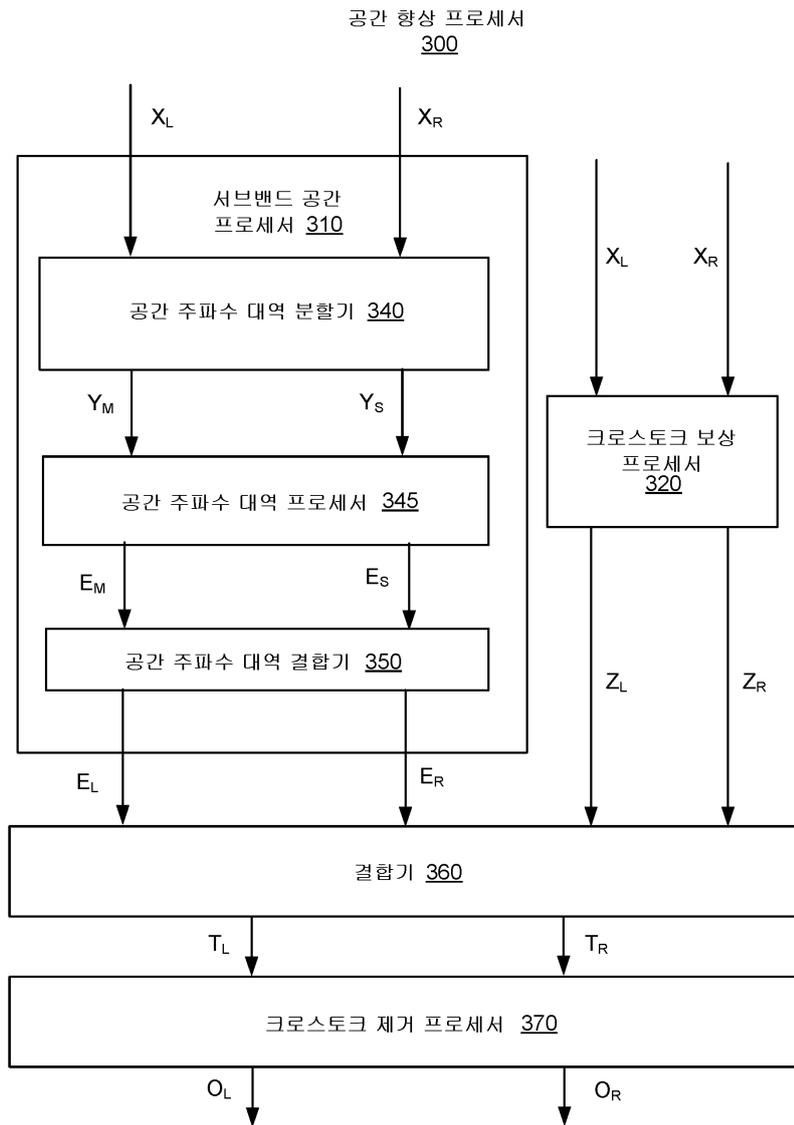
도면1



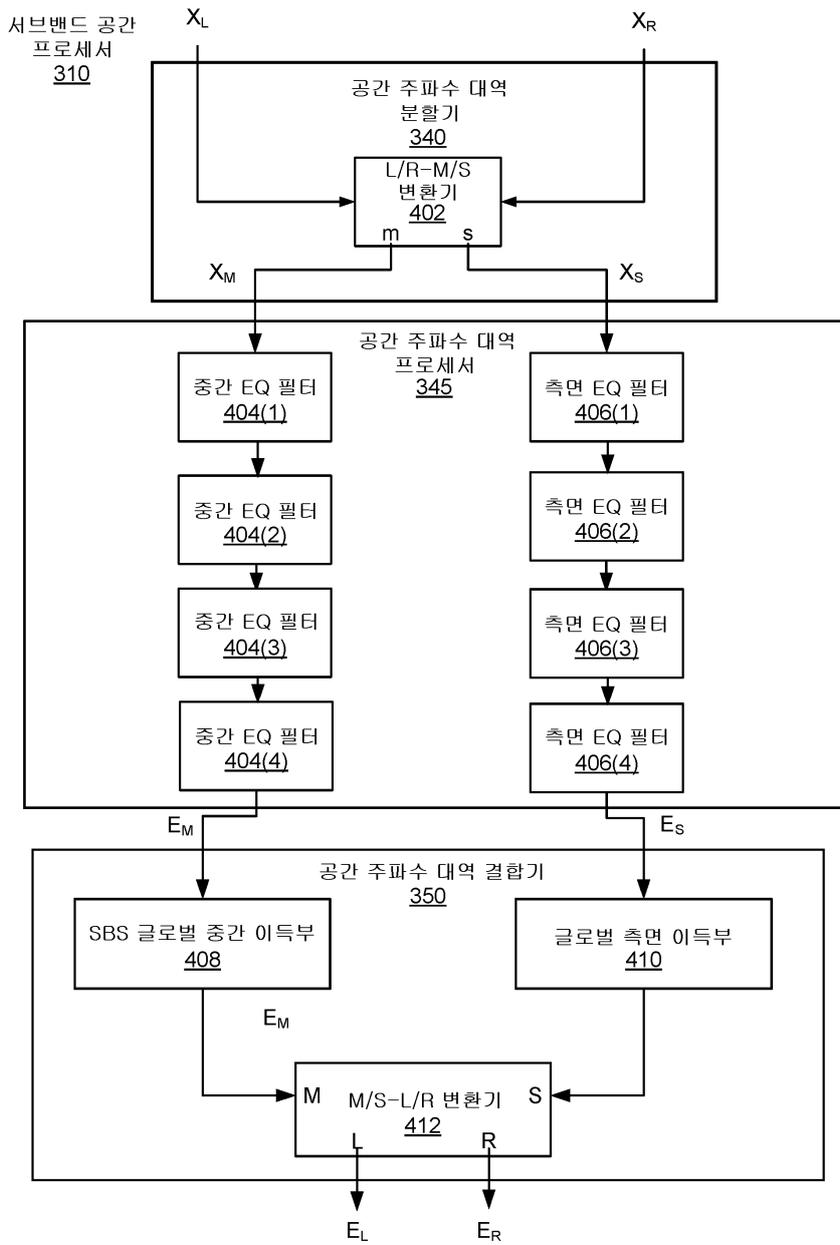
