



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0052586
(43) 공개일자 2017년05월12일

(51) 국제특허분류(Int. Cl.)
G10L 21/0208 (2013.01) G10L 21/0272 (2013.01)
H04S 7/00 (2006.01)
(52) CPC특허분류
G10L 21/0208 (2013.01)
G10L 21/0272 (2013.01)
(21) 출원번호 10-2017-7006418
(22) 출원일자(국제) 2015년07월23일
심사청구일자 없음
(85) 번역문제출일자 2017년03월08일
(86) 국제출원번호 PCT/US2015/041843
(87) 국제공개번호 WO 2016/039867
국제공개일자 2016년03월17일
(30) 우선권주장
14/483,044 2014년09월10일 미국(US)

(71) 출원인
하만인터내셔널인더스트리스인코포레이티드
미국 코네티컷 스태포드 애틀랜틱 스트리트400 (우: 06901)
(72) 발명자
디 센소, 다비드
미국, 캘리포니아주 94610, 오클랜드, 아파트. 106, 오렌지 스트리트 354
마티, 스테판
미국, 캘리포니아주 94618, 오클랜드, 스타뷰 드라이브 22
주네자, 아제이
미국, 캘리포니아주 94043, 마운틴 뷰, 아파트. 비, 2120 더블유. 미들필드 로드
(74) 대리인
백윤재

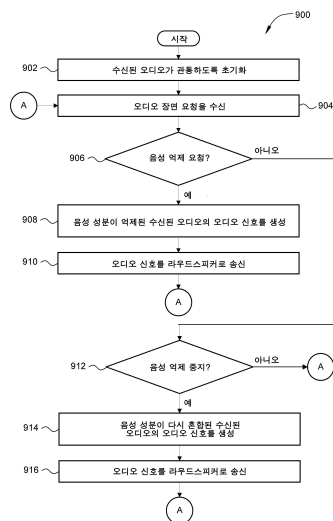
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 청각 디바이스들을 통해 다중 청취 환경을 생성하기 위한 기술들

(57) 요약

청각 장면들을 생성하기 위한 접근법들이 개시된다. 컴퓨팅 디바이스는 무선 네트워크 인터페이스 및 프로세서를 포함한다. 프로세서는 마이크를 통해, 제1 복수의 음성 성분을 포함하는 제1 청각 신호를 수신하도록 구성된다. 프로세서는 제1 복수의 음성 성분에 포함된 제1 음성 성분을 적어도 부분적으로 억제하기 위한 요청을 수신하도록 더 구성된다. 프로세서는 제1 음성 성분이 적어도 부분적으로 억제된 제1 복수의 음성 성분을 포함하는 제2 청각 신호를 생성하도록 더 구성된다. 프로세서는 제2 청각 신호를 출력용 스피커로 송신하도록 더 구성된다.

대표도 - 도9



(52) CPC특허분류

H04S 7/30 (2013.01)

G10L 2021/02087 (2013.01)

명세서

청구범위

청구항 1

제1 청각 장면 제어기와 연관된 청각 장면들을 생성하기 위한 컴퓨터-구현 방법으로서,
 마이크로폰으로부터, 제1 복수의 음성 성분을 포함하는 제1 청각 신호를 수신하고;
 상기 제1 복수의 음성 성분에 포함된 제1 음성 성분을 적어도 부분적으로 억제하기 위한 요청을 수신하며;
 상기 제1 음성 성분이 적어도 부분적으로 억제된 상기 제1 복수의 음성 성분을 포함하는 제2 청각 신호를 생성하고, 그리고
 상기 제2 청각 신호를 출력용 스피커로 송신하는 것을 포함하는, 컴퓨터-구현 방법.

청구항 2

청구항 1에 있어서, 상기 제2 청각 장면 제어기와 연관된 제2 복수의 음성 성분에 포함된 제2 음성 성분을 억제하기 위한 요청을 제2 청각 장면 제어기로 송신하는 것을 더 포함하는, 컴퓨터-구현 방법.

청구항 3

청구항 1에 있어서,
 상기 제1 음성 성분을 억제하는 것을 중단하기 위한 요청을 수신하고;
 상기 제1 음성 성분이 억제되지 않은 상기 제1 복수의 음성 성분을 포함하는 제3 청각 신호를 생성하고, 그리고
 상기 제3 청각 신호를 상기 출력용 스피커로 송신하는 것을 더 포함하는, 컴퓨터-구현 방법.

청구항 4

청구항 1에 있어서, 상기 제2 청각 신호를 생성하는 것은 상기 제1 음성 성분이 완전히 억제된 상기 제1 복수의 음성 성분을 포함하는 상기 제2 청각 신호를 생성하는 것을 포함하는, 컴퓨터-구현 방법.

청구항 5

청구항 1에 있어서, 상기 제1 복수의 음성 성분에 포함된 제2 음성 성분을 억제하기 위한 요청을 수신하는 것을 더 포함하고, 상기 제2 청각 신호를 생성하는 것은 상기 제1 청각 신호에 포함된 제2 음성 성분을 부분적으로 억제하는 것을 포함하는, 컴퓨터-구현 방법.

청구항 6

청구항 1에 있어서, 상기 제2 청각 신호를 생성하는 것은 상기 마이크로폰에 의해 수신된 배경 소음을 억제하는 것을 더 포함하는, 컴퓨터-구현 방법.

청구항 7

청구항 1에 있어서, 상기 마이크로폰 및 상기 스피커는 인-이어 청각 디바이스(in-ear auditory device)에 내장된, 컴퓨터-구현 방법.

청구항 8

청구항 1에 있어서, 상기 마이크로폰 및 상기 스피커가 헤드-장착 청각 디바이스(head-mounted auditory device)에 내장되는, 컴퓨터-구현 방법.

청구항 9

청구항 1에 있어서, 상기 마이크로폰 및 상기 스피커가 의자 또는 시트와 연관된 헤드레스트(headrest)에 내장

된, 컴퓨터-구현 방법

청구항 10

청구항 1에 있어서, 상기 스피커는 고 지향성 스피커(highly directional speaker)를 포함하는, 컴퓨터-구현 방법.

청구항 11

컴퓨팅 디바이스로서,

무선 네트워크 인터페이스; 및

프로세서를 포함하며, 상기 프로세서는:

마이크로폰을 통해, 제1 복수의 음성 성분을 포함하는 제1 청각 신호를 수신하도록;

상기 제1 복수의 음성 성분에 포함된 제1 음성 성분을 적어도 부분적으로 억제하기 위한 요청을 수신하도록;

상기 제1 음성 성분이 적어도 부분적으로 억제된 상기 제1 복수의 음성 성분을 포함하는 제2 청각 신호를 생성하도록; 그리고

상기 제2 청각 신호를 출력용 스피커로 송신하도록 구성된, 컴퓨팅 디바이스.

청구항 12

청구항 11에 있어서, 상기 제1 음성 성분을 억제하기 위한 상기 요청은 상기 무선 네트워크 인터페이스를 통해 수신되는, 컴퓨팅 디바이스.

청구항 13

청구항 11에 있어서, 입력 디바이스를 더 포함하고, 상기 제1 음성 성분을 억제하기 위한 상기 요청은 상기 입력 디바이스의 작동을 통해 수신되는, 컴퓨팅 디바이스.

청구항 14

청구항 11에 있어서, 상기 프로세싱 유닛은 제2 음성 성분을 억제하기 위한 제1 청각 장면 제어기로 지향된 요청을 상기 무선 네트워크 인터페이스를 통해 송신하도록 더 구성되고, 상기 제1 청각 장면 제어기는 상기 제2 음성 성분이 적어도 부분적으로 억제된 상기 제1 복수의 음성 성분을 포함하는 제3 청각 신호를 생성하도록 구성되는, 컴퓨팅 디바이스.

청구항 15

청구항 11에 있어서, 상기 프로세싱 유닛은 상기 컴퓨팅 디바이스와 연관된 음성 성분을 제외한 모든 음성 성분을 적어도 부분적으로 억제하기 위한 제1 청각 장면 제어기로 지향된 요청을 상기 무선 네트워크 인터페이스를 통해 송신하도록 더 구성되는, 컴퓨팅 디바이스.

청구항 16

청구항 15에 있어서, 상기 프로세싱 유닛은 상기 제1 청각 장면 제어기와 연관된 음성 성분을 제외한 모든 음성 성분을 억제하기 위한 요청을 상기 무선 네트워크 인터페이스를 통해 수신하도록 더 구성되는, 컴퓨팅 디바이스.

청구항 17

청구항 15에 있어서, 상기 프로세싱 유닛은 상기 제1 청각 장면 제어기 및 상기 제2 청각 장면 제어기와 연관된 음성 성분들을 제외한 모든 음성 성분을 억제하기 위한 요청을 상기 무선 네트워크 인터페이스를 통해 수신하도록 더 구성되는, 컴퓨팅 디바이스.

청구항 18

청구항 17에 있어서, 상기 제1 청각 장면 제어기는 상기 제2 청각 장면 제어기와 연관된 상기 음성 성분을 부분

적으로 억제하도록 그리고 상기 컴퓨팅 디바이스와 연관된 상기 음성 성분을 제외한 모든 기타 음성 성분을 완전히 억제하도록 구성되는, 컴퓨팅 디바이스.

청구항 19

프로세싱 유닛에 의해 실행될 때, 상기 프로세싱 유닛이 다음 단계들을 수행함으로써 청각 장면을 생성하게 하는, 프로그램 명령들을 저장하는 비-일시적 컴퓨터-판독가능한 매체로서, 상기 단계들은:

마이크로폰으로부터, 제1 복수의 음성 성분을 포함하는 제1 청각 신호를 수신하는 단계;

상기 제1 복수의 음성 성분에 포함된 제1 음성 성분을 억제하기 위한 요청을 수신하는 단계;

상기 제1 음성 성분이 억제된 상기 제1 복수의 음성 성분을 포함하는 제2 청각 신호를 생성하는 단계; 및

상기 제2 청각 신호를 출력용 스피커로 송신하는 단계인, 비-일시적 컴퓨터-판독가능한 매체.

청구항 20

청구항 19에 있어서, 상기 제2 청각 장면 제어기와 연관된 제2 복수의 음성 성분에 포함된 제2 음성 성분을 억제하기 위한 요청을 제2 청각 장면 제어기로 송신하는 단계를 더 포함하는, 비-일시적 컴퓨터-판독가능한 매체.

발명의 설명

기술 분야

[0001] 관련 출원들과 관련된 상호-인용

[0002] 본 출원은 2014년 9월 10일에 출원된 미국 특허 출원 제14/483,044호의 혜택을 주장하며, 이는 이에 의해 여기에 인용에 의한 삽입된다.

[0003] 기술분야

[0004] 본 발명의 실시예들은 일반적으로 인간-디바이스 인터페이스들, 보다 구체적으로, 청각 디바이스들을 통해 다중 청취 환경을 생성하기 위한 기술들에 관한 것이다.

배경 기술

[0005] 다양한 상황에서, 사람들은 보통 한 명 이상의 다른 사람이 있는 데서 사적인 대화를 말할 필요 또는 요구를 발견한다. 예를 들어, 그리고 제한 없이, 컨퍼런스 미팅에 참여하는 사람이 미팅 중 중요한 전화를 받을 수 있다. 미팅 중단을 막기 위해, 그 사람은 물리적으로 방을 나가거나 전화받지 않기를 선택할 수 있다. 다른 예에서, 그리고 제한 없이, 차를 탄 사람은 동승자들에 대하여 프라이버시를 유지하면서 전화를 걸거나 또는 동승자들 사이의 대화 중단을 방지하고 싶을 수 있다. 그러한 경우, 그 사람은 전화를 걸고 조용한 목소리로 말하거나 다른 사람이 없는 데서 통화할 수 있는 나중까지 통화를 늦출 수 있다. 또 다른 예에서, 그리고 제한 없이, 그룹 미팅에서의 주된 대화는 그룹 미팅 참여자들 중 부분 집합 사이의 부차적인 미팅의 필요를 만들 수 있다. 그러한 경우, 참여자들의 부분 집합은 다른 미팅 룬이 이용가능한 경우, 다른 미팅 룬으로 자리를 옮길 수 있거나 나중까지 부차적인 미팅을 늦출 수 있다.

