

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일

2022년 2월 24일 (24.02.2022)



(10) 국제공개번호

WO 2022/039486 A1

(51) 국제특허분류:

H04R 3/00 (2006.01)

G10L 17/02 (2013.01)

G06F 40/40 (2020.01)

G01S 3/808 (2006.01)

G10L 15/00 (2006.01)

G10L 17/22 (2013.01)

H04W 4/02 (2009.01)

G06F 40/58 (2020.01)

H04R 1/40 (2006.01)

(21) 국제출원번호:

PCT/KR2021/010939

(22) 국제출원일:

2021년 8월 18일 (18.08.2021)

(25) 출원언어:

한국어

(26) 공개언어:

한국어

(30) 우선권정보:

10-2020-0103909 2020년 8월 19일 (19.08.2020) KR

10-2020-0126504 2020년 9월 29일 (29.09.2020) KR

(71) 출원인: 주식회사 아모센스 (AMOSENSE CO., LTD.)
[KR/KR]; 31040 충청남도 천안시 서북구 직산읍 4산단 5길 90 천안 제4지방산업단지 19-1블럭, Chungcheongnam-do (KR).

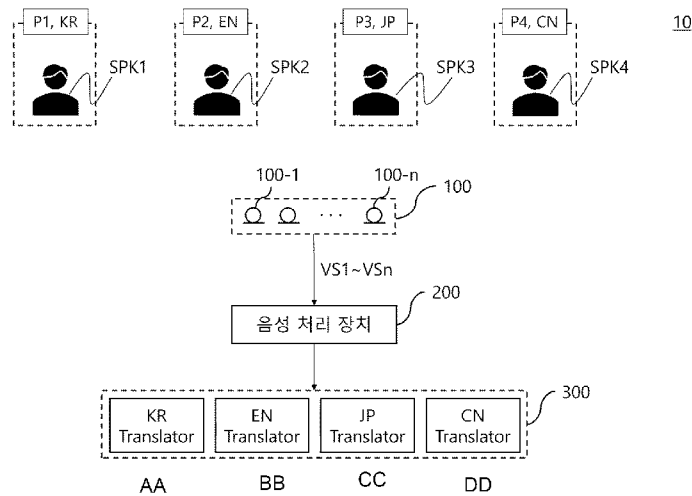
(72) 발명자: 김정민 (KIM, Jungmin); 31040 충청남도 천안시 서북구 직산읍 4산단 5길 90 천안 제4지방산업단지 19-1블럭, Chungcheongnam-do (KR).

(74) 대리인: 김철진 (KIM, Churchill); 06527 서울시 서초구 강남대로97길 37 티앤티빌딩 4층, Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK,

(54) Title: VOICE PROCESSING DEVICE FOR PROCESSING VOICE SIGNAL AND VOICE PROCESSING SYSTEM COMPRISING SAME

(54) 발명의 명칭: 음성 신호를 처리하기 위한 음성 처리 장치 및 이를 포함하는 음성 처리 시스템



200 ... Voice processing device
AA ... KR Translator
BB ... EN Translator
CC ... JP Translator
DD ... CN Translator

(57) Abstract: A voice processing device is disclosed. The voice processing device comprises: a voice data receiving circuit configured to receive input voice data associated with voices of speakers; a memory configured to store starting language data; a voice data output circuit configured to output output voice data associated with the voices of the speakers; and a processor configured to generate a control command for outputting the output voice data, wherein the processor uses the input voice data to generate first speaker position data indicating a position of a first speaker of the speakers and first output voice data associated with a voice of the first speaker, reads first starting language data corresponding to the first speaker position data with reference to the memory, and transmits, to the voice data output circuit, a control command for outputting the first output voice data to a translation environment for translating the first starting

MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

- 국제조사보고서와 함께 (조약 제21조(3))
- 청구범위 보정 기한 만료 전의 공개이며, 보정서를 접수하는 경우 그에 관하여 별도 공개함 (규칙 48.2(h))

language indicated by the first starting language data.

(57) 요약서: 음성 처리 장치가 개시된다. 음성 처리 장치는, 화자들의 음성과 연관된 입력 음성 데이터를 수신하도록 구성되는 음성 데이터 수신 회로, 출발 언어 데이터를 저장하도록 구성되는 메모리, 화자들의 음성과 연관된 출력 음성 데이터를 출력하도록 구성되는 음성 데이터 출력 회로 및 출력 음성 데이터를 출력하기 위한 제어 명령을 생성하도록 구성되는 프로세서를 포함하고, 프로세서는, 입력 음성 데이터를 이용하여, 화자들 중 제1화자의 위치를 나타내는 제1화자 위치 데이터 및 제1화자의 음성과 연관된 제1출력 음성 데이터를 생성하고, 메모리를 참조하여 제1화자 위치 데이터에 대응하는 제1출발 언어 데이터를 리드하고, 제1출발 언어 데이터에 의해 지시되는 제1출발 언어를 번역하기 위한 번역 환경으로 제1출력 음성 데이터를 출력하기 위한 제어 명령을 음성 데이터 출력 회로로 전송한다.

명세서

발명의 명칭: 음성 신호를 처리하기 위한 음성 처리 장치 및 이를 포함하는 음성 처리 시스템

기술분야

- [1] 본 발명의 실시 예들은 음성 신호를 처리하기 위한 음성 처리 장치 및 이를 포함하는 음성 처리 시스템에 관한 것이다.

배경기술

- [2] 마이크(microphone)는 음성을 인식하고, 인식된 음성을 전기적인 신호인 음성 신호로 변환하는 장치이다. 회의실이나 교실과 같이 복수의 화자(speaker)들이 위치하는 공간 내에 마이크가 배치되는 경우, 상기 마이크는 복수의 화자들로부터 나온 음성들을 모두 수신하고, 복수의 화자들의 음성에 연관된 음성 신호들을 생성한다.
- [3] 복수의 화자들이 동시에 발화하는 경우, 개별 화자들의 음성만을 나타내는 음성 신호를 분리하는 것이 필요하다. 또한, 복수의 화자들이 서로 다른 언어로 발화하는 경우, 복수의 화자들의 음성을 쉽게 번역하기 위해서는, 복수의 화자들의 음성의 원래 언어(즉, 출발 언어)를 파악해야 하는데, 음성 자체의 특징만으로 해당 음성의 언어를 파악하는 것은 시간이 많이 소요되고, 많은 리소스가 소요되는 문제가 있다.

발명의 상세한 설명

기술적 과제

- [4] 본 발명이 해결하고자 하는 과제는 입력 음성 신호를 이용하여 화자의 위치를 판단하고, 입력 음성 신호를 이용하여 각 화자의 음성을 나타내는 출력 음성 신호를 생성할 수 있는 음성 처리 장치 및 이를 포함하는 음성 처리 시스템을 제공하는 것에 있다.
- [5] 본 발명이 해결하고자 하는 과제는 음성 신호를 이용하여 화자의 위치를 판단하고, 화자의 위치에 대응하는 출발 언어를 결정하고, 결정된 출발 언어를 번역하기 위한 번역 환경으로 음성 신호를 전송할 수 있는 음성 처리 장치 및 이를 포함하는 음성 처리 시스템을 제공하는 것에 있다.
- [6] 본 발명이 해결하고자 하는 과제는 화자들의 각각의 음성과 연관된 분리 음성 신호를 이용하여, 화자들 각각의 음성에 대한 번역 결과를 생성하고, 생성된 번역 결과를 해당하는 코드리스 이어폰으로 전송할 수 있는 음성 처리 장치 및 이를 포함하는 음성 처리 시스템을 제공하는 것에 있다.

과제 해결 수단

- [7] 본 발명의 실시 예들에 따른 음성 처리 장치는, 화자들의 음성과 연관된 입력 음성 데이터를 수신하도록 구성되는 음성 데이터 수신 회로, 출발 언어 데이터를 저장하도록 구성되는 메모리, 화자들의 음성과 연관된 출력 음성 데이터를

출력하도록 구성되는 음성 데이터 출력 회로 및 출력 음성 데이터를 출력하기 위한 제어 명령을 생성하도록 구성되는 프로세서를 포함하고, 프로세서는, 입력 음성 데이터를 이용하여, 화자들 중 제1화자의 위치를 나타내는 제1화자 위치 데이터 및 제1화자의 음성과 연관된 제1출력 음성 데이터를 생성하고, 메모리를 참조하여 제1화자 위치 데이터에 대응하는 제1출발 언어 데이터를 리드하고, 제1출발 언어 데이터에 의해 지시되는 제1출발 언어를 번역하기 위한 번역 환경으로 제1출력 음성 데이터를 출력하기 위한 제어 명령을 음성 데이터 출력 회로로 전송한다.

- [8] 본 발명의 실시 예들에 따른 음성 처리 시스템은 복수의 코드리스 이어폰들 및 음성 처리 장치를 포함하고, 복수의 코드리스 이어폰들 각각은, 화자들로부터 발화된 음성들과 연관된 음성 신호를 생성하도록 구성되는 마이크 유닛, 음성 신호를 전송하도록 구성되는 통신 유닛 및 음성을 재생하도록 구성되는 스피커 유닛을 포함하고, 음성 처리 장치는, 복수의 코드리스 이어폰들로부터 전송된 음성 신호들을 수신하도록 구성되는 통신 회로, 복수의 코드리스 이어폰들 중 제1코드리스 이어폰으로부터 전송된 제1음성 신호들로부터 화자들 중 제1화자의 음성과 연관된 제1분리 음성 신호를 생성하고, 제1분리 음성 신호를 번역하여 제1번역 음성 신호를 생성하도록 구성되는 음성 처리 회로, 메모리 및 제1번역 음성 신호를 복수의 코드리스 이어폰들 중 제1코드리스 이어폰을 제외한 나머지 코드리스 이어폰으로 전송하도록 구성되는 통신 회로를 포함한다.

발명의 효과

- [9] 본 발명의 실시 예들에 따른 음성 처리 장치는 음성 신호를 이용하여 화자 위치를 파악할 수 있고, 화자 위치를 통해 음성 신호가 어떤 화자의 음성에 대응하는 것인지 구별할 수 있다. 이에 따라, 다수의 화자가 동시에 음성을 발화하더라도, 음성 분리 장치는 음성을 화자 별로 구분하여 분리할 수 있는 효과가 있다.
- [10] 본 발명의 실시 예들에 따른 음성 처리 장치는 음성의 음원 위치에 기초하여 특정 음원 위치로부터의 음성과 연관된 분리 음성 신호를 생성할 수 있으므로, 주변 소음의 영향을 최소화된 음성 신호를 생성할 수 있는 효과가 있다.
- [11] 본 발명의 실시 예들에 따른 음성 처리 장치는 전송된 음성 신호로부터 화자들 각각의 음성을 추출할 수 있을 뿐만 아니라, 음성들의 음원 위치에 기초하여 음성들의 번역 전 언어인 출발 언어를 판단하고, 판단된 출발 언어에 기초하여 해당 음성을 번역하여 번역 결과를 제공할 수 있는 효과가 있다.
- [12] 본 발명의 실시 예들에 따른 음성 처리 장치는 화자들의 각각의 음성과 연관된 분리 음성 신호를 이용하여, 화자들 각각의 음성에 대한 번역 결과를 생성하고, 생성된 번역 결과를 해당하는 코드리스 이어폰으로 전송할 수 있는 효과가 있다.

도면의 간단한 설명

- [13] 도 1은 본 발명의 실시 예들에 따른 음성 처리 시스템을 나타낸다.
- [14] 도 2는 본 발명의 실시 예들에 따른 음성 처리 장치를 나타낸다.
- [15] 도 3 내지 도 6은 본 발명의 실시 예들에 따른 음성 처리 장치의 동작을 설명하기 위한 도면이다.
- [16] 도 7은 본 발명의 실시 예들에 따른 음성 처리 장치의 작동 방법을 나타내는 플로우 차트이다.
- [17] 도 8은 본 발명의 실시 예들에 따른 음성 처리 장치의 작동을 설명하기 위한 도면이다.
- [18] 도 9 및 도 10은 본 발명의 실시 예들에 따른 음성 처리 장치의 작동을 설명하기 위한 도면이다.
- [19] 도 11은 본 발명의 실시 예들에 따른 음성 처리 시스템을 나타낸다.
- [20] 도 12는 본 발명의 실시 예들에 따른 코드리스 이어폰을 나타낸다.
- [21] 도 13은 본 발명의 실시 예들에 따른 음성 처리 장치를 나타낸다.
- [22] 도 14는 본 발명의 실시 예들에 따른 음성 처리 장치의 동작을 설명하기 위한 도면이다.
- [23] 도 15 내지 도 18은 본 발명의 실시 예들에 따른 음성 처리 장치의 번역 기능을 설명하기 위한 도면이다.
- [24] 도 19는 본 발명의 실시 예들에 따른 음성 처리 장치의 작동을 나타내는 플로우 차트이다.