도면2



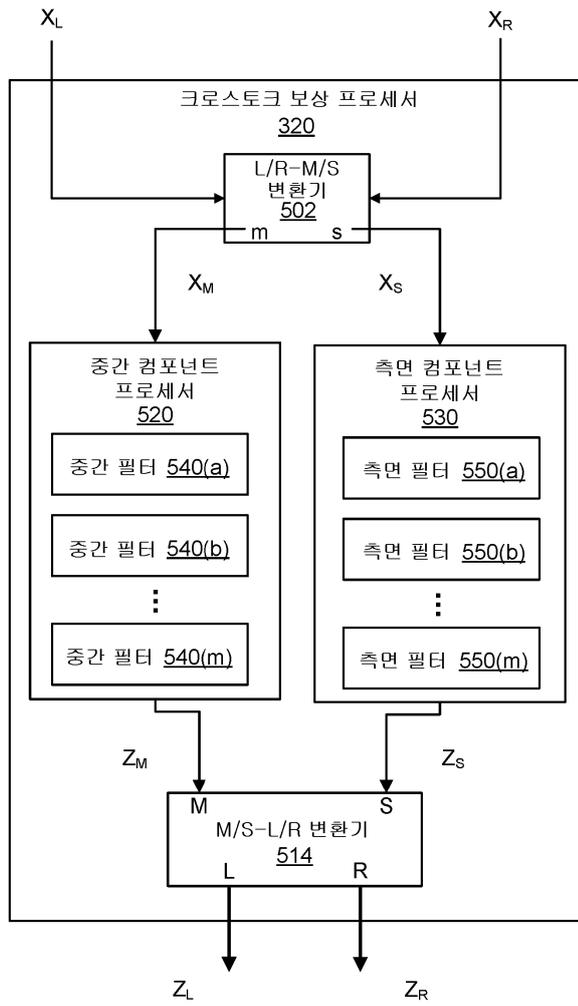
도면3