[0006] 이들 접근법에 하나의 잠재적인 문제는 중요한 또는 필요한 대화가 나중까지 불리하게 미뤄질 수 있다는 것, 또는 주된 대화가 부차적인 대화에 의해 방해받을 수 있다는 것이다. 이들 접근법에 다른 잠재적인 문제는 부차적인 대화가 원하는 정도의 프라이버시를 누릴 수 없거나 또는 작은 목소리로 이루어져, 대화를 참여자들이 이해하기 어렵게 만들 수 있다는 것이다.

[0007] 앞서 말한 내용이 예시한 바와 같이, 동시에 다중 대화를 수용할 새로운 기술이 유용할 수 있다.

발명의 내용

[0008] 하나 이상의 실시예는 무선 네트워크 인터페이스 및 프로세서를 포함하는 컴퓨팅 디바이스를 포함하는 것으로 제시된다. 프로세서는 마이크로폰을 통해, 제1 복수의 음성 성분을 포함하는 제1 청각 신호를 수신하도록 구성된다. 프로세서는 제1 복수의 음성 성분에 포함된 제1 음성 성분을 적어도 부분적으로 억제하기 위한 요청을 수신하도록 더 구성된다. 프로세서는 제1 음성 성분이 적어도 부분적으로 억제된 제1 복수의 음성 성분을 포함하

는 제2 청각 신호를 생성하도록 더 구성된다. 프로세서는 제2 청각 신호를 출력용 스피커로 송신하도록 더 구성된다.

[0009] 다른 실시예들은 제한 없이, 프로세싱 유닛이 개시된 상기 방법들의 하나 이상의 양상을 구현하게 하는 명령들을 포함하는 컴퓨터-판독가능한 매체를 포함한다. 다른 실시예들은 제한 없이, 개시된 상기 방법들의 하나 이상의 양상을 구현하기 위한 방법뿐만 아니라 상기 방법들의 하나 이상의 양상을 구현하도록 구성된 컴퓨팅 시스템을 포함한다.

[0010] 본 출원에 설명된 접근법의 적어도 하나의 이점은 그룹에의 참여자들이 각 대화에 대해 적절한 프라이버시를 유지하고 다른 대화들의 중단을 감소 또는 제거하면서 다중 대화에 참여할 수 있다는 것이다. 그 결과, 중요한 대화들이 미뤄지지 않고 각 별도의 대화를 수용할 별도의 물리적 공간을 찾을 필요 없이 다중 대화가 수용된다.

도면의 간단한 설명

[0011] 위에서 제시된 하나 이상의 실시예의 나열된 특징들이 상세하게 이해될 수 있도록, 위에서 간략하게 요약된 하나 이상의 실시예에 대한 보다 구체적인 설명이 특정 구체적인 실시예들을 참조하여 이루어질 수 있고, 그 일부가 첨부된 도면들에 예시된다. 그러나, 본 발명의 범위가 다른 실시예들을 또한 포괄하기 때문에, 첨부된 도면들이 단지 통상적인 실시예들을 예시하며 그에 따라 어떤 방식으로든 그것의 범위를 제한하는 것으로 간주되지 않아야 한다는 것을 주의해야 한다.

도 1은 다양한 실시예에 따라, 하나 이상의 청각 장면을 생성하도록 작동가능한 청각 장면 제어기를 예시한다;

도 2a는 다양한 실시예에 따라, 청각 장면들을 생성하기 위한 청각 장면 제어기들의 피어 투 피어 네트워크를 예시한다;

도 2b는 다양한 실시예에 따라, 청각 장면들을 생성하기 위한 시스템을 예시한다;

도 2c는 다양한 실시예에 따라, 청각 장면들을 생성하기 위한 시스템을 예시한다;

도 2d는 다양한 실시예에 따라, 청각 장면들을 생성하기 위한 시스템을 예시한다;

도 3은 다양한 실시예에 따라, 양 방향으로 분리된 청각 장면을 갖는 대표적인 사용 사례를 예시한다;

도 4는 다양한 실시예에 따라, 단일 방향으로 밖을 향해 분리된 청각 장면을 갖는 대표적인 사용 사례를 예시한다;

도 5는 다양한 실시예에 따라, 단일 방향으로 안쪽을 향해 분리된 청각 장면을 갖는 대표적인 사용 사례를 예시한다;

도 6은 다양한 실시예에 따라, 다수의 사용자의 양 방향으로 분리된 청각 장면을 갖는 대표적인 사용 사례를 예시한다;

도 7은 다양한 실시예에 따라, 다수의 사용자의 다 방향으로 분리된 청각 장면을 갖는 대표적인 사용 사례를 예시한다;

도 8은 다양한 실시예에 따라, 청각 장면 제어기를 초기화하고 다른 청각 장면 제어기들과 통신하도록 구성하기 위한 방법 단계들의 흐름도이다; 그리고

도 9는 다양한 실시예에 따라, 청각 장면 제어기를 통해 청각 장면을 생성하기 위한 방법 단계들의 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0012] 다음 설명에서, 많은 구체적인 세부사항이 특정 구체적인 실시예들의 보다 철저한 이해를 제공하기 위해 제시된다. 그러나, 해당 기술분야의 통상의 기술자에게 다른 실시예들이 이들 구체적인 세부사항 중 하나 이상 없이도 또는 추가 구체적인 세부사항들을 갖고 실시될 수 있다는 것이 명백할 것이다.

[0013] 시스템 개요

[0014] 도 1은 다양한 실시예에 따라, 하나 이상의 청각 장면을 생성하도록 작동가능한 청각 장면 제어기를 예시한다. 도시된 바와 같이, 청각 장면 제어기(100)는 제한 없이, 함께 결합되는, 마이크로폰 모듈(110), 프로세싱 유닛(120), 무선 트랜시버(130), 스피커 모듈(140), 및 메모리 유닛(150)을 포함한다. 메모리 유닛(150)은 본 출원에 더 설명될 바와 같이 다양한 청각 장면 구성을 생성하기 위한 소프트웨어 애플리케이션인 청각 장면 애플리

케이션(152)을 포함한다. 청각 장면 제어기(100)는 임의의 기술적으로 실행가능한 컴퓨팅 디바이스 내에 내장될 수 있다.

[0015] 마이크로폰 모듈(110)은 마이크로폰을 통해 오디오 신호를 수신하고 오디오 신호들을 기계 판독가능한 형태로 변환하도록 구성된 임의의 기술적으로 실행가능한 유형의 디바이스일 수 있다. 마이크로폰 모듈(110)은 아래에서 더 상세하게 설명될 바와 같이, 물리적 환경으로부터 오디오 신호들을 수신하고 프로세싱을 위한 프로세싱 유닛(120)에 의해 추가 프로세싱하기 위해 그것들의 오디오 신호들을 변환하도록 구성된다. 오디오 신호들은 미팅 또는 다른 물리적 공간에의 다양한 참여자로부터 발화된 음성들 뿐만 아니라 환경적 음원들 이л테면 배경 소음, 음악, 거리 소리들을 포함할 수 있다.

[0016] 프로세싱 유닛(120)은 예를 들어, 그리고 제한 없이, 중앙 처리 장치(CPU), 디지털 신호 처리기(DSP) 또는 응용 주문형 집적 회로(ASIC)를 포함하여, 데이터를 프로세싱하도록 그리고 소프트웨어 애플리케이션들을 실행하도록 구성된 임의의 기술적으로 실행가능한 유닛일 수 있다. 입력 디바이스들(125)은 예를 들어, 그리고 제한 없이, 입력을 수신하도록 구성된 디바이스들(이들테면, 제한 없이, 하나 이상의 버튼)을 포함할 수 있다. 프로세싱 유닛(120)에 의해 실행되는 애플리케이션에 관한 특정 기능들 또는 특징들은 입력 디바이스(125)를 작동시킴으로써, 이들테면 버튼을 누름으로써 액세스될 수 있다. 본 출원에 더 설명될 바와 같이, 프로세싱 유닛(120)은 다양한 사용자를 서로 완전히 또는 부분적으로 분리시키기 위한 하나 이상의 오디오 그룹 또는 대화 "버블(bubble)"을 생성하도록 작동가능하다.

[0017] 스피커 모듈(140)은 오디오 신호를 수신하고, 하나 이상의 라우드스피커 또는 스피커 디바이스를 구동할 수 있는 대응하는 신호를 생성하도록 구성된 임의의 기술적으로 실행가능한 유형의 디바이스일 수 있다. 오디오 신호는 마이크로폰 모듈(110)에 의해 수신된 오디오 입력 신호일 수 있거나, 또는 프로세싱 유닛(120)에 의해 생성된 오디오 신호일 수 있다. 프로세싱 유닛(120)으로부터 수신된 오디오 신호는 마이크로폰 유닛(110)에 의해 수신된, 그러나 하나 이상의 음성이 억제된 오디오 입력 신호의 대체 버전일 수 있다.

[0018] 무선 트랜시버(130)는 제한 없이, WiFi™ 트랜시버, 블루투스 트랜시버, RF 트랜시버 등을 포함하는, 다른 무선 디바이스와의 무선 통신 링크들을 수립하도록 구성된 임의의 기술적으로 실행가능한 디바이스일 수 있다. 무선 트랜시버(130)는 본 출원에 더 설명될 바와 같이, 다른 청각 장면 제어기들 및 중앙 통신 제어기와의 무선 링크들을 수립하도록 구성된다.

[0019] 메모리 유닛(150)은 예를 들어, 그리고 제한 없이, 제한 없이, 랜덤 액세스 메모리(RAM) 모듈 또는 하드 디스크를 포함하여, 데이터 및 프로그램 코드를 저장하도록 구성된 임의의 기술적으로 실행가능한 유닛일 수 있다. 메모리 유닛(150) 내 청각 장면 애플리케이션(152)은 프로세싱 유닛(120)에 의해 하나 이상의 청취 환경(본 출원에서 청각 장면들로도 지칭됨)을 생성하기 위해 실행될 수 있다. 청각 장면은 특정인에 대응하는 적어도 하나의 음성 성분이 청각 장면 내부의 개인들 또는 청각 장면 외부의 사람들 중 어느 하나에 의해 들리게 억제되는 청취 환경을 나타낸다. 일례로, 그리고 제한 없이, 한 사람을 포함하는 청각 장면은 그 밖에 누구도 그 사람의 음성을 듣지 않도록 생성될 수 있다. 다른 예로, 그리고 제한 없이, 한 사람을 포함하는 청각 장면은 그 밖에 누구의 음성도 들리지 않도록 생성될 수 있다. 다른 예로, 그리고 제한 없이, 한 사람을 포함하는 청각 장면은 그 밖에 누구도 그 사람의 음성을 듣지 않도록, 그리고 동시에 그 사람이 그 밖에 누구의 음성도 들리지 않도록 생성될 수 있다. 또 다른 예로, 임의의 수의 청각 장면이 생성될 수 있으며, 여기서 각 청각 장면은 임의의 수의 사람을 포함하고, 각 청각 장면은 다양한 음성을 억제하며, 각 청각 장면을 떠나게 또는 그에 진입하는 것이 방지된다. 이러한 방법으로, 청각 장면들은 매우 사용자 지정가능하고 설정가능하다. 그에 따라, 본 출원에 설명된 청각 장면들은 단지 대표적인 것이고 본 발명의 범위 내에서 생성될 수 있는 가능한 청각 장면들의 범위를 제한하지 않는다.

[0020] 청각 장면들을 생성할 때, 소프트웨어 애플리케이션(152)은 오디오 입력 신호와 연관된 주파수 및 진폭 데이터를 분석 및 파싱하기 위한 매우 다양한 상이한 오디오 프로세싱 알고리즘을 구현할 수 있다. 그러한 알고리즘들은 하나 이상의 기술에 의해 입력 오디오 신호로부터 하나 이상의 음성을 억제하도록 작동가능하다.