발명의 실시를 위한 형태

- [25] 이하, 첨부된 도면들을 참조하여 본 발명의 실시 예들을 설명한다.
- [26] 도 1은 본 발명의 실시 예들에 따른 음성 처리 시스템을 나타낸다. 도 1을 참조하면, 본 발명의 실시 예들에 따른 음성 처리 시스템(10)은 화자(speaker; SPK1~SPK4)의 음성에 대한 번역을 제공할 수 있다. 화자(SPK1~SPK4)는 각 위치(P1~P4)에 위치할 수 있다. 실시 예들에 따라, 각 위치(P1~P4)에 위치한 화자(SPK1~SPK4)는 화자(SPK1~SPK4)의 언어로 음성을 발화(pronounce)할 수 있다. 예컨대, 제1위치(P1)에 위치한 제1화자(SPK1)는 제1언어(예컨대, 한국어(KR))로 음성을 발화할 수 있고, 제2위치(P2)에 위치한 제2화자(SPK2)는 제2언어(예컨대, 영어(EN))로 음성을 발화할 수 있고, 제3위치(P3)에 위치한 제3화자(SPK3)는 제3언어(예컨대, 일본어(JP))로 음성을 발화할 수 있고, 제4위치(P4)에 위치한 제4화자(SPK4)는 제4언어(예컨대, 중국어(CN))로 음성을 발화할 수 있다.
- [27] 실시 예들에 따라, 음성 처리 시스템(10)은 화자(SPK1~SPK4)의 음성에 기초하여 화자(SPK1~SPK4)의 위치를 결정하고, 화자(SPK1~SPK4)의 음성을 결정된 위치에 대응하는 언어(즉, 출발 언어(source language)에서 다른 언어(즉, 도착 언어(target language))로 번역할 수 있다.
- [28] 즉, 본 발명의 실시 예들에 따른 음성 처리 시스템(10)은 화자(SPK1~SPK4)의

위치에 기초하여 화자(SPK1~SPK4)의 음성의 언어(즉, 출발 언어)를 결정하므로, 화자(SPK1~SPK4)의 음성의 언어를 별도로 인식하는 작업 없이 화자(SPK1~SPK4)의 음성을 번역할 수 있으므로, 번역에 소요되는 시간 및 리소스가 감소하는 효과가 있다.

- [29] 음성 처리 시스템(10)은 화자(SPK1~SPK4)의 음성을 수신하도록 구성되는 복수의 마이크로폰(100), 음성 처리 장치(200) 및 번역 환경(300)을 포함할 수 있다.
- [30] 화자(SPK1~SPK4)의 음성은 복수의 마이크로폰(100)에 의해 수신될 수 있다.
- [31] 복수의 마이크로폰(100)은 각 위치(P1~P4)에 위치한 화자(SPK1~SPK4)의 음성을 수신하고, 화자(SPK1~SPK4)의 음성과 연관된 음성 신호(VS1~VS_n)를 생성할 수 있다. 예컨대, 제1마이크로폰(100-1)은 화자(SPK1~SPK4)의 음성을 수신하고, 화자(SPK1~SPK4)의 음성과 연관된 제1음성 신호(VS1)를 생성할 수 있다. 제1마이크로폰(100-1)에 의해 생성된 제1음성 신호(VS1)는 적어도 하나 이상의 화자(SPK1~SPK4)의 음성에 대응할 수 있다.
- [32] 한편, 본 명세서에서 설명되는 음성 신호는 아날로그 타입의 신호 또는 디지털 타입의 데이터일 수 있다. 실시 예들에 따라, 아날로그 타입의 신호와 디지털 타입의 데이터는 상호 변환될 수 있고, 신호의 타입(아날로그 또는 디지털)이 바뀌더라도 포함된 정보는 실질적으로 동일하므로, 본 발명의 실시 예들에 대한 설명에 있어서, 디지털 타입의 음성 신호와 아날로그 타입의 음성 신호를 혼용하여 설명하도록 한다.
- [33] 이하, 첨부된 도면들을 참조하여 본 발명의 실시 예들을 설명한다.
- [34] 도 1은 본 발명의 실시 예들에 따른 음성 처리 시스템을 나타낸다. 도 1을 참조하면, 본 발명의 실시 예들에 따른 음성 처리 시스템(10)은 화자(speaker; SPK1~SPK4)의 음성에 대한 번역을 제공할 수 있다. 실시 예들에 따라, 음성 처리 시스템(10)은 화자(SPK1~SPK4)의 음성에 기초하여 화자(SPK1~SPK4)의 위치를 결정하고, 화자(SPK1~SPK4)의 음성을 결정된 위치에 대응하는 언어(즉, 출발 언어(source language)에서 다른 언어(즉, 도착 언어(target language))로 번역할 수 있다.
- [35] 즉, 본 발명의 실시 예들에 따른 음성 처리 시스템(10)은 화자(SPK1~SPK4)의 위치에 따라 화자(SPK1~SPK4)의 음성의 출발 언어를 결정하므로, 화자(SPK1~SPK4)의 음성의 언어를 별도로 인식하는 작업 없이 화자(SPK1~SPK4)의 음성을 번역할 수 있으므로, 번역에 소요되는 시간 및 리소스가 감소하는 효과가 있다.
- [36] 음성 처리 시스템(10)은 화자(SPK1~SPK4)의 음성을 수신하도록 구성되는 복수의 마이크로폰(100), 음성 처리 장치(200) 및 번역 환경(300)을 포함할 수 있다.
- [37] 복수의 마이크로폰(100)은 각 위치(P1~P4)에 위치한 화자(SPK1~SPK4)의 음성을 수신하고, 화자(SPK1~SPK4)의 음성을 전기적인 신호인 음성

신호(VS1~VS_n)로 변환할 수 있다. 예컨대, 제1마이크로폰(100-1)은 화자(SPK1~SPK4)의 음성을 수신하고, 화자(SPK1~SPK4)의 음성과 연관된 제1음성 신호(VS1)를 생성할 수 있다. 제1마이크로폰(100-1)에 의해 생성된 제1음성 신호(VS1)는 적어도 하나 이상의 화자(SPK1~SPK4)의 음성에 대응할 수 있다.

- [38] 복수의 마이크로폰(100)은 음성 신호들(VS1~VS_n)을 출력할 수 있다. 실시 예들에 따라, 복수의 마이크로폰(100)은 음성 신호들(VS1~VS_n)을 음성 처리 장치(200)로 전송할 수 있다. 예컨대, 복수의 마이크로폰(100)은 유선 방식 또는 무선 방식에 따라 음성 신호들(VS1~VS_n)을 음성 처리 장치(200)로 전송할 수 있다.
- [39] 복수의 마이크로폰(100)은 빔포밍(*beamforming*) 마이크로 구성되어, 다방향(*multi-direction*)으로부터 음성을 입력받을 수 있다. 실시 예들에 따라, 복수의 마이크로폰(100)은 서로 이격되어 배치되어, 하나의 마이크 어레이를 구성할 수 있으나, 본 발명의 실시 예들이 이에 한정되는 것은 아니다.
- [40] 복수의 마이크로폰(100) 각각은 임의의 특정 방향의 음성을 수신하도록 구성되는 지향성 마이크이거나, 또는, 모든 방향의 음성을 수신하도록 구성되는 무지향성 마이크일 수 있다.
- [41] 음성 처리 장치(200)는 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터를 수신하고, 입력 음성 데이터를 이용하여 화자(SPK1~SPK4)의 음성과 연관된 출력 음성 데이터를 생성하고, 출력 음성 데이터를 번역 환경(300)으로 전송할 수 있다.
- [42] 실시 예들에 따라, 입력 음성 데이터는 음성 신호들(VS1~VS_n)일 수 있다. 예컨대, 음성 처리 장치(200)는 복수의 마이크로폰(100)으로부터 전송되는 음성 신호들(VS1~VS_n)을 수신하고, 음성 신호들(VS1~VS_n)로부터 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터를 획득할 수 있다.
- [43] 한편, 본 명세서에서는 음성 처리 장치(200)가 복수의 마이크로폰(100)으로부터 음성 신호들(VS1~VS_n)을 수신하여 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터를 획득하는 것을 가정하고 설명하나, 실시 예들에 따라, 복수의 마이크로폰(100)은 음성 처리 장치(200)내에 포함될 수 있다.
- [44] 음성 처리 장치(200)는 연산 처리 기능을 갖는 컴퓨팅 장치일 수 있다. 실시 예들에 따라, 음성 처리 장치(200)는 컴퓨터, 노트북, 모바일 디바이스, 스마트폰 또는 웨어러블 디바이스로 구현될 수 있으나, 이에 한정되는 것은 아니다. 예컨대, 음성 처리 장치(200)는 연산 처리 기능을 갖는 적어도 하나의 집적 회로를 포함할 수 있다.
- [45] 음성 처리 장치(200)는 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터를 이용하여 화자(SPK1~SPK4)의 위치(즉, 음원의 위치)를 결정할 수 있다. 실시 예들에 따라, 음성 처리 장치(200)는 복수의 마이크로폰들(100) 사이의 거리,

복수의 마이크로폰들(100) 각각이 화자(SPK1~SPK4)의 음성을 수신한 시점(time)들 사이의 차이, 화자(SPK1~SPK4)의 음성의 크기 중 적어도 하나에 기초하여, 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터로부터 화자(SPK1~SPK4)의 위치를 나타내는 화자 위치 데이터를 생성할 수 있다.

- [46] 음성 처리 장치(200)는 화자(SPK1~SPK4)의 위치를 나타내는 화자 위치 데이터에 기초하여, 입력 음성 데이터를 화자의 위치(즉, 음원의 위치)에 따라 분리할 수 있다. 예컨대, 음성 처리 장치(200)는 입력 음성 데이터를 결정된 화자의 위치 별로 그룹핑할 수 있다.
- [47] 예컨대, 제1화자(SPK1)와 제2화자(SPK2)가 시간적으로 중첩해서 발화하는 경우, 제1화자(SPK1)와 제2화자(SPK2)의 음성은 중첩되므로, 입력 음성 데이터 또한 제1화자(SPK1)의 음성과 연관된 음성 데이터와 제2화자(SPK2)의 음성관 연관된 음성 데이터를 포함할 수 있다. 상술한 바와 같이, 음성 처리 장치(200)는 제1화자(SPK1)의 음성 및 제2화자(SPK2)의 음성과 연관된 입력 음성 데이터로부터 제1화자(SPK1) 및 제2화자(SPK2) 각각의 위치를 나타내는 화자 위치 데이터를 생성하고, 화자 위치 데이터에 기초하여 입력 음성 데이터로부터 제1화자(SPK1)의 음성을 나타내는 제1출력 음성 데이터와, 제2화자(SPK2)의 음성을 나타내는 제2출력 음성 데이터를 생성할 수 있다. 이 때, 제1출력 음성 데이터는 화자들(SPK1~SPK4)의 음성들 중 제1화자(SPK1)의 음성과 가장 높은 연관도를 가지는 음성 데이터일 수 있다. 다시 말하면, 제1출력 음성 데이터에 포함된 음성 성분 중에서 제1화자(SPK1)의 음성 성분의 비중이 가장 높을 수 있다.
- [48] 음성 처리 장치(200)는 입력 음성 데이터를 이용하여 화자(SPK1~SPK4)의 위치에 대응하는 출발 언어를 결정하고, 결정된 출발 언어를 도착 언어로 번역하기 위한 번역 환경(300)으로 출력 음성 데이터를 전송할 수 있다. 예컨대, 음성 처리 장치(200)는 입력 음성 데이터에 기초하여 결정된 화자의 위치가 제1위치(P1)인 경우, 제1위치(P1)에 대응하는 출발 언어(예컨대, 한국어(KR))를 번역하기 위한 번역 환경(300)으로 출력 음성 데이터를 전송할 수 있다.
- [49] 본 발명의 실시 예들에 따른 음성 처리 장치(200)는 화자(SPK1~SPK4)의 입력 음성 데이터를 이용하여 화자(SPK1~SPK4)의 위치를 나타내는 화자 위치 데이터를 생성하고, 화자(SPK1~SPK4)의 위치에 따라 화자(SPK1~SPK4)의 음성의 출발 언어를 결정하고, 출발 언어를 번역하기 위한 번역 환경(300)으로 출력 음성 데이터를 전송할 수 있다.
- [50] 따라서, 음성 처리 장치(200)는 화자(SPK1~SPK4)의 음성의 언어를 별도로 인식하는 작업 없이도, 화자(SPK1~SPK4)의 음성의 언어를 번역하기 위한 번역 환경으로 음성 데이터를 전송할 수 있는 효과가 있다.
- [51] 실시 예들에 따라, 음성 처리 장치(200)는 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터를 처리할 수 있다. 음성 처리 장치(200)는 입력 음성 데이터를 이용하여, 화자(SPK1~SPK4)의 음성과 연관된 텍스트를 포함하는 텍스트