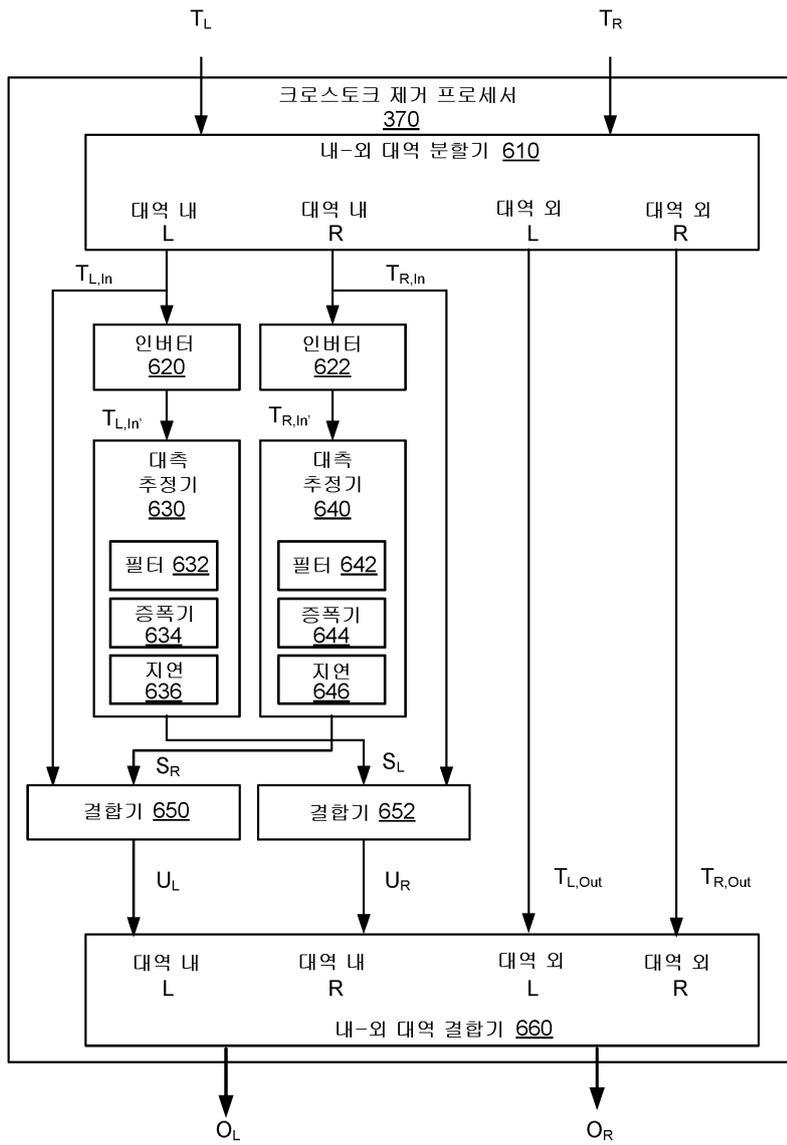
도면4



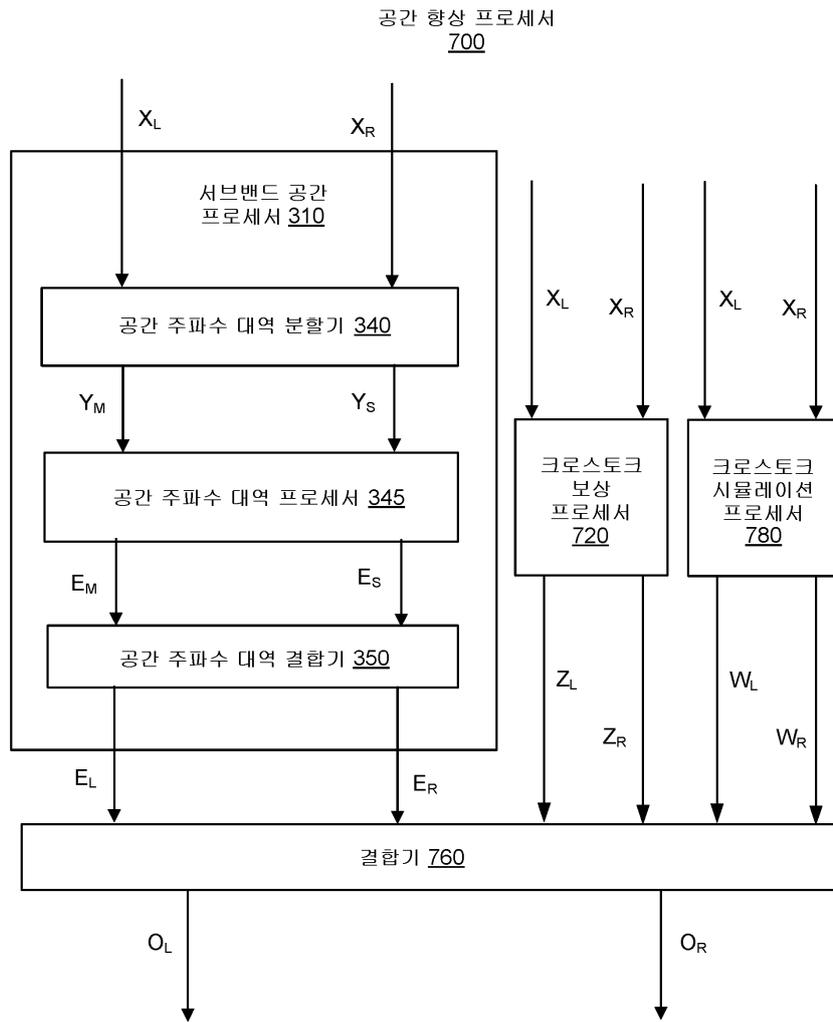