[0021] 일례로, 그리고 제한 없이, 청각 장면 애플리케이션(152)을 실행하는 프로세싱 유닛(120)은 억제될 하나 이상의 음성에 대응하는 오디오 입력 신호의 부분을 결정하고, 하나 이상의 음성에 대응하는 반전 신호를 나타내는 반전 오디오 신호를 생성하며, 반전 신호를 원래 오디오 입력 신호와 혼합할 수 있다. 다른 예로, 그리고 제한 없이, 청각 장면 애플리케이션(152)을 실행하는 프로세싱 유닛(120)은 다른 사용자의 청각 장면 제어기로부터 신호를 디지털 방식으로 수신하며, 여기서 수신된 신호는 예를 들어, 그리고 제한 없이, 대응하는 마이크로폰 모듈에 의해 캡처되는 바와 같은 연관된 사용자의 원래 또는 반전 음성을 나타낸다. 그 다음 프로세싱 유닛(120)

은 수신된 신호를 적절하게 반전시키고, 수신된 신호를 마이크로폰 모듈(110)로부터의 오디오 입력 신호와 혼합할 수 있다. 또 다른 예로, 그리고 제한 없이, 청각 장면 애플리케이션(152)을 실행하는 프로세싱 유닛(120)은 다른 사용자의 청각 장면 제어기로부터 타이밍 정보를 수신하여, 연관된 사용자가 말하고 있을 때 또는 말을 안 하고 있을 때를 식별할 수 있다. 프로세싱 유닛(120)은 프로세싱 유닛(120) 유닛이 마이크로폰 모듈(110)로부터 오디오 입력 신호를 억제하는 시간 간격들을 결정하기 위해 수신된 타이밍 정보를 프로세싱한다. 청각 장면 애플리케이션(152)은 그 다음 프로세싱된 오디오 신호를 스피커 모듈(140)로 송신하도록 구성된다.

[0022] 해당 기술분야의 통상의 기술자들은 도 1에 도시된 청각 장면 제어기(100)의 구체적인 구현예가 단지 예시적인 목적들로 제공되고, 본 발명의 범위를 제한하는 것으로 여겨지지 않는다는 것을 이해할 것이다. 실제로, 청각 장면 제어기(100)는 하드웨어 및 소프트웨어의 매우 다양한 상이한 조합에 의해 구현될 수 있다. 예를 들어, 그리고 제한 없이, 청각 장면 제어기(100)는 제한 없이, 위에서 설명된 기능을 수행하도록 구성된 집적 회로에 의해 구현될 수 있다. 다른 예로, 그리고 제한 없이, 청각 장면 제어기(100)는 제한 없이, 해당 기능을 수행하도록 구성된 시스템-온-칩에 의해 구현될 수 있다. 일반적 사실로서, 본 출원에 설명된 청각 장면 제어기(100)의 기능을 수행하도록 구성된 임의의 디바이스는 본 발명의 범위 내에 들어간다. 유사하게, 청각 장면 제어기(100)는 입력 오디오 신호로부터 하나 이상의 음성을 제거하기 위한 임의의 기술적으로 실행가능한 접근법을 수행하도록 구성될 수 있다.

[0023] 도 2a는 다양한 실시예에 따라, 청각 장면들을 생성하기 위한 청각 장면 제어기들(220, 222 및 224)의 피어 투 피어 네트워크(200)를 예시한다. 도시된 바와 같이, 피어 투 피어 네트워크는 제한 없이, 각각 사용자들(210, 212 및 214)에 대응하는 청각 장면 제어기들(220, 222, 및 224)을 포함한다. 청각 장면 제어기들(220, 222 및 224)은 네트워크(130)를 통해 통신한다.

[0024] 이러한 구성에서, 청각 장면 제어기들(220, 222 및 224)은 중앙 통신 제어기 없이 피어 투 피어 방식으로 서로 직접 통신한다. 따라서, 사용자(210)의 액션, 이를테면 버튼 누름에 응답하여, 청각 장면 제어기(220)는 사용자(210)의 음성을 억제하기 위한 요청을 청각 장면 제어기(222 및 224)로 송신한다. 이에 응답하여, 청각 장면 제어기들(222 및 224)은 사용자들(212 및 214)이 사용자(210)를 들을 수 없도록 사용자(210)의 음성을 억제한다. 사용자(210)의 제2 액션, 이를테면 다른 버튼 누름에 응답하여, 청각 장면 제어기(220)는 사용자(210)의 음성을 억제하는 것을 중단하기 위한 요청을 청각 장면 제어기들(222 및 224)로 송신한다. 이에 응답하여, 청각 장면 제어기들(222 및 224)은 사용자들(212 및 214)이 사용자(210)를 다시 들을 수 있도록 사용자(210)의 음성을 억제하는 것을 중단한다.

[0025] 도 2b는 다양한 실시예에 따라, 청각 장면들을 생성하기 위한 시스템(202)을 예시한다. 도시된 바와 같이, 시스템은 제한 없이, 중앙 통신 제어기(240) 및 각각 사용자들(210, 212 및 214)에 대응하는 청각 장면 제어기들(220, 222, 및 224)을 포함한다. 청각 장면 제어기들(220, 222 및 224)은 네트워크(130)를 통해 통신한다. 청각 장면 제어기들(220, 222 및 224) 및 네트워크(230)는 아래에 더 설명될 바를 제외하고는 도 2a와 함께 설명된 바와 실질적으로 동일한 기능을 한다.

[0026] 이러한 구성에서, 청각 장면 제어기들(220, 222 및 224)은 중앙 통신 제어기(240)를 통해 서로 통신한다. 중앙 통신 제어기(240)는 임의의 기술적으로 실행가능한 컴퓨팅 디바이스 내에 내장될 수 있다. 각 청각 장면 제어기들(220, 222 및 224) 중앙 통신 제어기(240). 적절하게, 중앙 통신 제어기(240)는 청각 장면 제어기들(220, 222 및 224)로부터 수신된 정보들을 다른 청각 장면 제어기들(220, 222 및 224)로 포워드한다. 덧붙여, 중앙 통신 제어기(240)는 청각 장면 제어기들(220, 222 및 224)로 지향되는 통신을 개시할 수 있다.

[0027] 따라서, 사용자(210)의 액션, 이를테면 버튼 누름에 응답하여, 청각 장면 제어기(220)는 사용자(210)의 음성을 억제하기 위한 요청을 통신 제어기(240)로 송신한다. 이에 응답하여, 통신 제어기(240)는 청각 장면 제어기들(222 및 224)로 요청을 포워드한다. 청각 장면 제어기들(222 및 224)은 사용자들(212 및 214)이 사용자(210)를 들을 수 없도록 사용자(210)의 음성을 억제한다. 사용자(210)의 제2 액션, 이를테면 다른 버튼 누름에 응답하여, 청각 장면 제어기(220)는 사용자(210)의 음성을 억제하는 것을 중단하기 위한 요청을 통신 제어기(240)로 송신한다. 이에 응답하여, 통신 제어기(240)는 청각 장면 제어기들(222 및 224)로 요청을 포워드한다. 청각 장면 제어기들(222 및 224)은 사용자들(212 및 214)이 사용자(210)를 다시 들을 수 있도록 사용자(210)의 음성을 억제하는 것을 중단한다.

[0028] 도 2c는 다양한 실시예에 따라, 청각 장면들을 생성하기 위한 시스템(204)을 예시한다. 도시된 바와 같이, 시스템은 제한 없이, 중앙 통신 제어기(240) 및 각각 사용자들(210, 212 및 214)에 대응하는 청각 장면 제어기들(220, 222, 및 224)을 포함한다. 또한 도시된 바와 같이, 시스템은 각각 사용자들(210, 212 및 214)에 대응하는

사용자 인터페이스 디바이스들(250, 252 및 254)을 포함한다. 청각 장면 제어기들(220, 222 및 224)은 네트워크(130)를 통해 통신한다. 청각 장면 제어기들(220, 222 및 224), 중앙 통신 제어기(240) 및 네트워크(230)는 아래에 더 설명될 바를 제외하고는 도 2a 및 도 2b와 함께 설명된 바와 실질적으로 동일한 기능을 한다.

[0029] 도시된 바와 같이, 사용자 인터페이스 디바이스(250)는 사용자(210)와 연관된 스마트폰이고, 사용자 인터페이스 디바이스(252)는 사용자(210)와 연관된 랩탑 컴퓨터이며, 사용자 인터페이스 디바이스(254)는 사용자(210)와 연관된 태블릿 컴퓨터이다. 대안적으로, 다양한 사용자는 제한 없이, 착용자의 안경에 부착되는, 착용자의 목걸이 또는 "애물렛 디바이스(amulet device)"에 부착되는, 손목 시계 또는 팔목 팔찌 상의, 헤드 밴드 또는 헤드 링에 내장되는, 의류 또는 벨트 버클에 부착되는, 사용자의 몸 어디든지에 부착 또는 착용되는 디바이스, 사용자의 스마트폰 또는 태블릿 컴퓨터에 부착되는 액세서리, 및 사용자와 연관된 운송 수단, 이륜차 또는 오토바이에 부착되는 것을 포함하여 임의의 조합으로 임의의 기술적으로 실행가능한 사용자 인터페이스 디바이스들과 연관될 수 있다.

[0030] 도 2c의 구성에서, 청각 장면 제어기들(220, 222 및 224)은 중앙 통신 제어기(240)를 통해 서로 통신한다. 각 청각 장면 제어기들(220, 222 및 224) 중앙 통신 제어기(240). 적절하게, 중앙 통신 제어기(240)는 청각 장면 제어기들(220, 222 및 224)로부터 수신된 정보들을 다른 청각 장면 제어기들(220, 222 및 224)로 포워딩한다. 덧붙여, 중앙 통신 제어기(240)는 청각 장면 제어기들(220, 222 및 224)로 지향되는 통신을 개시할 수 있다. 중앙 통신 제어기(240)는 또한 사용자 인터페이스 디바이스들(250, 252 및 254) 사이에서 정보들을 송수신한다. 사용자 인터페이스 디바이스들(250, 252 및 254) 상에서 실행하는 애플리케이션은 청각 장면 제어기들(220, 222 및 224)의 보다 간단한 사용자 인터페이스를 이용하여 가능한 것보다 복잡한 청각 장면들을 생성할 수 있을 수 있다.

[0031] 따라서, 사용자(210)의 액션, 이륜차면 사용자 인터페이스 디바이스(250) 상에서 실행하는 애플리케이션 상의 기능을 선택 하는 것에 응답하여, 사용자 인터페이스 디바이스(250)는 사용자(210)의 음성을 억제하기 위한 요청을 통신 제어기(240)로 송신한다. 이에 응답하여, 통신 제어기(240)는 청각 장면 제어기들(222 및 224)로 요청을 포워딩한다. 청각 장면 제어기들(222 및 224)은 사용자들(212 및 214)이 사용자(210)를 들을 수 없도록 사용자(210)의 음성을 억제한다. 사용자(210)의 제2 액션, 이륜차면 사용자 인터페이스 디바이스(250) 상에서 실행하는 애플리케이션 상의 기능을 선택 하는 것에 응답하여, 사용자 인터페이스 디바이스(250)는 사용자(210)의 음성을 억제하는 것을 중단하기 위한 요청을 통신 제어기(240)로 송신한다. 이에 응답하여, 통신 제어기(240)는 청각 장면 제어기들(222 및 224)로 요청을 포워딩한다. 청각 장면 제어기들(222 및 224)은 사용자들(212 및 214)이 사용자(210)를 다시 들을 수 있도록 사용자(210)의 음성을 억제하는 것을 중단한다.