- 데이터를 생성하고, 생성된 텍스트 데이터를 화자 위치 데이터와 매칭하여 저장할 수 있다.
- [52] 또한, 음성 처리 장치(200)는 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터를 텍스트 데이터로 변환하고, 텍스트 데이터를 번역 환경(300)으로 전송할 수도 있다.
- [53] 번역 환경(300)은 언어에 대한 번역을 제공하는 환경 또는 시스템을 의미할 수 있다. 실시 예들에 따라, 번역 환경(300)은 음성 처리 장치(200)로부터 화자(SPK1~SPK4)의 음성과 연관된 출력 음성 데이터를 수신할 수 있고, 다른 언어로 번역된 화자(SPK1~SPK4)의 음성과 연관된 데이터를 출력할 수 있다. 예컨대, 번역 환경(300)은 화자(SPK1~SPK4)의 언어에 대응하는 한국어, 영어, 일본어 및 중국어에 대한 번역을 제공할 수 있다.
- [54] 번역 환경(300)은 번역기들(310~340)을 포함할 수 있다. 번역기들(310)은 출발 언어로 표현된 데이터를 도착 언어로 표현된 데이터로 변환할 수 있는 장치, 또는, 언어 번역을 수행하는 번역자에게 음성을 제공하도록 구성되는 단말기를 의미할 수 있다.
- [55] 실시 예들에 따라, 번역 환경(300)은 언어 번역 기능을 지원하는 장치를 포함할 수 있다. 예컨대, 번역 환경(300)은 음성 데이터를 수신하고, 음성 데이터를 텍스트 데이터로 변환하고, 텍스트 데이터의 언어를 다른 언어로 변환할 수 있는 장치를 포함할 수 있다. 예컨대, 번역 환경(300)은 음성에 대응하는 텍스트 데이터를 수신하고, 텍스트 데이터의 언어를 다른 언어로 변환할 수 있는 장치를 포함할 수 있다.
- [56] 실시 예들에 따라, 번역 환경(300)은 언어 번역을 수행하는 번역자에게 화자(SPK1~SPK4)의 음성을 제공하도록 구성되는 단말기를 포함할 수 있다. 예컨대, 번역 환경(300)은 음성 데이터를 이용하여 번역자에게 음성 데이터에 대응하는 음성을 재생할 수 있는 단말기를 포함할 수 있다. 상기 단말기는, 예를 들면, 스피커 또는 이어폰일 수 있다.
- [57] 도 2는 본 발명의 실시 예들에 따른 음성 처리 장치를 나타낸다. 도 2를 참조하면, 음성 처리 장치(200)는 음성 데이터 수신 회로(210), 메모리(220), 프로세서(230) 및 음성 데이터 출력 회로(240)를 포함할 수 있다.
- [58] 음성 수신 회로(210)는 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터를 수신할 수 있다. 실시 예들에 따라, 음성 수신 회로(210)는 유선 통신 방식 또는 무선 통신 방식에 따라 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터를 수신할 수 있다.
- [59] 실시 예들에 따라, 음성 수신 회로(210)는 ADC(analog to digital) 컨버터를 포함하고, 복수의 마이크로폰(100)으로부터 아날로그 타입의 음성 신호들(VS1~VS_n)을 수신하고, 음성 신호들(VS1~VS_n)을 디지털 타입의 입력 음성 데이터로 변환하고, 변환된 입력 음성 데이터를 저장할 수 있다.
- [60] 실시 예들에 따라, 음성 수신 회로(210)는 무선 통신 방식에 따라 통신 가능한

- 통신 회로를 포함하고, 통신 회로를 통해 입력 음성 데이터를 수신할 수 있다.
- [61] 메모리(220)는 음성 처리 장치(200)의 작동에 필요한 데이터를 저장할 수 있다. 실시 예들에 따라, 메모리(220)는 비휘발성 메모리 및 휘발성 메모리 중 적어도 하나를 포함할 수 있다.
- [62] 메모리(220)는 등록 위치를 나타내는 위치 데이터와, 위치 데이터에 대응하는 출발 언어 데이터를 저장할 수 있다. 실시 예들에 따라, 위치 데이터와 출발 언어 데이터는 메모리(220)에 매칭되어 저장될 수 있다.
- [63] 출발 언어 데이터는 위치 데이터에 대응하는 위치에 위치하는 화자의 음성(또는 입력 음성 데이터)의 출발 언어를 나타낼 수 있다. 예컨대, 도 1에 도시된 바와 같이, 제1위치(P1)를 나타내는 위치 데이터와 대응하는 출발 언어 데이터는 제1위치(P1)에서 발화된 음성의 출발 언어(예컨대, 한국어)를 나타낼 수 있다. 즉, 출발 언어 데이터는 위치 데이터에 대응하는 위치에서 발화된 음성(또는 입력 음성 데이터)의 출발 언어를 나타낼 수 있다.
- [64] 프로세서(230)는 음성 처리 장치(200)의 전반적인 동작을 제어할 수 있다. 실시 예들에 따라, 프로세서(230)는 음성 데이터 수신 회로(210), 메모리(220) 및 음성 데이터 출력 회로(240)의 동작을 제어하기 위한 제어 명령을 생성하고, 제어 명령을 음성 데이터 수신 회로(210), 메모리(220) 및 음성 데이터 출력 회로(240) 각각으로 전송할 수 있다.
- [65] 프로세서(230)는 연산 처리 기능을 갖는 집적회로로 구현될 수 있다. 예컨대, 프로세서(230)는 CPU(central processing unit), MCU(micro controller unit), DSP(digital signal processor), GPU(graphics processing unit), ASIC(application specific integrated circuit) 또는 FPGA(field programmable gate array)를 포함할 수 있으나, 본 발명의 실시 예들이 이에 한정되는 것은 아니다.
- [66] 프로세서(230)는 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터를 이용하여 화자(SPK1~SPK4)의 위치(즉, 음성의 음원 위치)를 판단하고, 화자(SPK1~SPK4)의 위치를 나타내는 화자 위치 데이터를 생성할 수 있다. 예컨대, 프로세서(230)는 화자 위치 데이터를 메모리(220)에 저장할 수 있다.
- [67] 프로세서(230)는 복수의 마이크로폰들(100) 사이의 거리, 복수의 마이크로폰들(100) 각각이 화자(SPK1~SPK4)의 음성을 수신한 시점들 사이의 차이, 화자(SPK1~SPK4)의 음성의 크기 중 적어도 하나에 기초하여, 화자(SPK1~SPK4)의 음성과 연관된 입력 음성 데이터로부터 화자(SPK1~SPK4)의 위치를 나타내는 화자 위치 데이터를 생성할 수 있다.
- [68] 프로세서(230)는 화자(SPK1~SPK4)의 위치를 나타내는 화자 위치 데이터에 기초하여, 입력 음성 데이터를 화자의 위치(즉, 음원의 위치)에 따라 분리할 수 있다. 예컨대, 음성 처리 장치(200)는 위치에 따라 분리된 음성 데이터와 해당하는 화자 위치 데이터를 매칭하여 저장할 수 있다.
- [69] 실시 예들에 따라, 프로세서(230)는 제1화자(SPK1)의 음성 및 제2화자(SPK2)의 음성과 연관된 입력 음성 데이터로부터 제1화자(SPK1) 및 제2화자(SPK2)

각각의 위치를 나타내는 화자 위치 데이터를 생성하고, 화자 위치 데이터에 기초하여 입력 음성 데이터로부터 제1화자(SPK1)의 음성과 연관된 제1출력 음성 데이터와 제2화자(SPK2)의 음성과 연관된 제2출력 음성 데이터를 생성할 수 있다. 예컨대, 프로세서(230)는 제1출력 음성 데이터와 제1화자 위치 데이터를 매칭하여 저장하고, 제2출력 음성 데이터와 제2화자 위치 데이터를 매칭하여 저장할 수 있다.

- [70] 프로세서(230)는 화자 위치 데이터를 이용하여, 화자(SPK1~SPK4)의 음성의 출발 언어를 결정할 수 있다. 실시 예들에 따라, 프로세서(230)는 메모리(220)를 참조하여, 화자(SPK1~SPK4)의 화자 위치 데이터에 대응하는 위치 데이터를 결정하고, 결정된 위치 데이터에 매칭된 출발 언어 데이터를 결정하고, 결정된 출발 언어 데이터에 의해 지시되는 언어를 화자(SPK1~SPK4)의 음성의 출발 언어로서 결정할 수 있다. 예컨대, 프로세서(230)는 화자(SPK1~SPK4)의 음성과 연관된 (출력 또는 입력) 음성 데이터와 상기 음성의 출발 언어를 나타내는 출발 언어 데이터를 매칭하여 저장할 수 있다.
- [71] 실시 예들에 따라, 프로세서(230)는 화자(SPK1~SPK4)의 음성의 출발 언어를 번역하기 위한 번역 환경(300)으로 출력 음성 데이터를 전송하기 위한 제어 명령을 생성할 수 있다.
- [72] 음성 데이터 출력 회로(240)는 화자(SPK1~SPK4)의 음성과 연관된 출력 음성 데이터를 출력할 수 있다. 실시 예들에 따라, 음성 데이터 출력 회로(240)는 유선 통신 방식 또는 무선 통신 방식에 따라 화자(SPK1~SPK4)의 음성과 연관된 출력 음성 데이터를 출력할 수 있다.
- [73] 실시 예들에 따라, 음성 신호 출력 회로(250)는 통신 회로를 포함하고, 출력 음성 데이터를 외부 장치로 전송할 수 있다.
- [74] 음성 데이터 출력 회로(240)는 제어 명령에 응답하여, 화자(SPK1~SPK4)의 위치에 대응하는 출발 언어를 도착 언어로 번역하기 위한 번역 환경(300)으로 음성 데이터를 전송할 수 있다.
- [75] 도 3 내지 도 6은 본 발명의 실시 예들에 따른 음성 처리 장치의 동작을 설명하기 위한 도면이다.
- [76] 도 3을 참조하면, 제1화자(SPK1)는 제1위치(P1)에서 한국어(KR)로 된 음성을 발화하고, 제2화자(SPK2)는 제2위치(P2)에서 영어(EN)로 된 음성을 발화하고, 제3화자(SPK3)는 제3위치(P3)에서 일본어(JP)로 된 음성을 발화하고, 제4화자(SPK4)는 제4위치(P4)에서 중국어(CN)로 된 음성을 발화할 수 있다. 즉, 제1화자(SPK1)의 음성의 출발 언어는 한국어(KR)이고, 제2화자(SPK2)의 음성의 출발 언어는 영어(EN)이고, 제3화자(SPK3)의 음성의 출발 언어는 일본어(JP)이고, 제4화자(SPK4)의 음성의 출발 언어는 중국어(CN)가 된다.
- [77] 음성 처리 장치(200)는 위치 데이터(PD1~PD4) 및 위치 데이터(PD1~PD4)에 대응하는 출발 언어 데이터(SLD1~SLD4)를 저장할 수 있다. 위치 데이터(PD1~PD4)는 사전에 정의된 화자의 위치를 나타낼 수 있고, 출발 언어

데이터(SLD1~SLD4)는 해당 위치에 위치한 화자의 출발 언어를 나타낼 수 있다. 예컨대, 도 3의 경우, 제1위치 데이터(PD1)는 제1위치(P1)를 나타내고, 제1출발 언어 데이터(SLD1)는 한국어(KR)를 나타낼 수 있고, 제2위치 데이터(PD2)는 제2위치(P2)를 나타내고, 제2출발 언어 데이터(SLD2)는 영어(EN)를 나타낼 수 있다.

- [78] 도 4를 참조하면, 제2화자(SPK2)가 음성 "☆☆☆"를 영어(EN)로 발화하면, 음성 처리 장치(200)는 제2화자(SPK2)의 음성 "☆☆☆"에 대응하는 입력 음성 데이터를 수신할 수 있다. 예컨대, 복수의 마이크로폰(100)은 음성 "☆☆☆"에 대응하는 음성 신호들(VS1~VS_n)을 생성할 수 있고, 음성 처리 장치(200)는 제2화자(SPK2)의 음성 "☆☆☆"에 대응하는 음성 신호들(VS1~VS_n)을 수신하고, 음성 신호들(VS1~VS_n)로부터 입력 음성 데이터를 생성할 수 있다.
- [79] 음성 처리 장치(200)는 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 입력 음성 데이터를 이용하여, 음성 "☆☆☆"의 음원의 위치, 즉, 제2화자(SPK2)의 위치를 나타내는 제2화자 위치 데이터(SPD2)를 생성할 수 있다.
- [80] 음성 처리 장치(200)는 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 입력 음성 데이터를 이용하여, 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 제2출력 음성 데이터(OVD2)를 생성할 수 있다. 예컨대, 음성 처리 장치(200)는 제2출력 음성 데이터(OVD2)와 제2화자 위치 데이터(SPD2)를 매칭하여 저장할 수 있다.
- [81] 도 5를 참조하면, 음성 처리 장치(200)는 제2화자(SPK2)의 제2화자 위치 데이터(SPD2)에 기초하여, 제2화자(SPK2)의 음성 "☆☆☆"의 출발 언어를 나타내는 제2출발 언어 데이터(SLD2)를 메모리(220)로부터 리드(read)할 수 있다.
- [82] 실시 예들에 따라, 음성 처리 장치(200)는 메모리(220)에 저장된 위치 데이터(PD1~PD4) 중에서 제2화자 위치 데이터(SPD2)에 대응하는 제2위치 데이터(PD2)를 결정할 수 있다. 예컨대, 음성 처리 장치(200)는 위치 데이터(PD1~PD4) 중에서, 제2화자 위치 데이터(SPD2)와 동일 또는 유사한 위치를 나타내는 위치 데이터(예컨대, 제2위치 데이터(PD2))를 결정할 수 있다. 이후, 음성 처리 장치(200)는 제2위치 데이터(PD2)와 대응하는 제2출발 언어 데이터(SLD2)를 메모리(220)로부터 리드할 수 있다.
- [83] 이에 따라, 음성 처리 장치(200)는 제2출발 언어 데이터(SLD2)에 기초하여 제2화자(SPK2)의 음성 "☆☆☆"의 출발 언어를 결정할 수 있다.
- [84] 도 6을 참조하면, 음성 처리 장치(200)는 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 제2출력 음성 데이터(OVD2)를 번역 환경(300)으로 전송할 수 있다.
- [85] 실시 예들에 따라, 음성 처리 장치(200)는 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 제2출력 음성 데이터(OVD2)를, 제2화자(SPK2)의 음성 "☆☆☆"의 출발 언어(예컨대, 영어(EN))를 번역하기 위한 번역 환경(300)으로 전송할 수 있다. 예컨대, 음성 처리 장치(200)는 제2출력 음성 데이터(OVD2)를, 제2화자(SPK2)의 음성 "☆☆☆"의 출발 언어(예컨대, 영어(EN))로 표현되는 텍스트 데이터로