도면5



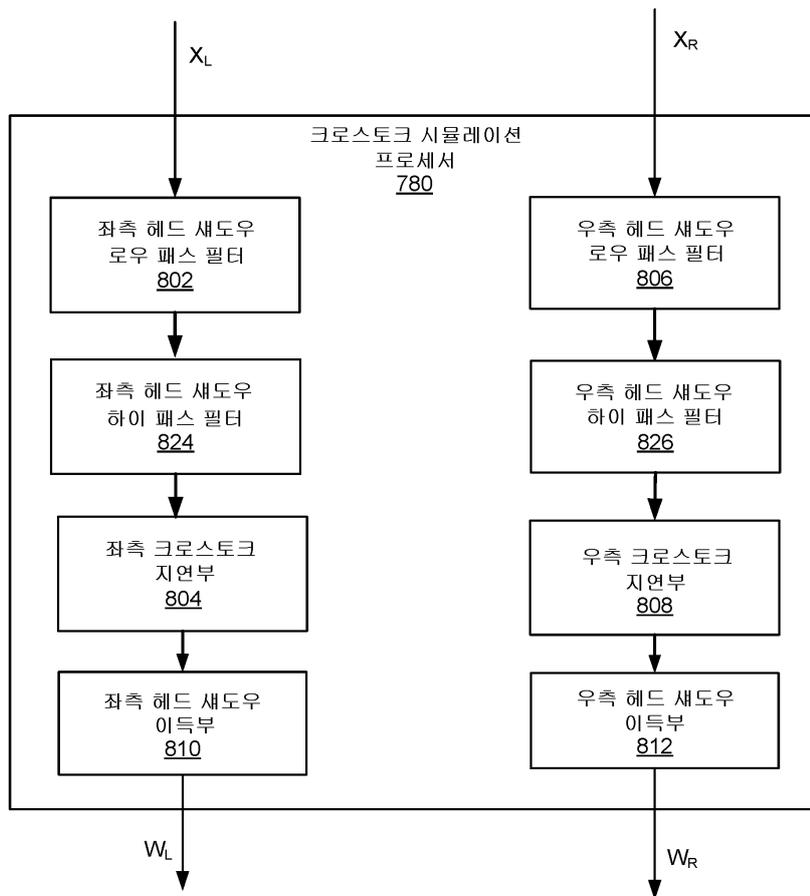
도면6