[0032] 도 2d는 다양한 실시예에 따라, 청각 장면들을 생성하기 위한 시스템(206)을 예시한다. 도시된 바와 같이, 시스템은 제한 없이, 중앙 통신 제어기(240) 및 고지향성 라우드스피커들(HDL들)(260(0) 내지 260(13))의 그룹을 포함한다. 중앙 통신 제어기(240)는 네트워크(130)를 통해 HDL들(260(0) 내지 260(13))과 통신한다. 대안적으로, 중앙 통신 제어기(240)는 HDL들(260(0) 내지 260(13))에 직접 연결될 수 있다. 중앙 통신 제어기(240) 및 네트워크(230)는 아래에 더 설명될 바를 제외하고는 도 2a 내지 도 2c와 함께 설명된 바와 실질적으로 동일한 기능을 한다.

[0033] HDL들(260)은 종래 라우드스피커들에 의해 생성된 보다 통상적인 무지향성 음파 패턴이 아니라, 비교적 고도의 지향성(협소)을 갖는 음파 패턴들을 생성하는 라우드스피커들이다. 따라서, 소정의 HDL(260)은 소리를 특정 청취자에 지향시킬 수 있고, 그에 따라 청취자는 HDL(260)에 의해 생성된 소리를 듣게 되나, 그 청취자의 바로 왼쪽 또는 바로 오른쪽에 앉은 다른 사람은 HDL(260)에 의해 생성된 소리를 듣지 않게 된다. 예를 들어, 그리고 제한 없이, HDL(260(1)) 및 HDL(260(2))은 소리를 사용자(210)의 각각 오른쪽 귀 및 왼쪽 귀에 지향시키도록 구성될 수 있다. HDL(260(5)) 및 HDL(260(6))은 소리를 사용자(212)의 각각 오른쪽 귀 및 왼쪽 귀에 지향시키도록 구성될 수 있다. HDL(260(10)) 및 HDL(260(11))은 소리를 사용자(214)의 각각 오른쪽 귀 및 왼쪽 귀에 지향시키도록 구성될 수 있다. 14개의 HDL(260(0) 내지 260(13))이 도시되지만, 본 발명의 범위 내에서 임의의 기술적으로 실행가능한 수의 사용자(210, 212 및 214)를 수용하기 위해, 임의의 기술적으로 실행가능한 수의 HDL(260)이 채용될 수 있다.

[0034] 도 2a 내지 도 2d의 다양한 구성요소는 본 발명의 범위 내에서 임의의 조합으로 사용될 수 있다. 일례로, 그리고 제한 없이, 사용자들은 도 2c에 도시된 바와 같이, 청각 장면들을 구성하기 위한 사용자 인터페이스 디바이스들을 갖고, 도 2d에 도시된 바와 같이, 고지향성 라우드스피커들로부터 오디오 신호들을 수신할 수 있다. 다른 예로, 그리고 제한 없이, 청각 장면 제어기들 및 사용자 인터페이스 디바이스들은 도 2c에 도시된 바와

같이, 중앙 통신 제어기 필요 없이, 도 2a에 도시된 바와 같이, 피어 투 피어 네트워크에서 서로 직접 통신할 수 있다.

[0035] 도시된 바와 같이, 청각 장면 제어기(100)의 기능은 사용자가 착용 또는 휴대할 수 있는 웨어러블 디바이스로 통합될 수 있다. 일 실시예에서, 청각 장면 제어기(100)는 사용자가 착용한 인-이어 디바이스로 통합될 수 있다. 대안적인 실시예들에서, 청각 장면 제어기(100)의 기능은 예를 들어 그리고 제한 없이, 블루투스 헤드셋, 어깨에 착용한 스피커들, 헤드폰들, 이어 버드들, 보청기들, 인-이어 모니터들, 헤드레스트에 내장된 스피커들, 또는 동일한 효과 또는 기능을 갖는 임의의 다른 디바이스를 포함하여, 마이크로폰 및 스피커 중 적어도 하나를 포함하는 헤드-장착 청각 디바이스로 통합될 수 있다. 청각 장면 제어기(100)는 제한 없이, 스마트폰, 컴퓨터 및 태블릿 컴퓨터를 포함하여, 청각 장면들을 구성하기 위한 사용자 인터페이스를 포함하는 디바이스에 결합될 수 있다. 청각 장면 제어기(100)는 제한 없이, 무선 링크, 하드웨어에 내장된 연결 및 네트워크 연결을 포함하여, 임의의 기술적으로 실행가능한 접근법을 통해 그러한 디바이스에 결합될 수 있다. 무선 링크들은 제한 없이, WiFi™ 링크, 블루투스 연결, 또는 포괄적인 라디오 주파수(RF) 연결을 포함하여, 임의의 기술적으로 실행가능한 무선 통신 링크를 통해 구성될 수 있다. 실제로, 청각 장면 제어기(100)는 예시된 것들을 너머 광범위한 상이한 무선 디바이스와의 통신 링크를 수립할 수 있다. 도 2c에 예시된 특정 디바이스들(250, 252 및 254)은 제한하는 것으로 여겨지는 것이 아니라 단지 대표적인 목적들로 도시된다.

[0036] **청각 장면 제어기의 작동을 예시하는 대표적인 사용 사례들**

[0037] 도 3은 다양한 실시예에 따라, 양 방향으로 분리된 청각 장면들을 갖는 대표적인 사용 사례(300)를 예시한다. 도시된 바와 같이, 사용 사례는 사용자들(310, 312 및 314) 및 양 방향으로 분리된 대화 버블(320)을 포함한다.

[0038] 도 3의 구성에서, 사용자(314)는 사용자들(310 및 312)에 들리지 않는 것 그리고 사용자들(310 및 312)의 음성들을 듣지 않는 것을 선택한다. 일례로서, 그리고 제한 없이, 사용자(314)는 사용자들(310 및 312)을 산만하게 하지 않고, 또는 사용자들(310 및 312)에 의해 산만하게 되지 않고 개인적인 전화를 걸기 위해 이러한 구성을 선택할 수 있다. 일례로, 그리고 제한 없이, 이러한 구성은 사용자(314)가 미팅 중에 있거나 버스 또는 택시에 타고 있을 때 전화 모바일 통화를 걸거나 받기 원할 때 생성될 수 있다. 그러한 경우, 사용자(314)와 연관된 청각 장면 제어기(100)는 사용자들(310 및 312)의 음성 성분들을 억제하도록 인입 오디오 신호들을 프로세싱한다. 사용자(314)와 연관된 청각 장면 제어기(100)는 그것들의 각각의 입력 오디오 신호들로부터 사용자(314)의 음성 성분을 억제하기 위한 요청을 사용자들(310 및 312)과 연관된 청각 장면 제어기들(100)로 전송한다. 그렇게 함으로써 양 방향으로 분리된 대화 버블(320)이 생성되어 하나는 사용자(314)를 포함하고 다른 하나는 사용자들(310 및 312)을 포함하는, 두 개의 청각 장면을 야기한다.

[0039] 도 4는 다양한 실시예에 따라, 단일 방향으로 밖을 향해 분리된 청각 장면들을 갖는 대표적인 사용 사례(400)를 예시한다. 도시된 바와 같이, 사용 사례(400)는 사용자들(410, 412 및 414) 및 단일 방향으로 밖을 향해 분리된 대화 버블(420)을 포함한다.

[0040] 도 4의 구성에서, 사용자(414)는 사용자들(410 및 412)에 들리지 않는 것, 그러나 사용자들(410 및 412)의 음성들을 듣는 것을 선택한다. 일례로, 그리고 제한 없이, 사용자(414)는, 이를테면 사용자(410)가 미팅 중에 있거나 버스 또는 택시에 타고 있을 때, 사용자들(410 및 412)을 산만하게 하지 않고 개인적인 전화를 걸기 위해 이러한 구성을 선택할 수 있으나, 계속해서 사용자들(410 및 412) 간에 발생하는 대화를 듣기 원할 수 있다. 그러한 경우, 사용자(414)와 연관된 청각 장면 제어기(100)는 그것들의 각각의 입력 오디오 신호들로부터 사용자(414)의 음성 성분을 억제하기 위한 요청을 사용자들(410 및 412)과 연관된 청각 장면 제어기들(100)로 전송한다. 사용자(414)와 연관된 청각 장면 제어기(100)는 사용자(414)의 기본 설정에 따라, 가장 높은 음량 또는 감음된 음량 중 어느 하나로 사용자들(410 및 412)의 음성 성분들을 억제하도록 인입 오디오 신호들을 프로세싱한다. 그렇게 함으로써 단일 방향으로 밖을 향해 분리된 대화 버블(420)이 생성되어 하나는 사용자(414)를 포함하고 다른 하나는 사용자들(410 및 412)을 포함하는, 두 개의 청각 장면을 야기한다.

[0041] 도 5는 다양한 실시예에 따라, 단일 방향으로 안쪽을 향해 분리된 청각 장면들을 갖는 대표적인 사용 사례(500)를 예시한다. 도시된 바와 같이, 사용 사례(500)는 사용자들(510, 512 및 514) 및 단일 방향으로 안을 향해 분리된 대화 버블(520)을 포함한다.

[0042] 도 5의 구성에서, 사용자(514)는 사용자들(510 및 512)에 들리는 것, 그러나 사용자들(510 및 512)의 음성들을 듣지 않는 것을 선택한다. 일례로, 그리고 제한 없이, 사용자(514)는 사용자들(510 및 512) 간 대화로부터의 주의산만을 제거하기 위해 이러한 구성을 선택할 수 있으나 사용자들(510 및 512)이 들을 수 있을 발언에 끼기 원

할 수 있다. 다른 예로, 그리고 제한 없이, 사용자(514)는 일시적으로 주의산만하지 않고 이메일에 회신하거나 다른 문제들을 처리하는 것에 집중하기 위해 이러한 구성을 선택할 수 있고 사용자들(510 및 512)이 대화를 나누고 있는 장소를 벗어나기 원치 않는다. 그러한 경우, 사용자(514)와 연관된 청각 장면 제어기(100)는 그것들의 각각의 입력 오디오 신호들을 갖는 사용자(514)의 음성 성분을 전달하기 위한 요청을 사용자들(510 및 512)과 연관된 청각 장면 제어기들(100)로 전송한다. 사용자(514)와 연관된 청각 장면 제어기(100)는 사용자(514)의 기본 설정에 따라, 가장 높은 음량 또는 감음된 음량 중 어느 하나로 사용자들(510 및 512)의 음성 성분들을 억제하도록 인입 오디오 신호들을 프로세싱한다. 그렇게 함으로써 단일 방향으로 안을 향해 분리된 대화 버블(520)이 생성되어 하나는 사용자(514)를 포함하고 다른 하나는 사용자들(510 및 512)을 포함하는, 두 개의 청각 장면을 야기한다.

[0043] 도 6은 다양한 실시예에 따라, 다수의 사용자의 양 방향으로 분리된 청각 장면을 갖는 대표적인 사용 사례(600)를 예시한다. 도시된 바와 같이, 사용 사례(600)는 사용자들(610, 612 및 614) 및 다수의 사용자를 갖는 양 방향으로 분리된 대화 버블(620)을 포함한다.

[0044] 도 6의 구성에서, 사용자들(610 및 614)은 사용자(612)에 들리지 않는 것 그리고 사용자(612)의 음성을 듣지 않는 것을 선택한다. 일례로, 그리고 제한 없이, 사용자들(610 및 614)은 사용자(612)의 청취 외에서 개인적인 대화를 나누기 위해 이러한 구성을 선택할 수 있다. 사용자들(610 및 614)은 사용자(612)를 산만하게 하지 않고 도서관 또는 커피숍에서 개인적인 대화를 나누기 위해 이러한 구성을 선택할 수 있다. 그러한 경우, 사용자들(610 및 614)과 연관된 청각 장면 제어기들(100) 중 하나 또는 양자는 입력 오디오 신호를 갖는 사용자들(610 및 614)의 음성 성분을 억제하기 위한 요청을 사용자(612)와 연관된 청각 장면 제어기(100)로 전송한다. 사용자들(610 및 614)과 연관된 청각 장면 제어기들(100)은 사용자(614)의 기본 설정에 따라, 가장 높은 음량 또는 감음된 음량 중 어느 하나로 사용자들(612)의 음성 성분들을 완전히 또는 부분적으로 억제하도록 인입 오디오 신호들을 프로세싱한다. 사용자들(610 및 614)과 연관된 청각 장면 제어기들(100)은 이를테면 사용자들(610 및 614)이 잡음 환경에서 대화를 나누고 있을 때, 임의로 배경 잡음을 억제하는 것을 선택할 수 있다. 그렇게 함으로써 양 방향으로 분리된 대화 버블(620)이 생성되어 하나는 사용자(612)를 포함하고 다른 하나는 사용자들(610 및 614)을 포함하는, 두 개의 청각 장면을 야기한다.