- 변환하고, 변환된 텍스트 데이터를 번역 환경(300)으로 전송할 수 있다.
- [86] 예컨대, 음성 처리 장치(200)는 영어 번역을 수행할 수 있는 영어 번역 장치로 제2출력 음성 데이터(OVD2)를 전송할 수 있다. 예컨대, 음성 처리 장치(200)는 영어 번역을 수행할 수 있는 번역가로 음성을 제공하는 단말기(예컨대, 스피커)로 제2출력 음성 데이터(OVD2)를 전송할 수 있다.
- [87] 본 발명의 실시 예들에 따른 음성 처리 장치(200)는 화자(SPK1~SPK4)의 위치에 따라 화자(SPK1~SPK4)의 음성의 출발 언어를 결정하고, 결정된 출발 언어를 번역하기 위한 번역 환경으로 화자(SPK1~SPK4)의 음성과 연관된 음성 데이터를 전송할 수 있다. 이에 따라, 음성 처리 장치(200)는 화자(SPK1~SPK4)의 음성에 대한 별도의 분석 작업(예컨대, 피치 분석 등) 또는 학습 작업 없이도, 화자(SPK1~SPK4)의 위치에 따라 화자(SPK1~SPK4)의 음성의 출발 언어를 판단할 수 있어 번역에 소요되는 시간 및 리소스가 감소하는 효과가 있다.
- [88] 도 7은 본 발명의 실시 예들에 따른 음성 처리 장치의 작동 방법을 나타내는 플로우 차트이다. 도 7을 참조하면, 음성 처리 장치(200)는 위치 데이터와 출발 언어 데이터를 저장할 수 있다(S110). 실시 예들에 따라, 음성 처리 장치(200)는 메모리(220)에 위치 데이터와 상기 위치 데이터에 대응하는 출발 언어 데이터를 저장할 수 있다. 예컨대, 출발 언어 데이터는 위치 데이터에 대응하는 위치에 위치하는 화자의 음성(또는 음성 데이터)의 출발 언어를 나타낼 수 있다.
- [89] 음성 처리 장치(200)는 화자(SPK1~SPK4)의 음성에 관련된 입력 음성 데이터를 수신할 수 있다(S120). 음성 처리 장치(200)는 수신된 입력 음성 데이터를 저장할 수 있다.
- [90] 예컨대, 음성 처리 장치(200)는 복수의 마이크로폰(100)으로부터 아날로그 타입의 음성 신호들을 수신하고, 음성 신호들로부터 입력 음성 데이터를 획득할 수 있다. 예컨대, 음성 처리 장치(200)는 무선 통신 방식에 따라 입력 음성 데이터를 수신할 수 있다.
- [91] 음성 처리 장치(200)는 입력 음성 데이터를 이용하여, 화자(SPK1~SPK4)의 위치를 나타내는 화자 위치 데이터를 생성할 수 있다(S130).
- [92] 음성 처리 장치(200)는 입력 음성 데이터를 이용하여, 입력 음성 데이터와 연관된 음성의 음원의 위치를 계산할 수 있다. 음원의 위치가 곧 화자(SPK1~SPK4)의 위치이므로, 음성 처리 장치(200)는 계산된 음원의 위치를 화자(SPK1~SPK4)의 화자 위치 데이터로서 생성할 수 있다.
- [93] 음성 처리 장치(200)는 입력 언어 데이터를 이용하여, 화자(SPK1~SPK4)의 음성과 연관된 출력 음성 데이터를 생성할 수 있다. 예컨대, 음성 처리 장치(200)는 입력 음성 데이터에 기초하여, 계산된 화자의 위치에서 발화된 음성들과만 연관된 출력 음성 데이터를 생성할 수 있다.
- [94] 음성 처리 장치(200)는 화자 위치 데이터와 위치 데이터를 비교하고, 화자 위치 데이터에 대응하는 출발 언어 데이터를 리드할 수 있다(S140).
- [95] 실시 예들에 따라, 음성 처리 장치(200)는 저장된 위치 데이터 중에서 화자 위치

데이터와 대응하는 위치 데이터를 결정하고, 결정된 위치 데이터에 대응하는 출발 언어 데이터를 메모리(220)로부터 리드할 수 있다. 상술한 바와 같이, 메모리(220)에는 위치 데이터와 대응하는 출발 언어 데이터가 매칭되어 저장되어 있으므로, 음성 처리 장치(200)는 화자 위치 데이터를 이용하여, 화자(SPK1~SPK4)의 위치와 대응하는 출발 언어를 나타내는 출발 언어 데이터를 결정할 수 있다.

- [96] 음성 처리 장치(200)는 출발 언어 데이터를 이용하여, 출발 언어를 번역하기 위한 번역 환경으로 출력 음성 데이터를 전송할 수 있다(S150).
- [97] 실시 예들에 따라, 음성 처리 장치(200)는 출발 언어 데이터가 나타내는 출발 언어를 번역하기 위한 번역 환경으로 출력 음성 데이터를 전송할 수 있다.
- [98] 예컨대, 음성 처리 장치(200)는 각각이 여러 출발 언어를 번역하도록 구성되는 복수의 번역 장치들 중에서, 리드된 출발 언어 데이터에 의해 지시되는 출발 언어를 번역하도록 구성되는 번역 장치로 출력 음성 데이터를 전송할 수 있다.
- [99] 예컨대, 도 3 내지 도 6에 도시된 예시의 경우, 음성 처리 장치(200)는 한국어, 영어, 일본어 및 중국어를 번역하도록 구성되는 복수의 번역 장치들 중에서, 제2화자(SPK2)의 언어의 출발 언어에 대응하는 영어를 번역하도록 구성되는 번역 장치로 출력 음성 데이터를 전송할 수 있다.
- [100] 도 8은 본 발명의 실시 예들에 따른 음성 처리 장치의 작동을 설명하기 위한 도면이다. 도 8을 참조하면, 음성 처리 장치(200)는 위치 데이터(PD1~PD4), 위치 데이터(PD1~PD4)에 대응하는 출발 언어 데이터(SLD1~SLD4) 및 도착 언어 데이터(TLD1~TLD4)를 저장할 수 있다.
- [101] 도착 언어 데이터(TLD1~TLD4)는 해당 위치에 위치한 화자의 음성의 도착 언어를 나타낼 수 있다. 예컨대, 도착 언어는 화자(SPK1~SPK4) 별로 다르게 설정될 수 있으나, 이에 한정되는 것은 아니다.
- [102] 도 3 내지 도 6과 비교할 때, 도 8을 참조하여 설명되는 음성 처리 장치(200)는 화자 위치 데이터에 대응하는 출발 언어 데이터(SLD1~SLD4) 뿐만 아니라, 도착 언어 데이터(TLD1~TLD4)를 추가적으로 리드할 수 있다.
- [103] 실시 예들에 따라, 음성 처리 장치(200)는 화자 위치 데이터와 위치 데이터를 비교하고, 화자 위치 데이터에 대응하는 출발 언어 데이터(SLD1~SLD4) 및 도착 언어 데이터(TLD1~TLD4)를 리드할 수 있다. 예컨대, 음성 처리 장치(200)는 저장된 위치 데이터 중에서 화자 위치 데이터와 대응하는 위치 데이터를 결정하고, 결정된 위치 데이터에 대응하는 출발 언어 데이터(SLD1~SLD4) 및 도착 언어 데이터(TLD1~TLD4)를 메모리(220)로부터 리드할 수 있다.
- [104] 음성 처리 장치(200)는 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 제2출력 음성 데이터(OVD2)를, 제2화자(SPK2)의 음성 "☆☆☆"의 출발 언어(예컨대, 영어(EN))를 도착 언어(예컨대, 한국어(KR))로 번역하기 위한 번역 환경(300)으로 전송할 수 있다.
- [105] 본 발명의 실시 예들에 따른 음성 처리 장치(200)는 화자(SPK1~SPK4)의

위치에 따라 화자(SPK1~SPK4)의 음성의 출발 언어 및 도착 언어를 결정하고, 결정된 출발 언어를 도착 언어로 번역하기 위한 번역 환경으로 화자(SPK1~SPK4)의 음성과 연관된 음성 데이터를 전송할 수 있다.

- [106] 도 9 및 도 10은 본 발명의 실시 예들에 따른 음성 처리 장치의 작동을 설명하기 위한 도면이다. 도 9를 참조하면, 제1화자(SPK1)가 음성 "㉠㉠㉠"을 발화하고, 제2화자(SPK2)가 음성 "☆☆☆"를 영어(EN)로 발화하면, 음성 처리 장치(200)는 제1화자(SPK1)의 음성 "㉠㉠㉠"과 연관된 제1입력 음성 데이터와 제2화자(SPK2)의 음성 "☆☆☆"에 대응하는 입력 음성 데이터를 수신할 수 있다.
- [107] 실시 예들에 따라, 음성 처리 장치(200)는 제1화자(SPK1)의 음성 "㉠㉠㉠"과 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 입력 음성 데이터를 수신하고, 입력 음성 데이터로부터 제1화자(SPK1) 및 제2화자(SPK2) 각각의 위치를 나타내는 화자 위치 데이터(SPD1 및 SPD2)를 생성하고, 화자 위치 데이터(SPD1 및 SPD2)에 기초하여 입력 음성 데이터로부터 제1화자(SPK1)의 음성 "㉠㉠㉠"과 연관된 제1출력 음성 데이터(OVD1)와 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 제2출력 음성 데이터(OVD2)를 생성할 수 있다.
- [108] 예컨대, 음성 처리 장치(200)는 제1화자 위치 데이터(SPD1)와 제1출력 음성 데이터(OVD1)를 매칭하여 저장하고, 제2화자 위치 데이터(SPD2)와 제2출력 음성 데이터(OVD2)를 매칭하여 저장할 수 있다.
- [109] 음성 처리 장치(200)는 제1화자(SPK1)의 제1화자 위치 데이터(SPD1)에 기초하여, 제1화자(SPK1)의 음성 "㉠㉠㉠"의 출발 언어를 나타내는 제1출발 언어 데이터(SLD1)를 메모리(220)로부터 리드할 수 있다. 또한, 음성 처리 장치(200)는 제2화자(SPK2)의 제2화자 위치 데이터(SPD2)에 기초하여, 제2화자(SPK2)의 음성 "☆☆☆"의 출발 언어를 나타내는 제2출발 언어 데이터(SLD2)를 메모리(220)로부터 리드할 수 있다. 예컨대, 상술한 바와 같이, 메모리(220)에 저장된 위치 데이터(PD1~PD4) 중에서 제1화자 위치 데이터(SPD1)와 동일 또는 유사한 위치를 나타내는 위치 데이터(예컨대, 제1위치 데이터(PD1))를 결정하고, 제1위치 데이터(PD1)와 대응하는 제1출발 언어 데이터(SLD1)를 메모리(220)로부터 리드할 수 있다.
- [110] 이에 따라, 음성 처리 장치(200)는 출발 언어 데이터(SLD1 및 SLD2)에 기초하여 제1화자(SPK1)의 음성 "㉠㉠㉠"의 출발 언어와 제2화자(SPK2)의 음성 "☆☆☆"의 출발 언어를 결정할 수 있다. 예컨대, 음성 처리 장치(200)는 제1화자(SPK1)의 음성 "㉠㉠㉠"과 연관된 제1출력 음성 데이터(OVD1)와 제1출발 언어 데이터(SLD1)를 매칭하여 저장하고, 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 제2출력 음성 데이터(OVD2)와 제2출발 언어 데이터(SLD2)를 매칭하여 저장할 수 있다.
- [111] 또한, 음성 처리 장치(200)는 출발 언어 데이터(SLD1 및 SLD2)에 기초하여 제1화자(SPK1)의 음성 "㉠㉠㉠"의 출발 언어와 제2화자(SPK2)의 음성 "☆☆☆"의 도착 언어를 결정할 수 있다. 실시 예들에 따라, 음성 처리 장치(200)는

화자(SPK1~SPK4)의 음성의 도착 언어를, 화자(SPK1~SPK4) 자신을 제외한 나머지 화자의 출발 언어에 기초하여 결정할 수 있다.