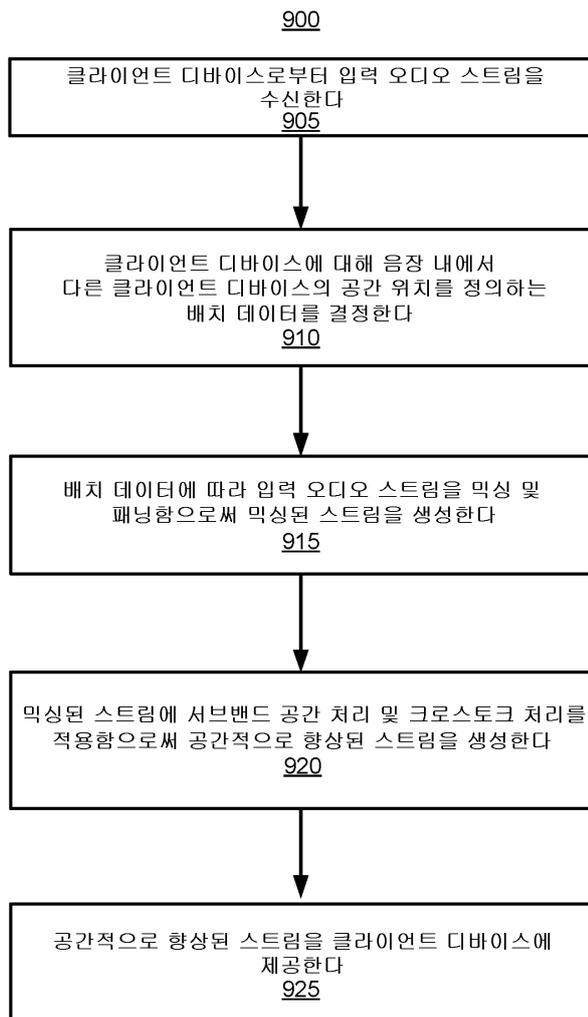
도면7



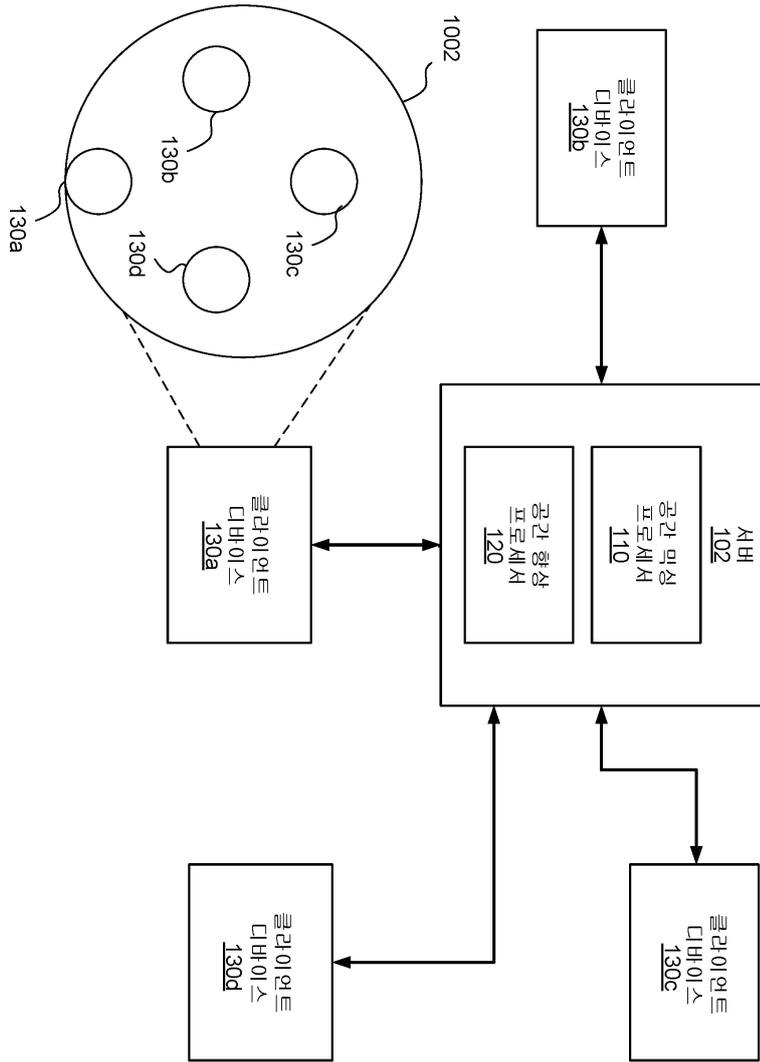