[0045] 도 7은 다양한 실시예에 따라, 다수의 사용자의 다 방향으로 분리된 청각 장면을 갖는 대표적인 사용 사례(700)를 예시한다. 도시된 바와 같이, 사용 사례는 사용자들(710, 712, 714 및 716) 및 다 방향으로 분리된 대화 버블(720, 722 및 724)을 포함한다.

[0046] 도 7의 구성에서, 사용자들(710 및 716)이 서로 대화하기 원할 수 있는 한편, 사용자들(712 및 714)이 서로 대화하기 원할 수 있다. 덧붙여, 사용자(712)는 사용자(710)의 음성을 듣기 원할 수 있다. 일례로서, 그리고 제한 없이, 사용자(710)가 하나의 언어로 연설하고 있는 동안, 사용자(712)는 연설을 제2 언어로 번역하고 있는 상황들에 대한 사용자들(710, 712, 714 및 716). 사용자(716)는 710에 의해 발화된 언어로 연설을 들으나, 사용자들(712 또는 714)의 음성들을 듣지 않는다. 사용자(714)는 사용자(712)의 음성을 들으나, 사용자(710)의 음성은 사용자(714)의 기본 설정시 사용자(714)에 대해 완전히 또는 부분적으로 억제된다. 그러한 경우, 사용자들(710, 712, 714 및 716)과 연관된 청각 장면 제어기들(100)은 적절한 음성 성분들을 억제하기 위한 요청들을 서로에 전송한다. 사용자들(710, 712, 714 및 716)과 연관된 청각 장면 제어기들(100)은 적절하게, 가장 높은 음량 또는 감음된 음량 중 어느 하나로 다양한 사용자의 음성 성분들을 완전히 또는 부분적으로 억제하도록 인입 오디오 신호들을 프로세싱한다. 그렇게 함으로써 다 방향으로 분리된 대화 버블(720, 722 및 724)이 생성되어 하나는 사용자들(710 및 716)을 포함하고, 다른 하나는 사용자들(710 및 712)을 포함하며, 다른 하나는 사용자들(712 및 714)을 포함하는, 세 개의 청각 장면을 야기한다.

[0047] 해당 기술분야의 통상의 기술자들은 도 3 내지 도 7과 함께 위에서 설명된 대표적인 사용 사례 시나리오들은 단지 다양한 청각 장면 구성을 생성하기 위해 청각 장면 제어기(100)가 구현할 수 있는 상이한 기술들을 예시하기 위한 예시적인 목적들로 제공된다는 것이 이해될 것이다. 각 청각 장면이 임의의 수의 사용자를 포함하는, 임의의 수의 청각 장면의 많은 기타 구성이 본 발명의 범위 내에서 설명된 기술들을 사용하여 구현될 수 있다. 나아가, 위에서 논의된 예들은 특정 명령들, 디바이스들, 및 동작들을 참조하여 제시되었지만, 본 발명의 범위를 그것들의 특이성들로 제한하는 것으로 여겨지지 않는다.

[0048] 지금까지 청각 장면들의 다양한 구성을 생성하기 위한 다양한 사용 사례 및 시스템이 설명되었고, 이제 청각 장면 제어기(100)에 의해 구현될 수 있는 대표적인 알고리즘들이 설명된다. 지금까지 설명된 기능을 구현함으로써, 청각 장면 제어기(100)는 동일한 공간에서 서로 방해하지 않고 다양한 대화를 동시에 수행할 수

있는 개인들의 능력을 향상시킬 수 있다.

- [0049] **청각 장면 제어기에 의해 구현되는 대표적인 알고리즘들**
- [0050] 도 8은 다양한 실시예에 따라, 청각 장면 제어기를 초기화하고 다른 청각 장면 제어기들과 통신하도록 구성하기 위한 방법 단계들의 흐름도이다. 방법 단계들이 도 1 내지 도 7의 시스템과 함께 설명되지만, 해당 기술분야의 통상의 기술자들은 방법 단계들을 수행하도록 구성된 임의의 시스템이 임의의 순서로 본 발명의 범위 내에 있다는 것을 이해할 것이다.
- [0051] 도시된 바와 같이, 방법(800)은 청각 장면 제어기(100)가 제한 없이, 다른 청각 장면 제어기들 및 중앙 통신 제어기를 포함하여, 가까운 무선 디바이스들을 발견하는 단계(802)로 시작된다. 청각 장면 제어기(100)는 제한 없이, WiFi™ 액세스 포인트의 위치를 찾고 그 다음 그것에 결합되는 다른 디바이스들을 식별하는 것, 가까운 블루투스 디바이스들과 직접 상호작용하는 것, 또는 RF 신호들을 사용하여 무선 디바이스들과 통칭 핸드셰이킹을 수행하는 것을 포함하여, 임의의 기술적으로 실행가능한 형태의 디바이스 발견을 수행할 수 있다.
- [0052] 단계(804)에서, 청각 장면 제어기(100)는 각 발견된 디바이스로부터, 다른 것들 중에서도, 디바이스 성능들을 반영하는 디바이스 정보를 획득한다. 성능들은 예를 들어, 그리고 제한 없이, 기본 설정 무선 연결 프로토콜(예를 들어, 제한 없이, WiFi™, 블루투스), 디바이스에 의해 지원되는 청각 장면들의 최대 수 등을 포함할 수 있다. 다른 디바이스 정보는 예를 들어, 그리고 제한 없이, 디바이스 위치, 디바이스 배터리 레벨 들을 포함할 수 있다.
- [0053] 단계(806)에서, 청각 장면 제어기(100)는 발견된 디바이스들 중 하나 이상과 페어링한다. 그렇게 하면서, 청각 장면 제어기(100)는 임의의 관련 프로토콜에 의존할 수 있다. 덧붙여, 청각 장면 제어기(100)는 상이한 프로토콜들을 의존하는 상이한 디바이스들과 페어링할 수 있다.
- [0054] 단계(808)에서, 청각 장면 제어기(100)는 요구될 때, 페어링된 디바이스들에 대한 명령 라우팅 기본 설정들을 구성한다. 그렇게 하면서, 청각 장면 제어기(100)는 피어 투 피어 네트워크에서 다른 청각 장면 제어기들과 직접 통신할 수 있다. 대안적으로, 청각 장면 제어기(100)는 다른 청각 장면 제어기들과 함께, 단지 중앙 통신 제어기(240)와 직접 통신하고, 중앙 통신 제어기(240)는 청각 장면 제어기들의 각각과 별도로 통신한다. 아래에 논의될 도 9는 청각 장면 제어기(100)의 정규 동작을 단계적인 방식으로 설명한다.
- [0055] 도 9는 다양한 실시예에 따라, 청각 장면 제어기를 통해 청각 장면을 생성하기 위한 방법 단계들의 흐름도이다. 방법 단계들이 도 1 내지 도 7의 시스템과 함께 설명되지만, 해당 기술분야의 통상의 기술자들은 방법 단계들을 수행하도록 구성된 임의의 시스템이 임의의 순서로 본 발명의 범위 내에 있다는 것을 이해할 것이다.
- [0056] 도시된 바와 같이, 방법(900)은 청각 장면 제어기(100)가 마이크로폰 모듈(110)로부터 수신되는 오디오 입력 신호가 변경 없이 스피커 모듈(140)로 송신되는 상태로 초기화하는 단계(902)로 시작된다. 단계(904)에서, 청각 장면 제어기(100)는 예를 들어, 그리고 제한 없이, 다른 오디오 장면 제어기로부터 또는 통신 모듈(140)로부터 요청을 수신함으로써, 청각 장면 요청을 수신한다.
- [0057] 단계(906)에서, 청각 장면 제어기(100)는 오디오 장면 요청이 오디오 음성 성분, 이를테면 다른 청각 장면 제어기와 연관된 음성을 억제하기 위한 요청이었는지 여부를 결정한다. 오디오 장면 요청이 음성 억제 요청인 경우, 방법(900)은 청각 장면 제어기(100)가 요청된 음성 성분이 억제된 수신된 오디오 입력 신호를 포함하는 오디오 신호를 생성하는 단계(908)로 진행한다. 단계(910)에서, 청각 장면 제어기(100)는 생성된 오디오 신호를 라우드스피커 모듈(140)로 송신한다. 그 다음 방법(900)은 위에서 설명된 단계(904)로 진행한다.
- [0058] 단계(906)에서, 오디오 장면 요청이 음성 억제 요청이 아닌 경우, 방법(900)은 청각 장면 제어기(100)가 오디오 장면 요청이 오디오 음성 성분, 이를테면 다른 청각 장면 제어기와 연관된 음성을 억제하는 것을 중단하기 위한 요청이었는지 여부를 결정하는 단계(912)로 진행한다. 오디오 장면 요청이 음성 억제 중지 요청인 경우, 방법(900)은 청각 장면 제어기(100)가 수신된 오디오 입력 신호를 포함하고 다시 그 신호로 혼합되는 요청된 음성 성분을 갖는 오디오 신호를 생성하는 단계(914)로 진행한다. 단계(916)에서, 청각 장면 제어기(100)는 생성된 오디오 신호를 라우드스피커 모듈(140)로 송신한다. 그 다음 방법(900)은 위에서 설명된 단계(904)로 진행한다.
- [0059] 단계(906)에서 오디오 장면 요청이 음성 억제 중지 요청이 아닌 경우, 방법(900)은 위에서 설명된 단계(904)로 진행한다.
- [0060] 요약하면, 청각 장면 제어기는 물리적 환경에서 다수의 청각 장면을 생성하도록 구성된다. 청각 장면 제어기는 인입 오디오 신호에서의 모든 음성을 억제하고 사용자의 음성이 다른 사용자들에 의해 들리는 것을 불허하기 위

해 사용자의 음성을 억제하기 위한 요청을 다른 청각 장면 제어기로 전송함으로써 청각 신호의 사용자를 양 방향으로 분리할 수 있다. 대안적으로, 청각 장면 제어기는 인입 오디오 신호에서의 모든 음성을 억제하나, 사용자의 음성이 다른 사용자들에 의해 들리는 것을 허용함으로써 청각 신호의 사용자를 단일 방향으로 분리할 수 있다. 대안적으로, 청각 장면 제어기는 인입 오디오 신호에서의 모든 음성이 그 사용자에게 의해 들리는 것을 허용하나, 그 사용자의 음성이 다른 사용자들에 의해 들리는 것을 불허하기 위해 그 사용자의 음성을 억제하기 위한 요청을 다른 청각 장면 제어기로 전송함으로써 청각 신호의 사용자를 단일 방향으로 분리할 수 있다. 대화 버블들은 몇몇 사람의 서브 그룹이 서브 그룹에서 서로 대화할 수 있게 하도록, 그러나 주된 그룹에서의 다른 사용자들의 대화에서 분리되도록 생성될 수 있다.

[0061] 본 출원에 설명된 접근법의 적어도 하나의 이점은 그룹에의 참여자들이 각 대화에 대해 적절한 프라이버시를 유지하고 다른 대화들의 중단을 감소 또는 제거하면서 다중 대화에 참여할 수 있다는 것이다. 그 결과, 중요한 대화들이 미뤄지지 않고 각 별도의 대화를 수용할 별도의 물리적 공간을 찾을 필요 없이 다중 대화가 수용된다.