- [112] 예컨대, 음성 처리 장치(200)는 제2화자(SPK2)의 음성 "☆☆☆"의 도착 언어를, 제2화자(SPK2) 자신을 제외한 제1화자(SPK1)의 음성 "㉠㉠㉠"의 출발 언어(예컨대, 한국어)로 설정할 수 있다. 또한, 예컨대, 음성 처리 장치(200)는 제1화자(SPK1)의 음성 "㉠㉠㉠"의 도착 언어를, 제1화자(SPK1) 자신을 제외한 제2화자(SPK2)의 음성 "☆☆☆"의 출발 언어(예컨대, 영어)로 설정할 수 있다.
- [113] 도 10을 참조하면, 음성 처리 장치(200)는 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 제2출력 음성 데이터(OVD2)를 번역 환경(300)으로 전송할 수 있다. 실시 예들에 따라, 음성 처리 장치(200)는 제2화자(SPK2)의 음성 "☆☆☆"과 연관된 제2출력 음성 데이터(OVD2)를, 제2화자(SPK2)의 음성 "☆☆☆"의 출발 언어(예컨대, 영어(EN))를 도착 언어(예컨대, 한국어(KR))로 번역하기 위한 번역 환경(300)으로 전송할 수 있다.
- [114] 음성 처리 장치(200)는 영어로 표현되는 음성 "☆☆☆"과 연관된 제2출력 음성 데이터(OVD2)를 번역 환경(300)으로 전송하고, 번역 환경(300)으로부터 제2출력 음성 데이터(OVD2)의 번역 결과를 수신할 수 있다.
- [115] 실시 예들에 따라, 음성 처리 장치(200)는 제2출력 음성 데이터(OVD2)를, 제2화자(SPK2)의 음성 "☆☆☆"의 출발 언어(예컨대, 영어(EN))로 표현되는 텍스트 데이터로 변환하고, 변환된 텍스트 데이터를 번역 환경(300)으로 전송할 수 있다. 또한, 음성 처리 장치(200)는 번역 환경(300)으로부터 음성 "☆☆☆"의 도착 언어(예컨대, 한국어(KR))로 표현되는 텍스트 데이터를 수신하고, 텍스트 데이터를 이용하여 음성 "☆☆☆"의 도착 언어(예컨대, 한국어(KR))로 표현되는 데이터(음성 데이터 또는 텍스트 데이터)를 제1화자(SPK1)로 제공할 수 있다.
- [116] 본 발명의 실시 예들에 따른 음성 처리 장치(200)는 화자(SPK1~SPK4)의 위치에 따라 화자(SPK1~SPK4)의 음성의 출발 언어를 결정하고, 결정된 출발 언어를 번역하기 위한 번역 환경으로 화자(SPK1~SPK4)의 음성과 연관된 음성 데이터를 전송할 수 있다. 이에 따라, 음성 처리 장치(200)는 화자(SPK1~SPK4)의 음성에 대한 별도의 분석 작업(예컨대, 피치 분석 등) 또는 학습 작업 없이도, 화자(SPK1~SPK4)의 위치에 따라 화자(SPK1~SPK4)의 음성의 출발 언어를 판단할 수 있어 번역에 소요되는 시간 및 리소스가 감소하는 효과가 있다.
- [117] 도 11은 본 발명의 실시 예들에 따른 음성 처리 시스템을 나타낸다. 도 11을 참조하면, 음성 처리 시스템(10A)은 음성 처리 장치(100)와 코드리스 이어폰들(300-1~300-4)을 포함할 수 있다. 도 1와 비교할 때, 도 11의 음성 처리 시스템(10A)은 코드리스 이어폰들(300-1~300-4)을 포함하는 것이 차이가 있다.
- [118] 코드리스 이어폰들(300-1~300-4)은 화자들의 음성을 수신하고, 화자들의 음성과 연관된 음성 신호를 음성 처리 장치(200)로 전송할 수 있다. 또한, 코드리스 이어폰들(300-1~300-4)은 음성 처리 장치(200)로부터 전송된 음성 신호를 재생할 수 있다.

- [119] 코드리스 이어폰들(300-1~300-4)은 무선 방식의 이어폰(earphone)으로서, 음성 처리 장치(200)와 무선 방식으로 연동 또는 연결될 수 있다. 실시 예들에 따라, 코드리스 이어폰들(300-1~300-4)은 Bluetooth, WiFi, ZigBee, RFID, NFC 등과 같은 무선 통신 방식에 따라 음성 처리 장치(200)와 신호를 주고받을 수 있다. 예컨대, 코드리스 이어폰들(300-1~300-4)은 완전 무선(true wireless stereo (TWS)) 이어폰일 수 있다.
- [120] 코드리스 이어폰들(300-1~300-4) 각각은 왼쪽 이어폰 및 오른쪽 이어폰을 포함할 수 있다.
- [121] 실시 예들에 따라, 코드리스 이어폰들(300-1~300-4) 각각은 화자들(SPK1~SPK4)에 의해 착용된 장치일 수 있다. 예컨대, 제1코드리스 이어폰(300-1)은 제1화자(SPK1)에 의해 착용되고, 제2코드리스 이어폰(300-2)은 제2화자(SPK2)에 의해 착용되고, 제3코드리스 이어폰(300-3)은 제3화자(SPK3)에 의해 착용되고, 제4코드리스 이어폰(300-4)은 제4화자(SPK4)에 의해 착용될 수 있다.
- [122] 한편, 도 1에는 코드리스 이어폰들(300-1~300-4)이 두 개의 이어폰 유닛으로 구성된 것으로 도시되어 있으나, 본 발명의 실시 예들이 이에 한정되는 것은 아니다.
- [123] 코드리스 이어폰들(300-1~300-4)은 공간 내에서 발생한 음성에 응답하여, 음성 신호를 생성하며, 생성된 음성 신호를 음성 처리 장치(200)로 전송할 수 있다. 실시 예들에 따라, 코드리스 이어폰들(300-1~300-4)은 화자들(SPK1~SPK4)의 음성에 응답하여, 화자들(SPK1~SPK4)의 음성과 연관된 음성 신호를 생성하고, 음성 신호를 음성 처리 장치(200)로 전송할 수 있다.
- [124] 예컨대, 코드리스 이어폰들(300-1~300-4)은 복수의 마이크로폰(microphone)들을 포함할 수 있다.
- [125] 코드리스 이어폰들(300-1~300-4)은 음성 처리 장치(200)로부터 전송된 음성 신호를 청각적인 방식에 따라 출력할 수 있다. 실시 예들에 따라, 코드리스 이어폰들(300-1~300-4)은 전송된 음성 신호에 해당하는 음성을 재생할 수 있다.
- [126] 음성 처리 장치(200)는 코드리스 이어폰들(300-1~300-4)로부터 화자들(SPK1~SPK4)에 의해 발화된 음성과 연관된 음성 신호를 수신할 수 있다. 음성 신호는 특정 시간동안 발화된 음성들과 연관된 신호로서, 복수의 화자들의 음성을 나타내는 신호일 수 있다.
- [127] 음성 처리 장치(200)는 코드리스 이어폰들(300-1~300-4)으로부터 전송된 음성 신호의 음원 위치를 판단하고, 음원 위치에 기초하여 음원 분리를 수행함으로써, 코드리스 이어폰들(300-1~300-4)으로부터 전송된 음성 신호로부터 화자들(SPK1~SPK4) 각각의 음성과 연관된 분리 음성 신호를 추출(또는 생성)할 수 있다. 즉, 이하 본 명세서에서 설명되는 분리 음성 신호는 도 1 내지 도 10을 참조하여 설명된 출력 음성 데이터에 대응한다.
- [128] 음성 처리 장치(200)는 화자들(SPK1~SPK4) 각각의 음성에 대한 번역을

- 코드리스 이어폰들(300-1~300-4) 각각으로 제공할 수 있다.
- [129] 도 12는 본 발명의 실시 예들에 따른 코드리스 이어폰을 나타낸다. 도 12를 참조하면, 코드리스 이어폰(300)은 도 1에 도시된 코드리스 이어폰들(300-1~300-4)을 대표적으로 나타낸다.
- [130] 코드리스 이어폰(300)은 마이크 유닛(310), 스피커 유닛(320), 제어 유닛(330), 통신 유닛(340) 및 배터리(350)를 포함할 수 있다.
- [131] 마이크 유닛(310)은 음성을 수신하고, 음성에 응답하여 음성 신호를 생성할 수 있다. 실시 예들에 따라, 마이크 유닛(310)은 음성으로 인한 공기의 진동을 검출하고, 검출 결과에 따라 진동에 대응하는 전기적인 신호인 음성 신호를 생성할 수 있다.
- [132] 실시 예들에 따라, 마이크 유닛(310)은 복수의 마이크들을 포함할 수 있고, 복수의 마이크들 각각은 음성에 응답하여 음성 신호를 생성할 수 있다. 예컨대, 코드리스 이어폰(300)은 하나의 음성에 응답하여 복수의 음성 신호를 생성할 수 있다. 이 때, 마이크들이 배치된 위치는 서로 다를 수 있으므로, 마이크들 각각에 의해 생성된 음성 신호들은 서로 위상 차(또는 시간 지연)를 가질 수 있다.
- [133] 스피커 유닛(320)은 음성 신호에 해당하는 음성을 출력할 수 있다. 실시 예들에 따라, 스피커 유닛(320)은 음성 신호에 대응하는 진동을 형성함으로써, 상기 음성 신호와 연관된 음성을 재생할 수 있다.
- [134] 제어 유닛(330)은 코드리스 이어폰(300)의 전반적인 동작을 제어할 수 있다. 실시 예들에 따라, 제어 유닛(330)은 연산 처리 기능을 갖는 프로세서를 포함할 수 있다. 예컨대, 제어 유닛(330)은 CPU(central processing unit), MCU(micro controller unit), DSP(digital signal processor), ADC 컨버터(analog to digital converter) 또는 DAC 컨버터(digital to analog converter)를 포함할 수 있으나, 이에 한정되는 것은 아니다.
- [135] 제어 유닛(330)은 마이크 유닛(310)에 의해 생성된 음성 신호들을 아날로그-디지털 변환할 수 있다. 디지털로 변환된 음성 신호는 통신 유닛(340)을 통해 출력될 수 있다. 또한, 제어 유닛(330)은 통신 유닛(340)에 의해 수신된 디지털 형태의 번역 음성 신호를 디지털-아날로그 변환하고, 아날로그로 변환된 번역 음성 신호를 스피커 유닛(320)으로 전송할 수 있다.
- [136] 통신 유닛(340)은 무선 통신 방식에 따라 음성 처리 장치(200A)와 데이터를 주고받을 수 있다. 실시 예들에 따라, 통신 유닛(340)은 WiFi, ZigBee, RFID, NFC 등과 같은 무선 통신 방식에 따라 음성 처리 장치(200A)와 신호를 주고받을 수 있다.
- [137] 통신 유닛(340)은 음성 처리 장치(200A)로 음성 신호를 전송할 수 있고, 음성 처리 장치(200A)로부터 번역 음성 신호를 수신할 수 있다.
- [138] 배터리(350)는 코드리스 이어폰(300)의 동작에 필요한 전력을 공급할 수 있다. 실시 예들에 따라, 배터리(350)는 코드리스 이어폰(300)에 포함된 마이크 유닛(310), 스피커 유닛(320), 제어 유닛(330) 및 통신 유닛(340)으로 전력을

공급할 수 있다.

- [139] 도 13은 본 발명의 실시 예들에 따른 음성 처리 장치를 나타낸다. 도 13을 참조하여 설명되는 음성 처리 장치(200A)는 도 2를 참조하여 설명된 음성 처리 장치(200A)의 기능을 수행할 수 있다. 이하, 차이점에 대해서만 설명한다.
- [140] 도 13을 참조하면, 음성 처리 장치(200A)는 통신 회로(210A), 음성 처리 회로(220A) 및 메모리(230A)를 포함할 수 있다.
- [141] 통신 회로(210A)는 도 2를 참조하여 설명된 음성 데이터 수신 회로(210) 및 음성 데이터 출력 회로(240)에 대응할 수 있고, 음성 처리 회로(220A)는 도 2를 참조하여 설명된 프로세서(230)에 대응할 수 있고, 메모리(230A)는 도 2를 참조하여 설명된 메모리(220)에 대응할 수 있다.
- [142] 통신 회로(210A)는 코드리스 이어폰들(300-1~300-4)로부터 화자들(SPK1~SPK4)의 음성과 연관된 음성 신호를 수신할 수 있다. 실시 예들에 따라, 통신 회로(210A)는 복수의 통신 모듈들을 포함할 수 있고, 복수의 통신 모듈들 각각은 코드리스 이어폰들(300-1~300-4)와 페어링(pairing)을 수행할 수 있다.
- [143] 통신 회로(210A)는 코드리스 이어폰들(300-1~300-4) 각각의 식별자를 수신할 수 있다. 상기 식별자는 코드리스 이어폰들(300-1~300-4)의 단말 ID 또는 MAC 주소일 수 있으나, 이에 한정되는 것은 아니다.
- [144] 통신 회로(210A)는 코드리스 이어폰들(300-1~300-4)로 번역 결과를 전송할 수 있다. 이에 대해서는 후술한다.
- [145] 음성 처리 회로(220A)는 음성 신호를 처리할 수 있다. 실시 예들에 따라, 음성 처리 회로(220A)는 코드리스 이어폰들(300-1~300-4)로부터 전송된 음성 신호를 이용하여, 화자들(SPK1~SPK4) 각각의 음성과 연관된 분리 음성 신호를 추출(또는 생성)할 수 있다.
- [146] 음성 처리 회로(220A)는 코드리스 이어폰들(300-1~300-4) 각각에 대한 음원의 상대적인 위치를 결정하고, 음원 위치에 기초하여, 화자들(SPK1~SPK4) 각각의 음성과 연관된 분리 음성 신호를 생성할 수 있다. 예컨대, 음성 처리 회로(220A)는 음성들의 음원 위치에 기초하여, 제1화자(SPK1)의 음성과 연관된 제1분리 음성 신호를 생성할 수 있다.
- [147] 또한, 실시 예들에 따라, 음성 처리 회로(220A)는 음원의 위치를 나타내는 음원 위치 정보를 분리 음성 신호와 매칭하여 저장할 수 있다. 음원 위치 정보는 도 1 내지 도 10을 참조하여 설명된 음원 위치 데이터를 의미할 수 있다.
- [148] 음성 처리 회로(220A)는 분리 음성 신호를 이용하여, 화자들(SPK1~SPK4) 각각의 음성에 대한 번역을 수행하고, 번역 결과를 생성할 수 있다. 상기 번역 결과는 도착 언어로 표현된 화자들(SPK1~SPK4) 각각의 음성과 연관된 텍스트 데이터 또는 음성 신호일 수 있다.
- [149] 메모리(230A)는 음성 처리 장치(200A)의 동작에 필요한 데이터를 저장할 수 있다.