도면8



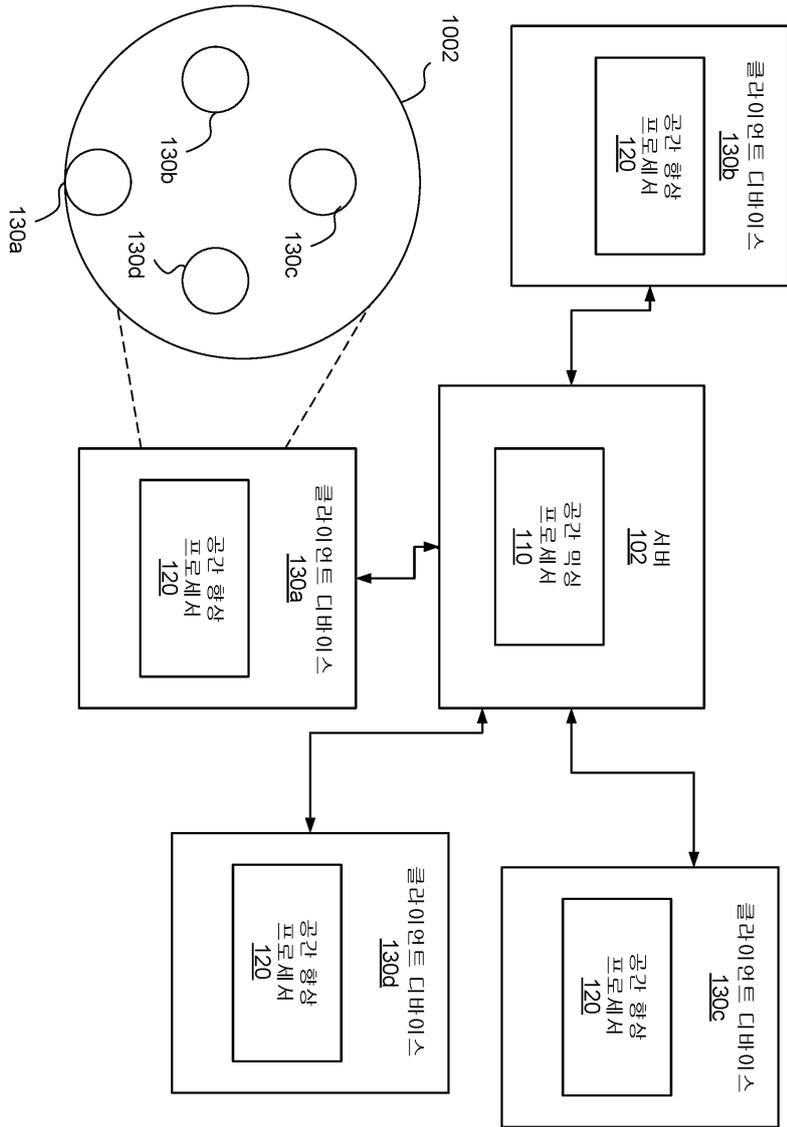
도면9