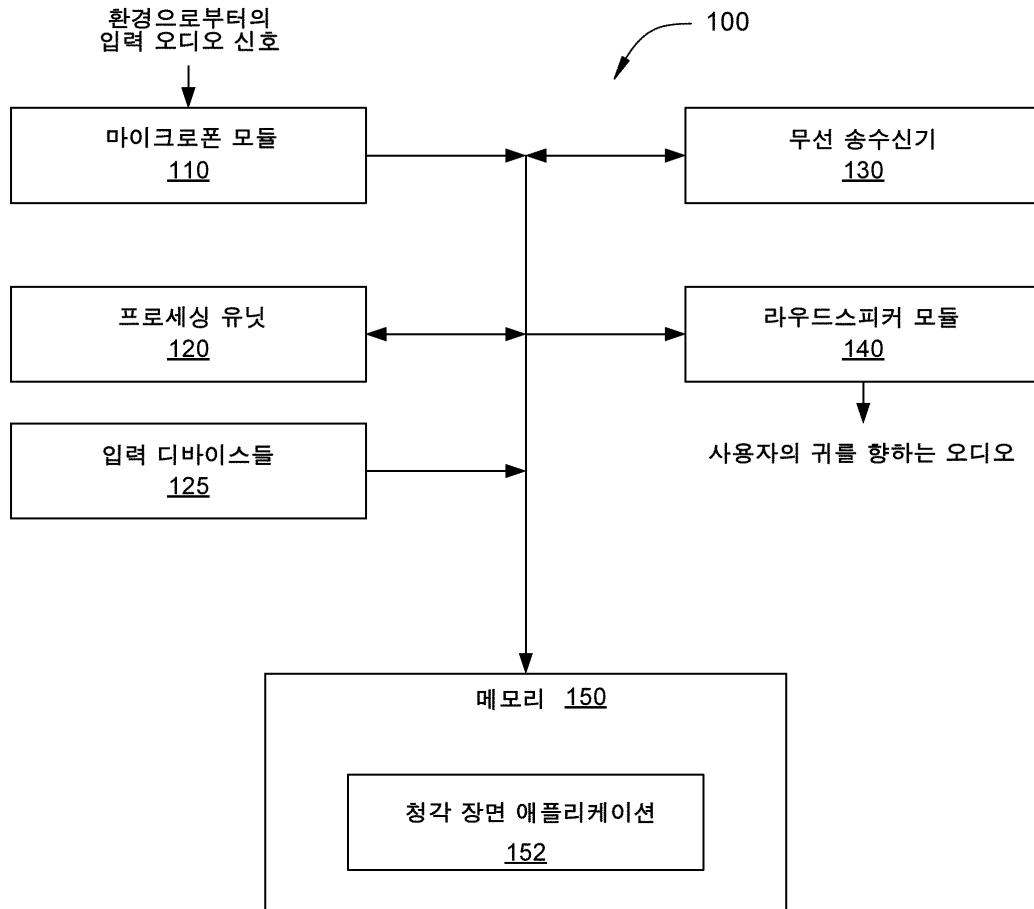
[0062] 본 발명의 일 실시예는 컴퓨터 시스템용 프로그램 제품으로 구현될 수 있다. 프로그램 제품의 프로그램(들)은 실시예들의 기능들(본 출원에 설명된 방법들을 포함하여)을 정의하고 다양한 컴퓨터-판독가능한 저장 매체 상에 포함될 수 있다. 예시적인 컴퓨터-판독가능한 저장 매체는 이에 제한되는 것은 아니나, 다음을 포함한다: (i) 정보가 영구적으로 저장되는 비-기록가능한 저장 미디어(예를 들어, 컴퓨터 내 판독 전용 메모리 디바이스들, 이를테면 CD-ROM 드라이브에 의해 판독가능한 콤팩트 디스크 판독 전용 메모리(CD-ROM) 디스크들, 플래시 메모리, 판독 전용 메모리(ROM) 칩들 또는 임의의 유형의 고체-상태 비-휘발성 반도체 메모리); 및 (ii) 변경가능한 정보가 저장되는 기록가능한 저장 미디어(예를 들어, 디스켓 드라이브 또는 하드 디스크 드라이브 내 플로피 디스크들 또는 임의의 유형의 고체-상태 랜덤-액세스 반도체 메모리).

[0063] 본 발명은 위에서 구체적 실시예들을 참조하여 설명되었다. 그러나, 해당 기술분야의 통상의 기술자들은 다양한 변형 및 변경이 첨부된 청구항들에 제시된 바와 같이 본 발명의 광범위한 사상 및 범위로부터 벗어나지 않고 그것 내에서 이루어질 수 있다는 것을 이해할 것이다. 그에 따라, 앞서 말한 설명 및 도면들은 제한적인 의미가 아니라 예시적인 의미로 간주될 것이다.

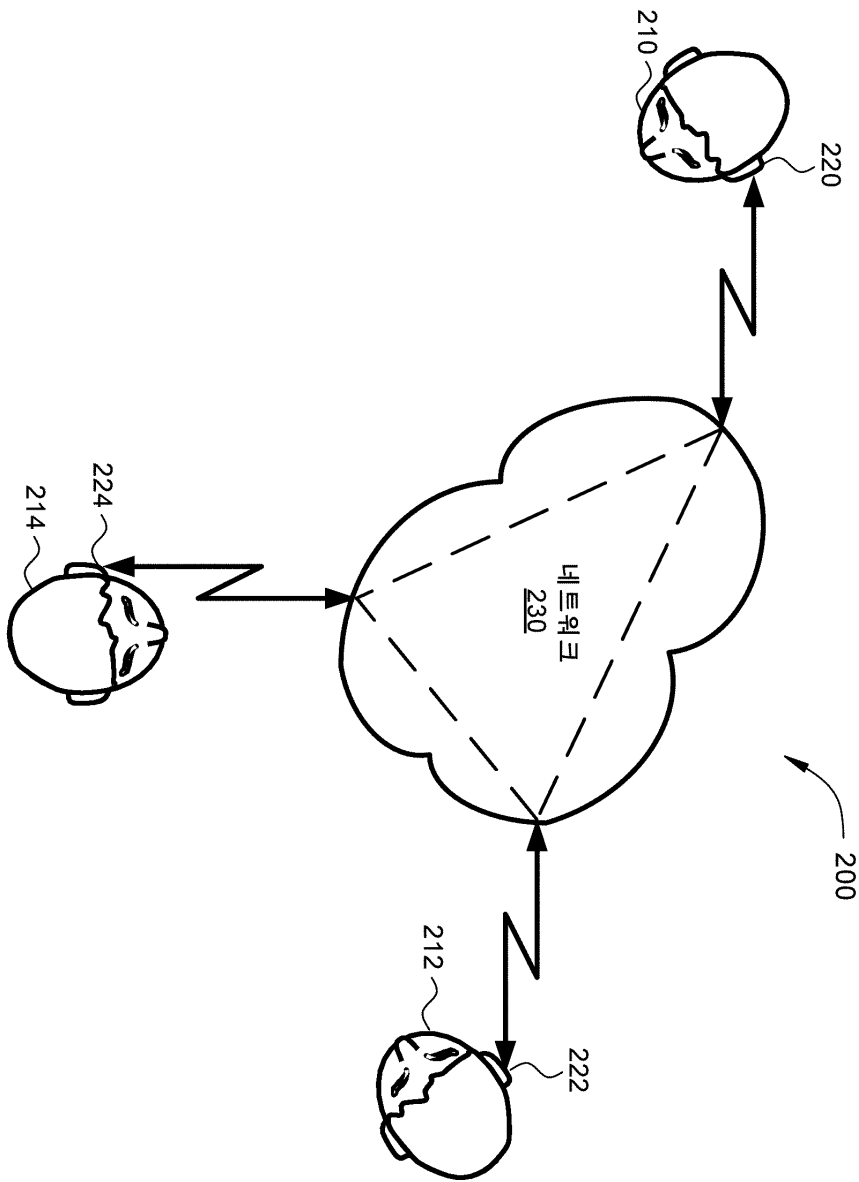
[0064] 따라서, 본 발명의 실시예들의 범위는 뒤따르는 청구항들에 제시된다.