- [150] 도 14는 본 발명의 실시 예들에 따른 음성 처리 장치의 동작을 설명하기 위한 도면이다. 도 14를 참조하면, 각 위치(P1~P4)에 위치한 화자들(SPK1~SPK4) 각각이 발화할 수 있다. 예컨대, 제1화자(SPK1)는 음성 “AAA”를 발화할 수 있고, 제2화자(SPK2)는 음성 “BBB”를 발화할 수 있고, 제3화자(SPK3)는 음성 “CCC”를 발화할 수 있고, 제4화자(SPK4)는 음성 “DDD”를 발화할 수 있다.
- [151] 음성 처리 장치(200A)는 화자들(SPK1~SPK4)의 음성과 연관된 음성 신호들(VS1~VS4)을 수신하고, 화자들(SPK1~SPK4) 각각의 음성의 음원 위치를 나타내는 음원 위치 정보를 생성할 수 있다. 예컨대, 음성 처리 장치(200A)는 제1화자(SPK1)의 음성의 제1음원 위치인 “P1”을 나타내는 제1음원 위치 정보를 생성 및 저장할 수 있다.
- [152] 음성 처리 장치(200A)는 음원 위치 정보에 기초하여, 화자들(SPK1~SPK4) 각각의 음성과 연관된 분리 음성 신호를 생성할 수 있다. 예컨대, 음성 처리 장치(200A)는 수신된 음성 신호들(VS1~VS4)로부터 음성 “AAA”, “BBB”, “CCC” 및 “DDD”와 연관된 분리 음성 신호들을 생성할 수 있다.
- [153] 음성 처리 장치(200A)는 화자들(SPK1~SPK4) 각각의 음성과 연관된 분리 음성 신호와, 상기 각 음성을 발화한 화자가 착용한 코드리스 이어폰(300-1~300-4)의 식별자를 매칭하여 저장할 수 있다. 예컨대, 음성 처리 장치(200A)는 제1코드리스 이어폰(300-1)로부터 전송된 제1음성 신호들(VS1)을 이용하여 제1화자(SPK1)의 음성 “AAA”와 연관된 제1분리 음성 신호를 생성하고, 제1분리 음성 신호와 제1코드리스 이어폰(300-1)의 식별자 “EID1”를 매칭하여 저장할 수 있다.
- [154] 이를 통해, 특정 음성이 코드리스 이어폰들(300-1~300-4) 중 어느 코드리스 이어폰을 착용한 화자로부터 발화된 것인지 파악할 수 있게 한다. 예컨대, 제1화자(SPK1)의 음성과 연관된 제1분리 음성 신호와 제1코드리스 이어폰(300-1)의 제1식별자가 매칭되어 저장되므로, 제1화자(SPK1)는 제1코드리스 이어폰(300-1)을 착용했음을 알 수 있다.
- [155] 따라서, 결과적으로, 화자들(SPK1~SPK4) 각각의 음성과 연관된 분리 음성 신호는 코드리스 이어폰들(300-1~300-4)의 식별자에 의해 식별될 수 있다.
- [156] 도 15 내지 도 18은 본 발명의 실시 예들에 따른 음성 처리 장치의 번역 기능을 설명하기 위한 도면이다.
- [157] 도 15를 참조하면, 제1화자(SPK1)는 음성 “AAA”를 한국어(KR)로 발화하고, 제2화자(SPK2)는 음성 “BBB”를 영어(EN)로 발화하고, 제3화자(SPK3)는 음성 “CCC”를 중국어(CN)로 발화하고, 제4화자(SPK4)는 음성 “DDD”를 일본어(JP)로 발화한다.
- [158] 음성 처리 장치(200A)는 화자들(SPK1~SPK4) 각각의 음성에 대한 분리 음성 신호와, 상기 각 음성을 발화한 화자가 착용한 코드리스 이어폰(300-1~300-4)의 식별자를 매칭하여 저장할 수 있다.
- [159] 실시 예들에 따라, 음성 처리 장치(200A)는 화자들(SPK1~SPK4) 각각의 음성의

음원 위치를 나타내는 음원 위치 정보를 생성 및 저장할 수 있다.

- [160] 본 발명의 실시 예들에 따른 음성 처리 장치(200A)는 화자들(SPK1~SPK4) 각각의 음성과 연관된 분리 음성 신호를 이용하여, 화자들(SPK1~SPK4) 각각의 음성의 언어에 대한 출발 언어로부터 도착 언어로의 번역을 제공할 수 있다.
- [161] 실시 예들에 따라, 출발 언어와 도착 언어는 코드리스 이어폰들(300-1~300-4) 별로 결정될 수 있다. 즉, 코드리스 이어폰들(300-1~300-4)의 각각의 착용자의 음성에 대한 출발 언어와 도착 언어가 결정될 수 있다.
- [162] 도 16을 참조하면, 코드리스 이어폰들(300-1~300-4) 각각에 대해서, 출발 언어가 설정될 수 있다. 예컨대, 음성 처리 장치(200A)와 연동 가능한 단말기를 이용하여, 각 코드리스 이어폰들(300-1~300-4)을 착용한 화자의 언어를 번역하기 위한 출발 언어를 설정할 수 있다. 설정된 값은 음성 처리 장치(200A)로 전송될 수 있다. 음성 처리 장치(200A)는 각 코드리스 이어폰들(300-1~300-4)을 착용한 화자의 언어를 번역하기 위한 출발 언어를 나타내는 출발 언어 정보를 저장할 수 있다.
- [163] 단말기는 코드리스 이어폰들(300-1~300-4)들에 대한 출발 언어를 나타내는 출발 언어 정보를 음성 처리 장치(200A)로 전송할 수 있고, 음성 처리 장치는 출발 언어 정보를 코드리스 이어폰들(300-1~300-4)의 식별자와 매칭하여 저장할 수 있다.
- [164] 또한, 실시 예들에 따라, 코드리스 이어폰들(300-1~300-4) 각각에 대한 출발 언어 정보는 음성 처리 장치(200A)에 미리 저장되어 있을 수 있다.
- [165] 도 17을 참조하면, 음성 처리 장치(200A)는 분리 음성 신호들 각각에 대응하는 코드리스 이어폰들(300-1~300-4)의 식별자를 이용하여, 화자들(SPK1~SPK4) 각각의 음성을 번역하기 위한 출발 언어와 도착 언어를 결정하고, 화자들(SPK1~SPK4) 각각의 음성에 대한 번역 결과를 생성하고, 번역 결과를 출력할 수 있다.
- [166] 실시 예들에 따라, 음성 처리 장치(200A)는 코드리스 이어폰들(300-1~300-4)의 식별자를 이용하여, 각 코드리스 이어폰들(300-1~300-4)에 대응하는 출발 언어 정보를 리드함으로써, 화자들(SPK1~SPK4) 각각의 음성을 번역하기 위한 출발 언어를 결정할 수 있다. 예컨대, 음성 처리 장치(200A)는 제1코드리스 이어폰(300-1)의 제1식별자(EID1)를 이용하여, 메모리(230A)로부터 제1식별자(EID1)에 대응하는 제1출발 언어 정보를 리드할 수 있다. 리드된 제1출발 언어 정보는 제1화자(SPK1)(즉, 제1코드리스 이어폰(300-1)의 착용자)의 음성 “AAA”의 출발 언어가 한국어(KR)임을 지시한다.
- [167] 실시 예들에 따라, 음성 처리 장치(200A)는 코드리스 이어폰들(300-1~300-4) 중 번역 대상이 아닌 나머지 코드리스 이어폰들에 대한 출발 언어에 기초하여, 번역 대상의 코드리스 이어폰에 대한 도착 언어를 결정할 수 있다. 예컨대, 음성 처리 장치(200A) 제1코드리스 이어폰(300-1)을 제외한 나머지 코드리스 이어폰들(100-2~100-4)의 출발 언어를 제1코드리스 이어폰(300-1)의 도착

- 언어로서 결정할 수 있다. 즉, 제1도착 언어 정보는 제1화자(SPK1)(즉, 제1코드리스 이어폰(300-1)의 착용자)의 음성 “AAA”의 도착 언어가 나머지 언어들인 영어(EN), 중국어(CN) 및 일본어(JP)임을 지시할 수 있다.
- [168] 즉, 본 발명의 실시 예들에 따른 음성 처리 장치(200A)는 복수의 코드리스 이어폰들(300-1~300-4) 중 제1코드리스 이어폰(300-1)의 착용자(즉, 제1화자(SPK1))의 음성의 언어를, 나머지 코드리스 이어폰들(100-2~100-4)의 착용자들(즉, 제2화자(SPK2) 내지 제4화자(SPK4))의 언어로 번역할 수 있다.
- [169] 음성 처리 장치(200A)는 결정된 출발 언어 및 도착 언어에 기초하여, 화자들(SPK1~SPK4)의 음성들에 대한 번역을 제공할 수 있다. 실시 예들에 따라, 음성 처리 장치(200A)는 화자들(SPK1~SPK4)의 음성들 각각의 번역 결과를 생성할 수 있다.
- [170] 본 명세서에서, 음성 처리 장치(200A)에 의해 출력되는 번역 결과는 도착 언어로 표현된 텍스트 데이터이거나 혹은 도착 언어로 발화된 음성과 연관된 음성 신호일 수 있으나, 이에 한정되는 것은 아니다.
- [171] 본 명세서에서, 음성 처리 장치(200A)가 번역 결과를 생성한다는 것은, 음성 처리 장치(200A)의 음성 처리 회로(220A) 자체의 연산을 통해 언어를 번역함으로써 번역 결과를 생성하는 것뿐만 아니라, 음성 처리 장치(200A)가 번역 기능을 갖는 서버와의 통신을 통해, 상기 서버로부터 번역 결과를 수신함으로써 번역 결과를 생성하는 것을 포함한다.
- [172] 예컨대, 음성 처리 회로(220A)는 메모리(230A)에 저장된 번역 애플리케이션을 실행함으로써, 화자들(SPK1~SPK4) 각각의 음성에 대한 번역 결과를 생성할 수 있다.
- [173] 예컨대, 음성 처리 장치(200A)는 분리 음성 신호, 출발 언어 정보 및 도착 언어 정보를 번역기(translator)로 전송하고, 번역기로부터 분리 음성 신호에 대한 번역 결과를 수신할 수 있다. 번역기는 언어에 대한 번역을 제공하는 환경 또는 시스템을 의미할 수 있다. 실시 예들에 따라, 번역기는 분리 음성 신호, 출발 언어 정보 및 도착 언어 정보를 이용하여, 화자들(SPK1~SPK4) 각각의 음성에 대한 번역 결과를 출력할 수 있다.
- [174] 도 17에 도시된 바와 같이, 예컨대, 음성 처리 장치(200A)는 한국어(KR)로 표현되는 제1화자(SPK1)의 음성 “AAA (KR)”과 연관된 분리 음성 신호를 이용하여, 영어(EN)로 표현되는 제1화자(SPK1)의 음성에 대한 번역 결과 “AAA (EN)”를 생성할 수 있다. 또한, 예컨대, 음성 처리 장치(200A)는 중국어(CN) 및 일본어(JP)로 표현되는 제1화자(SPK1)의 음성에 대한 번역 결과 “AAA (CN)” 및 “AAA (JP)”를 생성할 수 있다.
- [175] 음성 처리 장치(200A)는 화자들(SPK1~SPK4) 각각의 음성에 대한 번역 결과를 코드리스 이어폰들(300-1~300-4)로 전송할 수 있다. 실시 예들에 따라, 음성 처리 장치(200A)는 화자들(SPK1~SPK4) 각각의 음성에 대한 번역 결과를, 번역된 언어(즉, 도착 언어)에 대응하는 코드리스 이어폰 각각으로 전송할 수 있다.