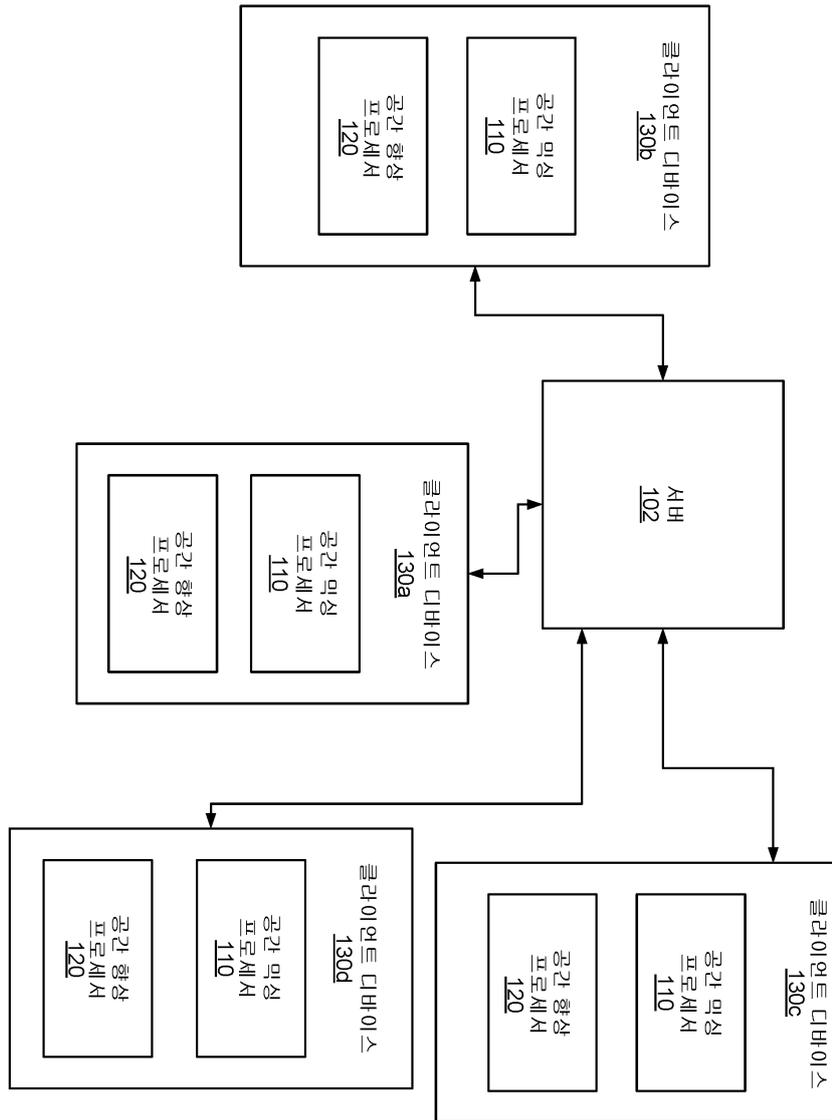
도면10



도면11



도면12



도면13