도면

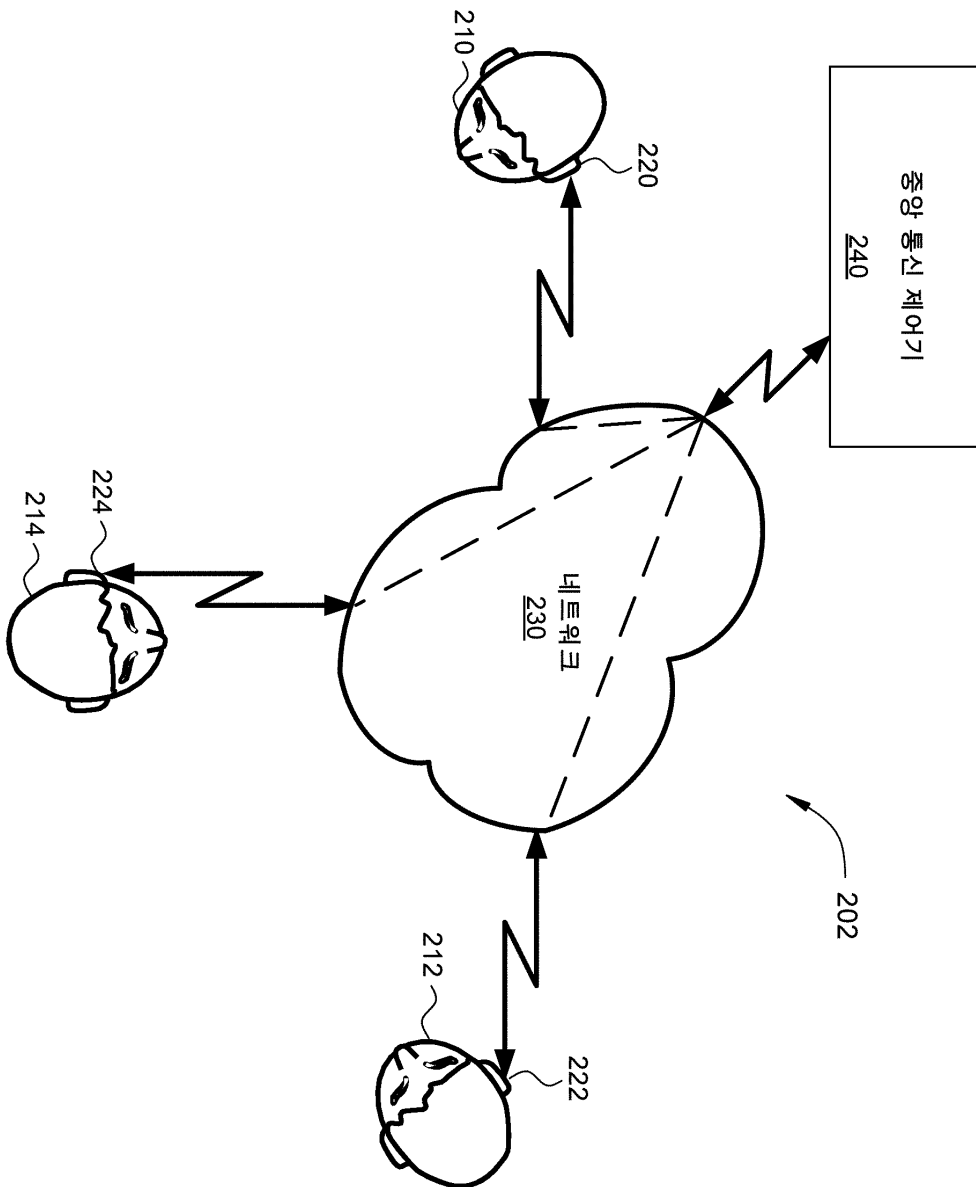
도면1



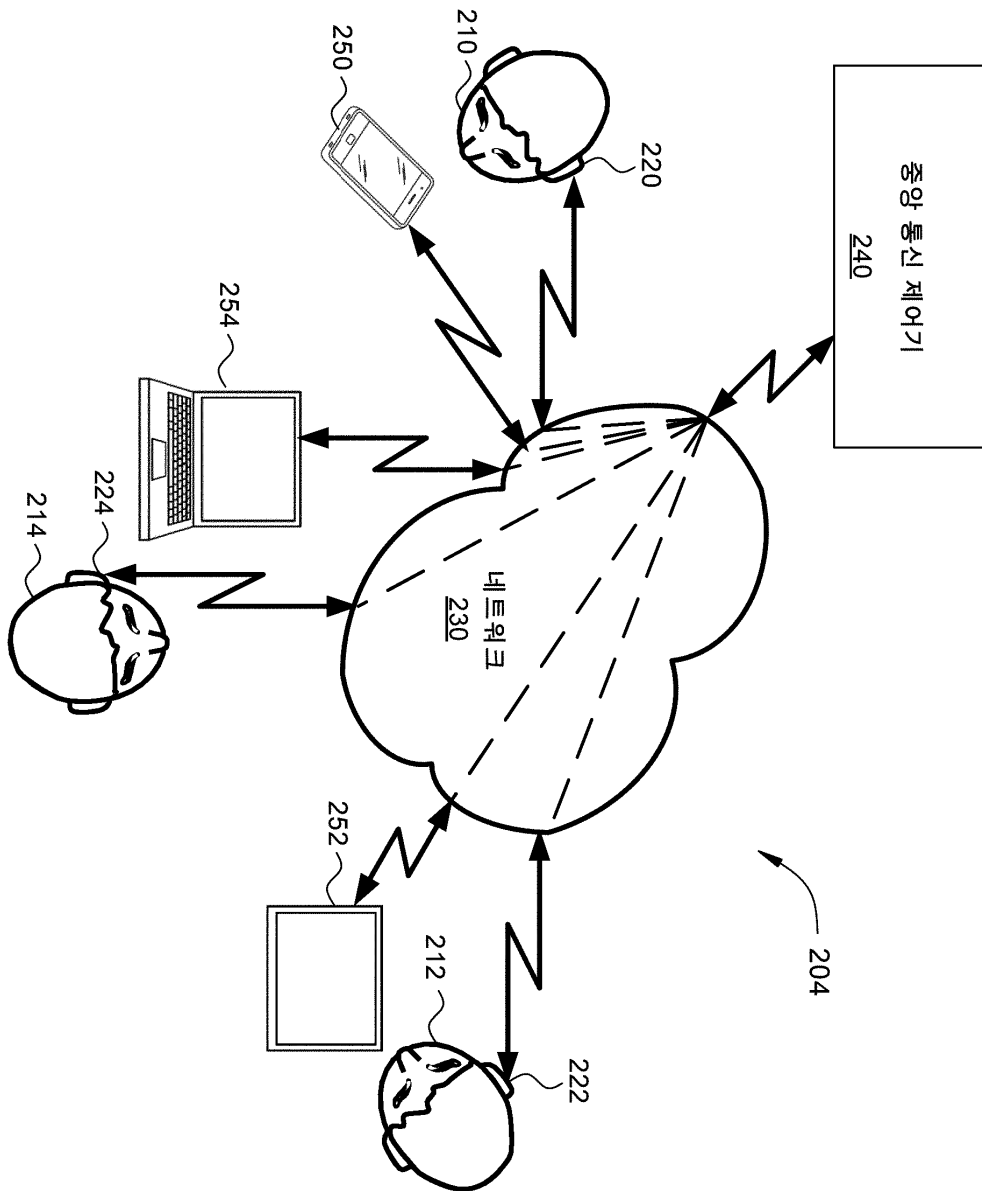
도면2a



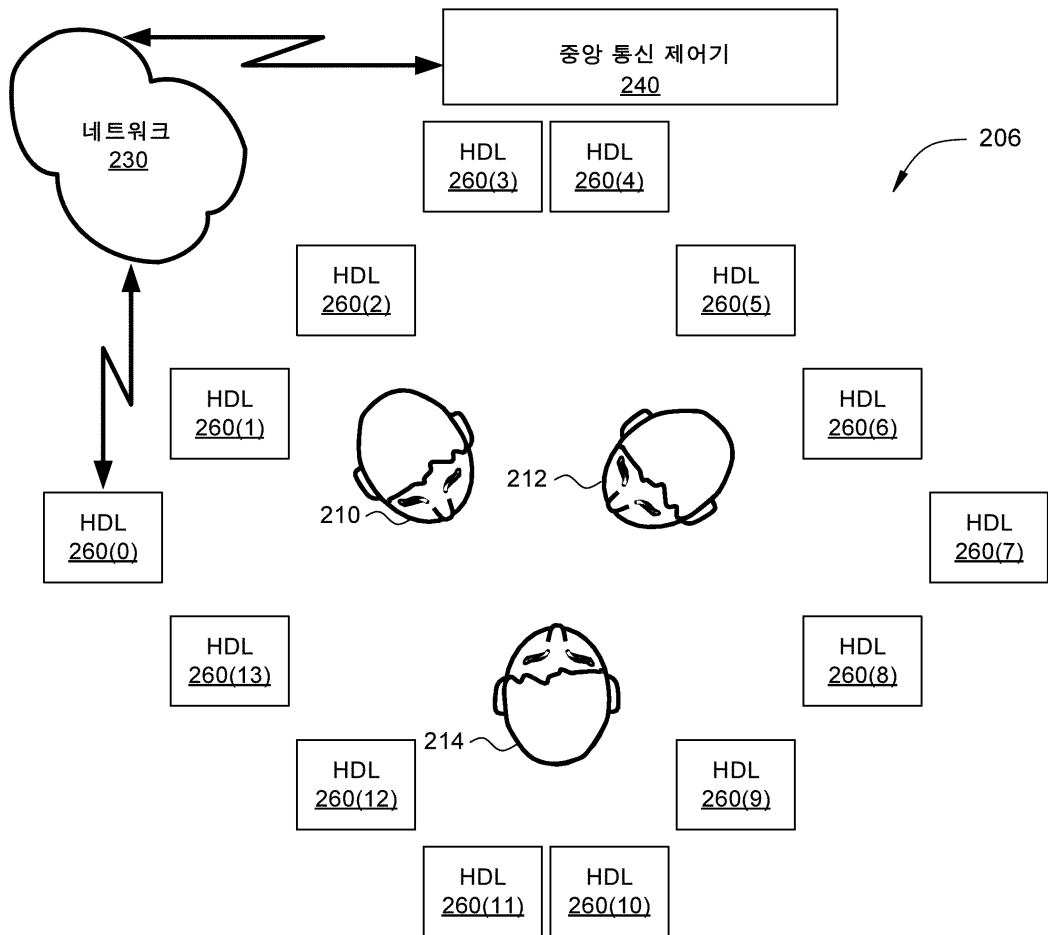
도면2b



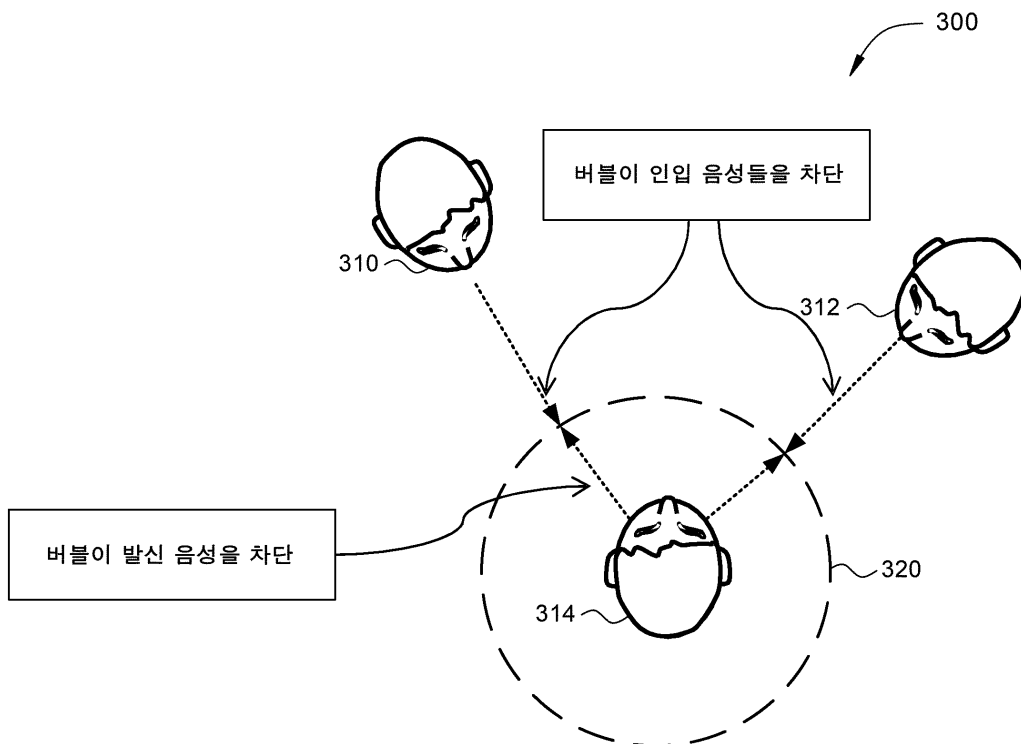
도면2c



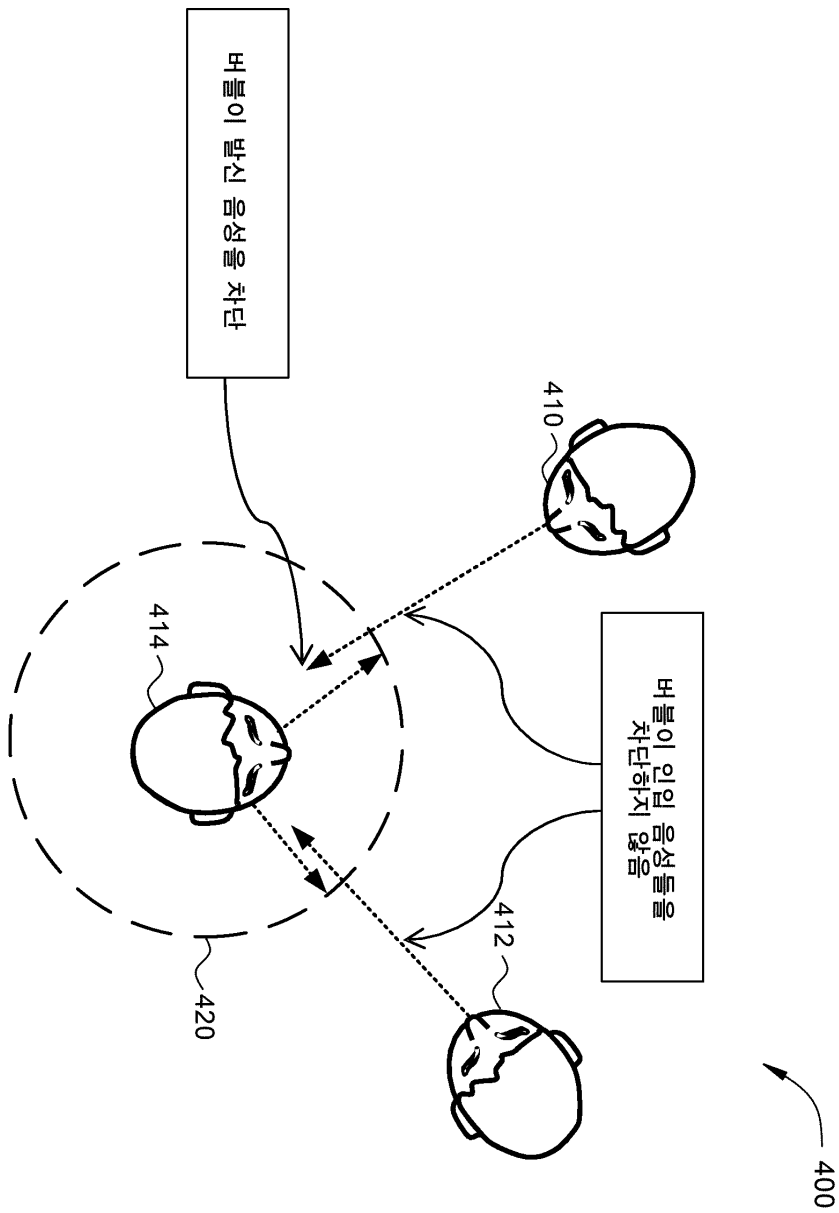