- [176] 실시 예들에 따라, 음성 처리 장치(200A)는 메모리(230A)를 참조하여, 번역 결과의 도착 언어와 동일한 언어를 나타내는 출발 언어 정보와 매칭되는 코드리스 이어폰의 식별자를 리드하고, 리드된 식별자를 이용하여 해당하는 코드리스 이어폰으로 번역 결과를 전송할 수 있다.
- [177] 예컨대, 도 17에 도시된 바와 같이, 음성 처리 장치(200A)는 제1화자(SPK1)의 음성에 대한 번역 결과를 코드리스 이어폰들(100-2~100-4)로 전송할 수 있다.
- [178] 따라서, 본 발명의 실시 예들에 따른 음성 처리 장치(200A)는 화자들(SPK1~SPK4) 각각의 음성을 번역하여 번역 결과를 생성하고, 생성된 번역 결과를 화자들(SPK1~SPK4) 각각이 착용한 코드리스 이어폰들(300-1~300-4)로 전송할 수 있다. 이에 따라, 화자들(SPK1~SPK4)의 언어가 서로 다르더라도, 음성 처리 시스템(10A)을 통해 서로의 언어로 소통할 수 있는 효과가 있다.
- [179] 예컨대, 또한, 도 18에 도시된 바와 같이, 음성 처리 장치(200A)는 제2화자(SPK2)의 음성에 대한 번역 결과를 코드리스 이어폰들(100-1, 100-3 및 100-4)로 전송할 수 있다.
- [180] 도 19는 본 발명의 실시 예들에 따른 음성 처리 장치의 작동을 나타내는 플로우 차트이다. 도 19를 참조하면, 음성 처리 장치(200A)는 코드리스 이어폰들(300-1~300-4)로부터 화자들(SPK1~SPK4)의 음성과 연관된 음성 신호를 수신할 수 있다(S210). 예컨대, 음성 처리 장치(200A)는 제1화자(SPK1)에 의해 착용된 제1코드리스 이어폰(300-1)로부터 화자들(SPK1~SPK4)의 음성과 연관된 음성 신호들(VS1)을 수신할 수 있다.
- [181] 음성 처리 장치(200A)는 코드리스 이어폰들(300-1~300-4)로부터 전송된 음성 신호들로부터 화자들(SPK1~SPK4) 각각의 음성과 연관된 분리 음성 신호를 생성할 수 있다(S220). 실시 예들에 따라, 음성 처리 장치(200A)는 코드리스 이어폰들(300-1~300-4)로부터 전송된 음성 신호들에 대응하는 음성들의 음원 위치에 기초하여, 코드리스 이어폰들(300-1~300-4) 각각을 착용한 화자들(SPK1~SPK4)의 음성과 연관된 분리 음성 신호를 생성할 수 있다.
- [182] 음성 처리 장치(200A)는 화자들(SPK1~SPK4) 각각의 음성의 번역을 위한 출발 언어와 도착 언어를 결정할 수 있다(S230). 실시 예들에 따라, 음성 처리 장치(200A)는 메모리(230A)를 참조하여, 코드리스 이어폰들(300-1~300-4) 각각의 식별자와 매칭되어 저장된 출발 언어 정보를 이용하여 출발 언어를 결정할 수 있고, 결정된 출발 언어에 따라 도착 언어 또한 결정할 수 있다.
- [183] 음성 처리 장치(200A)는 분리 음성 신호를 이용하여, 화자들(SPK1~SPK4) 각각의 음성에 대한 번역 결과를 생성할 수 있다(S240). 실시 예들에 따라, 음성 처리 장치(200A)는 음성 처리 장치(200A) 내에 저장된 자체 번역 알고리즘을 통해 번역 결과를 생성하거나, 또는, 통신 가능한 번역기로 분리 음성 신호, 도착 언어 및 출발 언어 정보를 전송하고, 번역기로부터 번역 결과를 수신할 수 있다.
- [184] 음성 처리 장치(200A)는 생성된 번역 결과를 코드리스

이어폰들(300-1~300-4)로 전송할 수 있다(S250). 실시 예들에 따라, 음성 처리 장치(200A)는 화자들(SPK1~SPK4) 각각의 음성에 대한 번역 결과를, 번역된 언어(즉, 도착 언어)에 대응하는 코드리스 이어폰 각각으로 전송할 수 있다.

- [185] 본 발명의 실시 예들에 따른 음성 처리 시스템은 코드리스 이어폰들(300-1~300-4)을 이용하여 화자들(SPK1~SPK4)의 음성과 연관된 음성 신호를 생성하고, 음성 신호들을 처리함으로써 화자들(SPK1~SPK4) 각각의 음성과 연관된 분리 음성 신호를 생성할 수 있다.
- [186] 또한, 음성 처리 시스템은 분리 음성 신호를 이용하여, 화자들(SPK1~SPK4)의 음성을 번역하고, 번역 결과를 해당하는 코드리스 이어폰들로 출력할 수 있다. 이에 따라, 화자들(SPK1~SPK4)의 사용 언어가 다르더라도, 화자들(SPK1~SPK4) 각각은 자신이 사용하는 언어로 발화할 수 있고, 다른 언어를 사용하는 화자들의 음성을 자신이 사용하는 언어로 번역하여 들을 수 있는 효과가 있다.
- [187] 이상과 같이 실시 예들이 비록 한정된 실시 예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.
- [188] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

산업상 이용가능성

- [189] 본 발명의 실시 예들은 음성을 처리하기 위한 장치 및 이의 작동 방법에 관한 것이다.

청구범위

- [청구항 1] 화자들의 음성과 연관된 입력 음성 데이터를 수신하도록 구성되는 음성 데이터 수신 회로;
출발 언어 데이터를 저장하도록 구성되는 메모리;
상기 화자들의 음성과 연관된 출력 음성 데이터를 출력하도록 구성되는 음성 데이터 출력 회로; 및
상기 출력 음성 데이터를 출력하기 위한 제어 명령을 생성하도록 구성되는 프로세서를 포함하고,
상기 프로세서는,
상기 입력 음성 데이터를 이용하여, 상기 화자들 중 제1화자의 위치를 나타내는 제1화자 위치 데이터 및 상기 제1화자의 음성과 연관된 제1출력 음성 데이터를 생성하고,
상기 메모리를 참조하여 상기 제1화자 위치 데이터에 대응하는 제1출발 언어 데이터를 리드하고,
상기 제1출발 언어 데이터에 의해 지시되는 제1출발 언어를 번역하기 위한 번역 환경으로 상기 제1출력 음성 데이터를 출력하기 위한 제어 명령을 상기 음성 데이터 출력 회로로 전송하는,
음성 처리 장치.
- [청구항 2] 제1항에 있어서,
상기 입력 음성 데이터는 복수의 마이크로폰들에 의해 생성된 음성 신호들로부터 생성된 것인,
음성 처리 장치.
- [청구항 3] 제2항에 있어서, 상기 프로세서는,
상기 복수의 마이크로폰들 사이의 거리 및 상기 음성 신호가 상기 복수의 마이크로폰에 의해 수신되는 시점에 기초하여, 상기 제1화자 위치 데이터를 생성하는,
음성 처리 장치.
- [청구항 4] 제1항에 있어서,
상기 메모리는 상기 출발 언어 데이터와 대응하는 위치 데이터를 상기 출발 언어 데이터와 매칭하여 저장하고,
상기 프로세서는 저장된 위치 데이터 중 상기 제1화자 위치 데이터와 대응하는 제1위치 데이터를 결정하고, 출발 언어 데이터 중 상기 제1위치 데이터와 매칭되어 저장된 상기 제1출발 언어 데이터를 결정하는,
음성 처리 장치.
- [청구항 5] 제1항에 있어서,
상기 프로세서는, 상기 제1화자의 음성과 연관된 상기 제1출력 음성 데이터를 상기 제1출발 언어로 표현되는 텍스트 데이터로 변환하고,

상기 음성 데이터 출력 회로는, 상기 프로세서의 제어에 따라 변환된 텍스트 데이터를 상기 번역 환경으로 전송하는,
음성 처리 장치.

[청구항 6] 제1항에 있어서, 상기 프로세서는,
상기 입력 음성 데이터를 이용하여, 상기 화자들 중 제2화자의 위치를 나타내는 제2화자 위치 데이터를 생성하고,
상기 메모리를 참조하여 상기 제2화자 위치 데이터에 대응하는 제2출발 언어 데이터를 리드하고,
상기 제1출발 언어를 상기 제2출발 언어 데이터에 의해 지시되는 제2출발 언어로 번역하기 위한 번역 환경으로 상기 제1출력 음성 데이터를 출력하기 위한 제어 명령을 상기 음성 데이터 출력 회로로 전송하는,
음성 처리 장치.

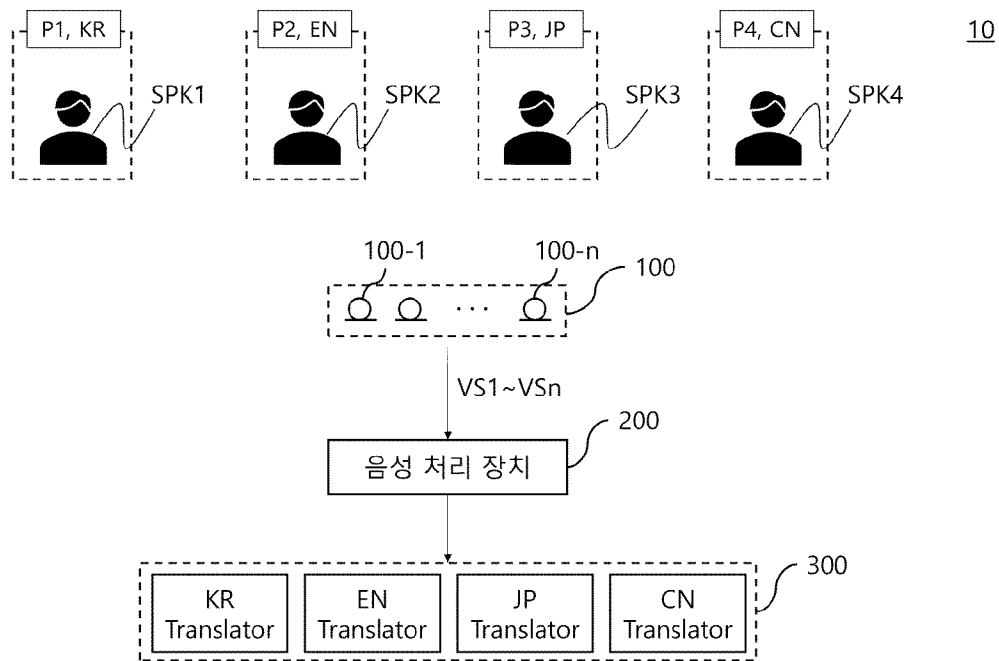
[청구항 7] 제6항에 있어서, 상기 프로세서는,
상기 입력 음성 데이터를 이용하여, 상기 제2화자의 음성과 연관된 제2출력 음성 데이터를 생성하고,
상기 제2출발 언어를 상기 제1출발 언어로 번역하기 위한 번역 환경으로 상기 제1출력 음성 데이터를 출력하기 위한 제어 명령을 상기 음성 데이터 출력 회로로 전송하는,
음성 처리 장치.

[청구항 8] 복수의 코드리스 이어폰들 및 음성 처리 장치를 포함하는 음성 처리 시스템에 있어서,
상기 복수의 코드리스 이어폰들 각각은,
화자들로부터 발화된 음성들과 연관된 음성 신호를 생성하도록 구성되는 마이크 유닛, 음성 신호를 전송하도록 구성되는 통신 유닛 및 음성을 재생하도록 구성되는 스피커 유닛을 포함하고,
상기 음성 처리 장치는,
상기 복수의 코드리스 이어폰들로부터 전송된 음성 신호들을 수신하도록 구성되는 통신 회로;
상기 복수의 코드리스 이어폰들 중 제1코드리스 이어폰으로부터 전송된 제1음성 신호들로부터 상기 화자들 중 제1화자의 음성과 연관된 제1분리 음성 신호를 생성하고, 상기 제1분리 음성 신호를 번역하여 제1번역 음성 신호를 생성하도록 구성되는 음성 처리 회로;
메모리; 및
상기 제1번역 음성 신호를 상기 복수의 코드리스 이어폰들 중 제1코드리스 이어폰을 제외한 나머지 코드리스 이어폰으로 전송하도록 구성되는 통신 회로를 포함하는,
음성 처리 시스템.