도면2d



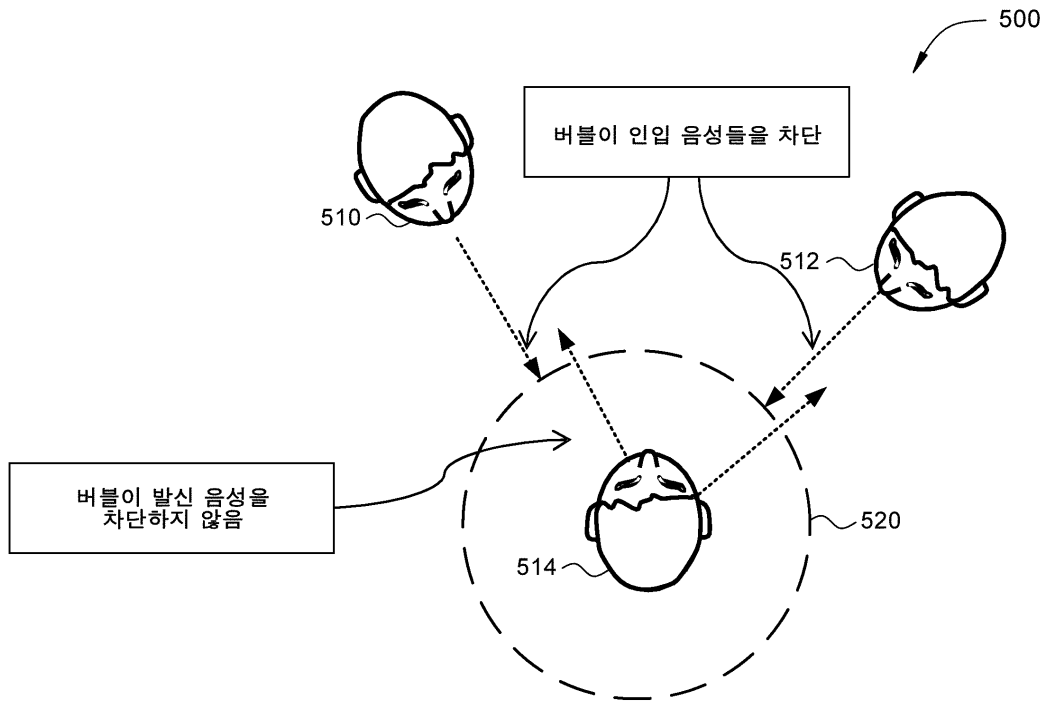
도면3



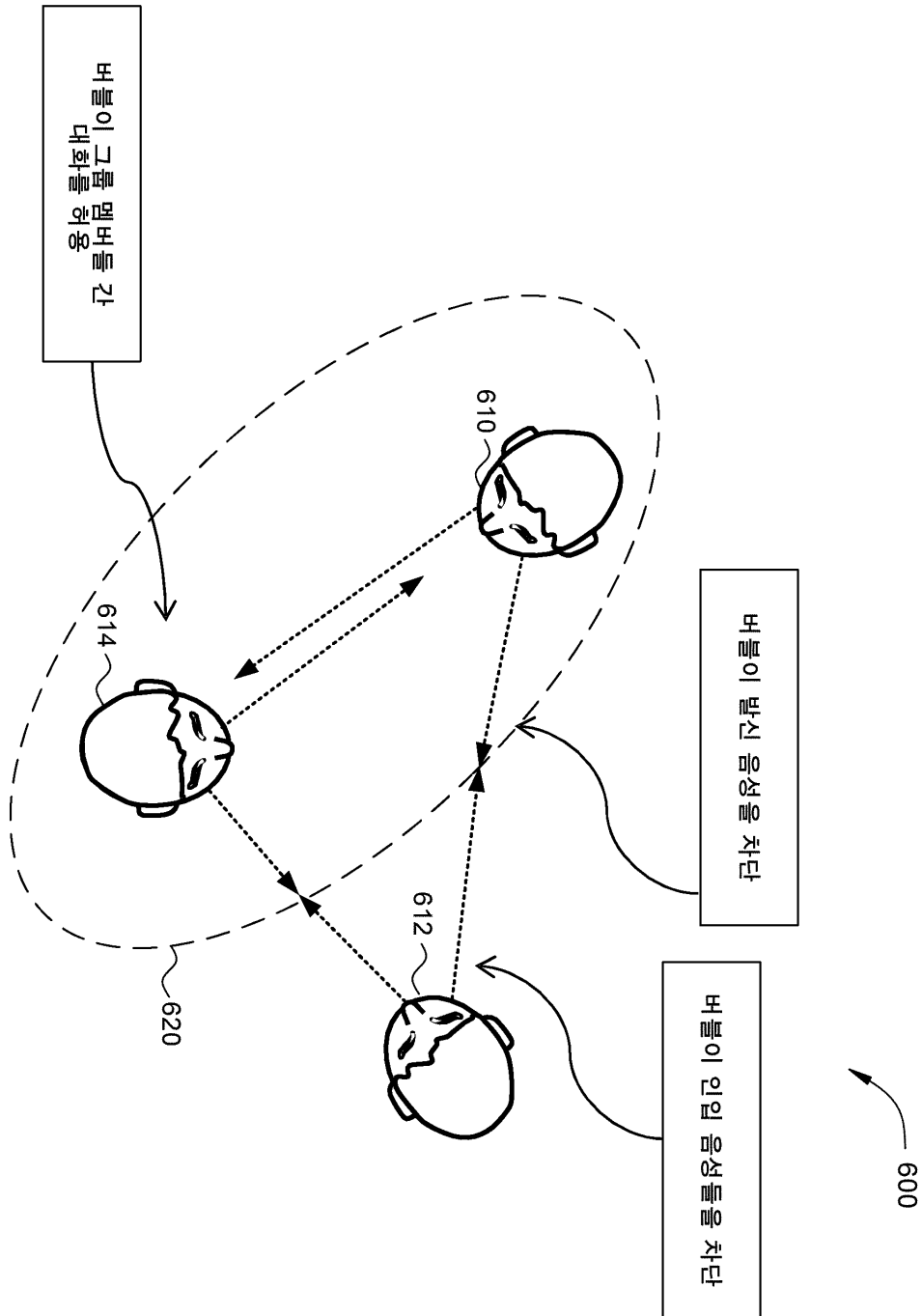
도면4



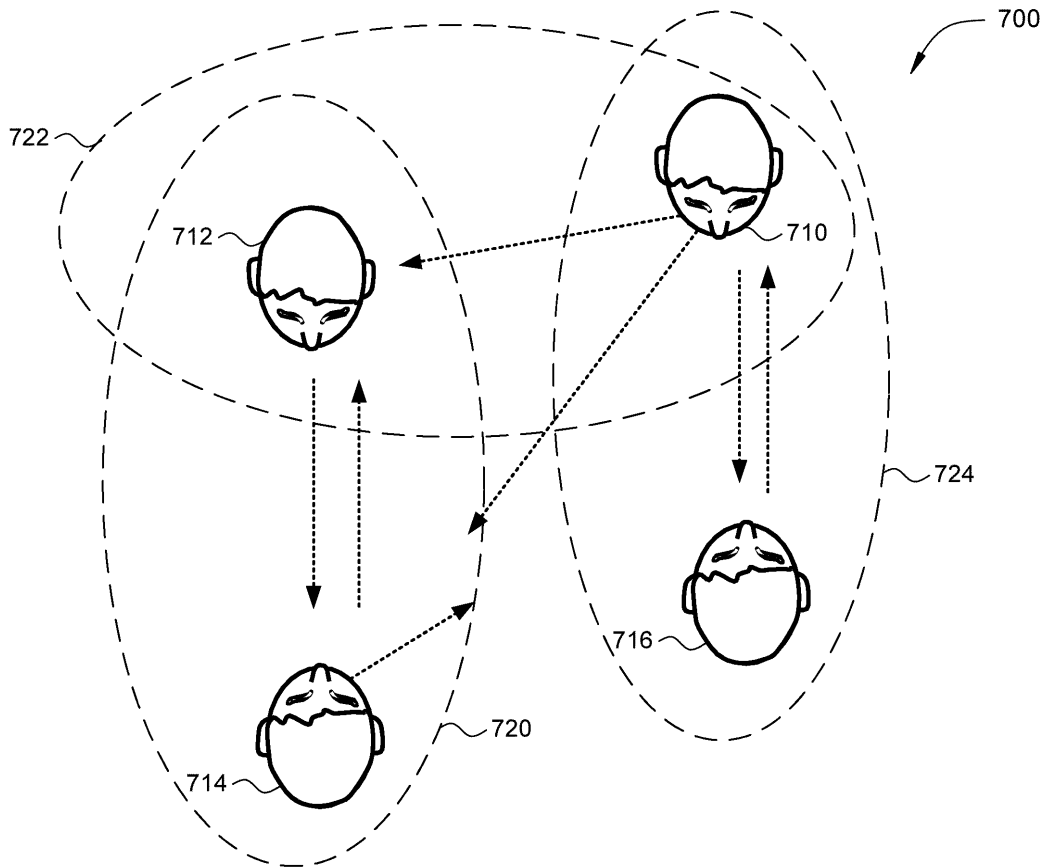
도면5



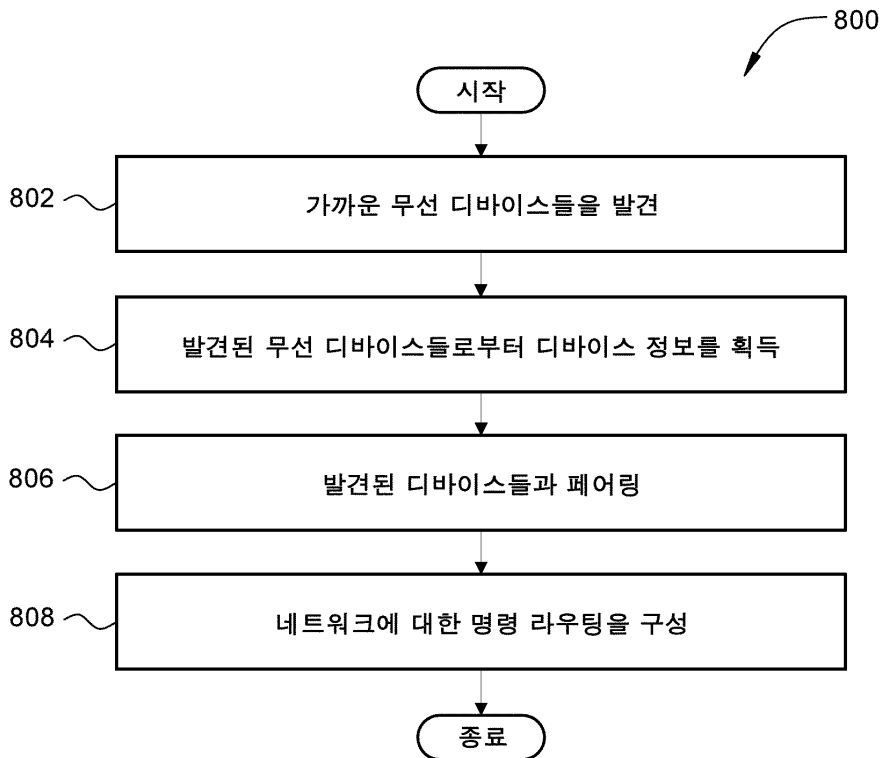
도면6



도면7



도면8



도면9