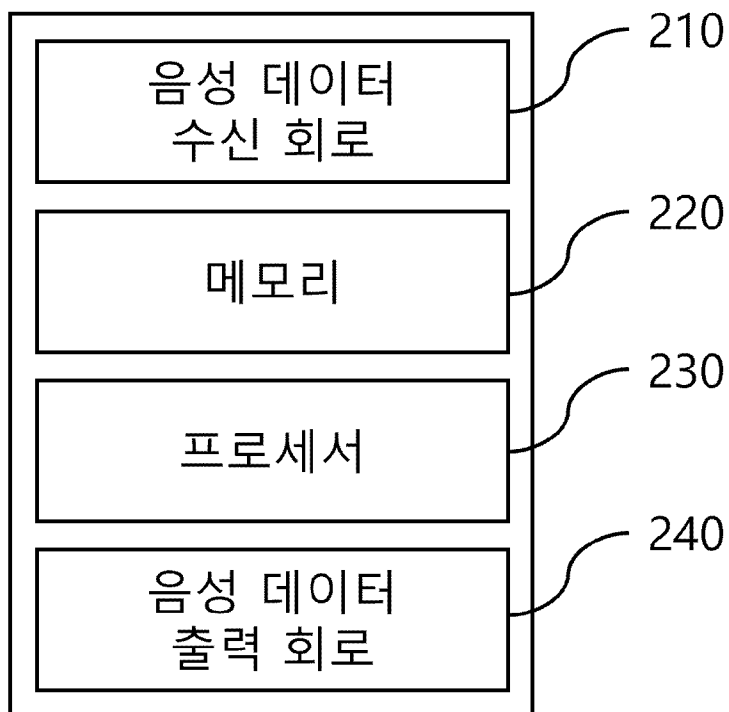
[청구항 9] 제8항에 있어서,

- 상기 제1번역 음성 신호는 상기 제1화자의 음성을 번역한 음성과 연관된 음성 신호인,
음성 처리 시스템.
- [청구항 10] 제8항에 있어서,
상기 복수의 코드리스 이어폰들의 상기 마이크 유닛은 복수의 마이크들을 포함하고,
상기 복수의 마이크들은 상기 화자들의 음성들에 응답하여 상기 음성 신호를 생성하도록 구성되는,
음성 처리 시스템.
- [청구항 11] 제8항에 있어서, 상기 음성 처리 회로는,
상기 제1코드리스 이어폰으로부터 전송된 제1음성 신호들 사이의 시간 지연에 기초하여, 상기 화자들의 음성들 각각의 음원 위치를 판단하고, 판단된 음원 위치에 기초하여 상기 제1화자의 음성과 연관된 상기 제1분리 음성 신호를 생성하는,
음성 처리 시스템.
- [청구항 12] 제11항에 있어서,
상기 제1분리 음성 신호는 상기 화자들의 음성들 중에서 상기 제1코드리스 이어폰과 가장 가까운 위치의 음원 위치를 갖는 음성과 연관된 신호인,
음성 처리 시스템.
- [청구항 13] 제8항에 있어서,
상기 통신 회로는 상기 복수의 코드리스 이어폰들로부터 상기 복수의 코드리스 이어폰들 각각의 식별자들을 수신하고,
상기 음성 처리 회로는 상기 복수의 코드리스 이어폰들의 식별자들과 상기 코드리스 이어폰들을 착용한 화자의 음성의 언어를 나타내는 출발 언어 정보를 서로 매칭하여 상기 메모리에 저장하는,
음성 처리 시스템.
- [청구항 14] 제13항에 있어서, 상기 음성 처리 회로는,
상기 식별자들 및 상기 출발 언어 정보를 이용하여, 상기 복수의 코드리스 이어폰들 중 제2코드리스 이어폰을 착용한 화자의 음성의 언어로 상기 제1분리 음성 신호의 언어를 번역하고, 상기 제1번역 음성 신호를 생성하는,
음성 처리 시스템.
- [청구항 15] 제14항에 있어서, 상기 음성 처리 회로는,
상기 제1번역 음성 신호를 상기 제2코드리스 이어폰으로 전송하는,
음성 처리 시스템.

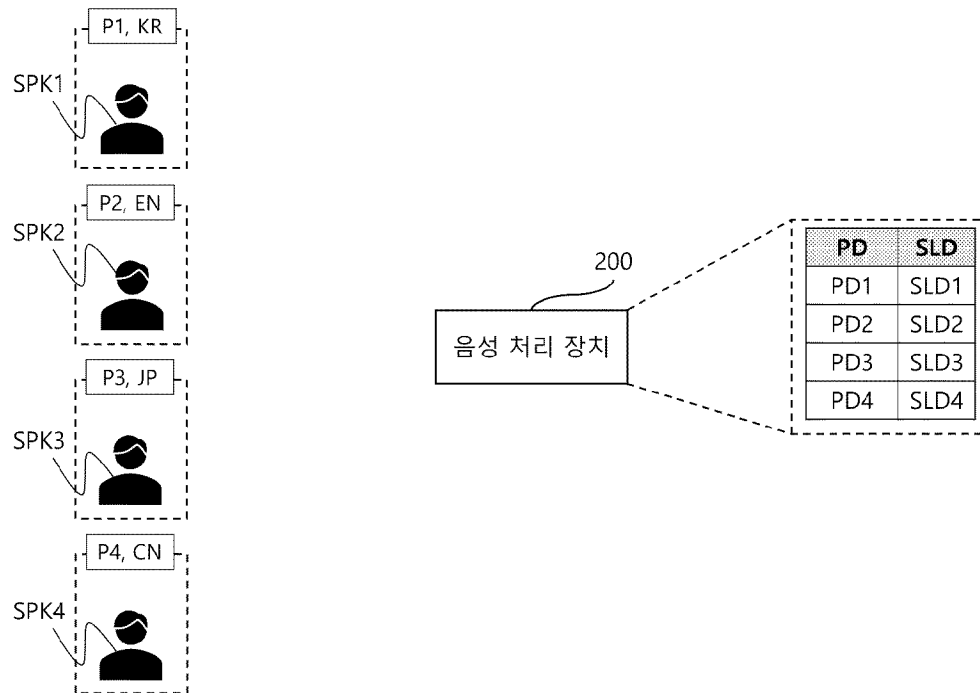
[도1]



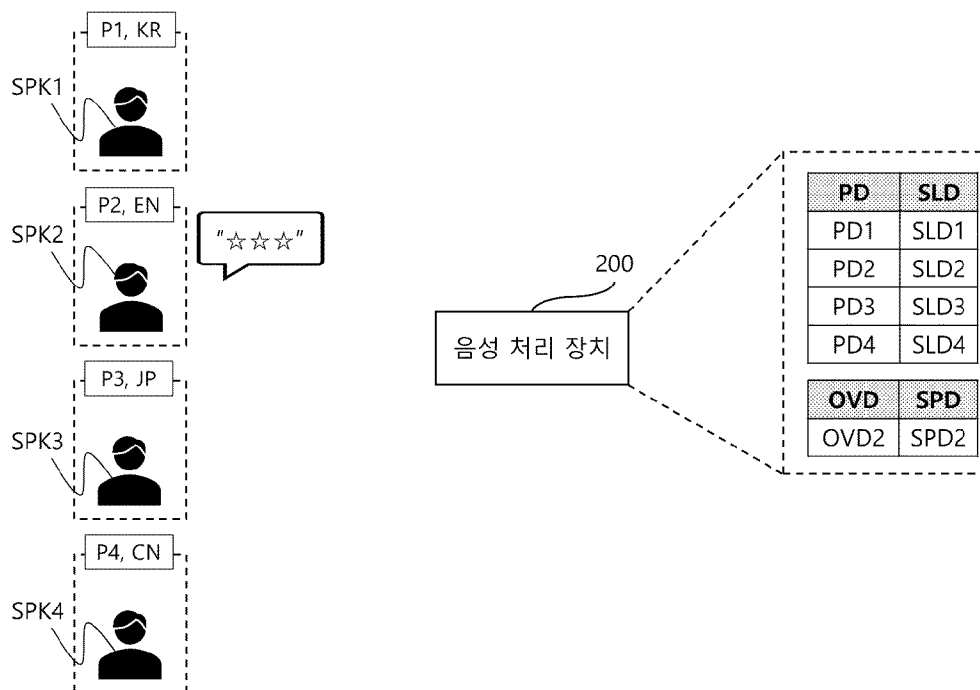
[도2]

200

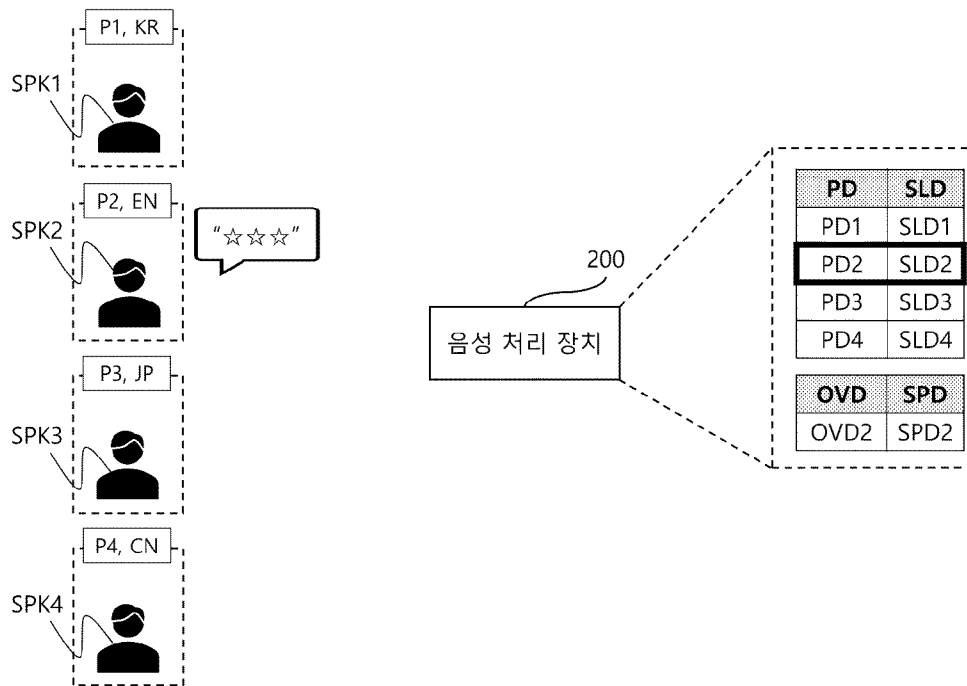
[도3]



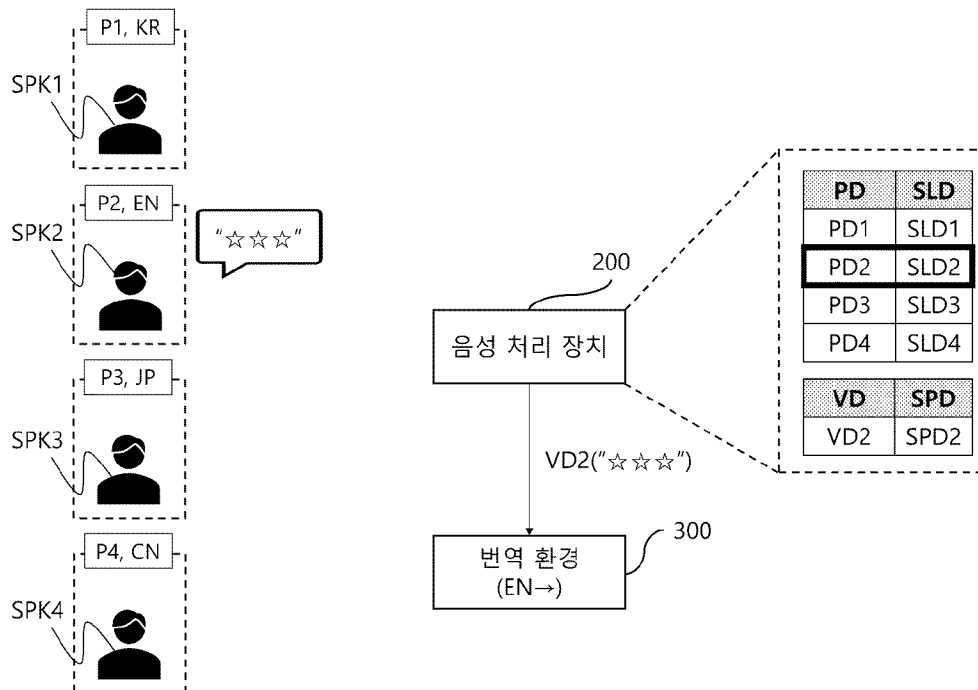
[도4]



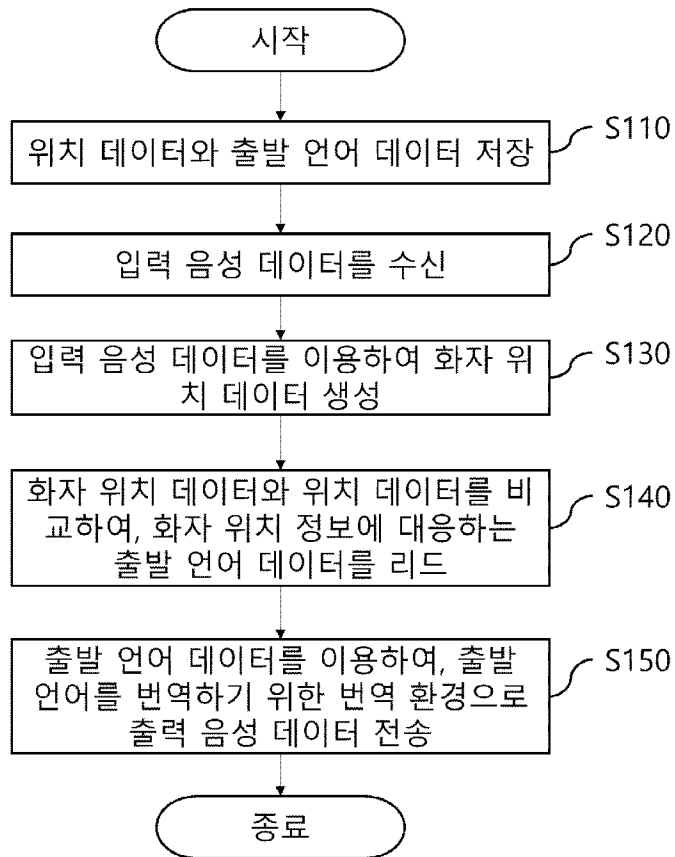
[도5]



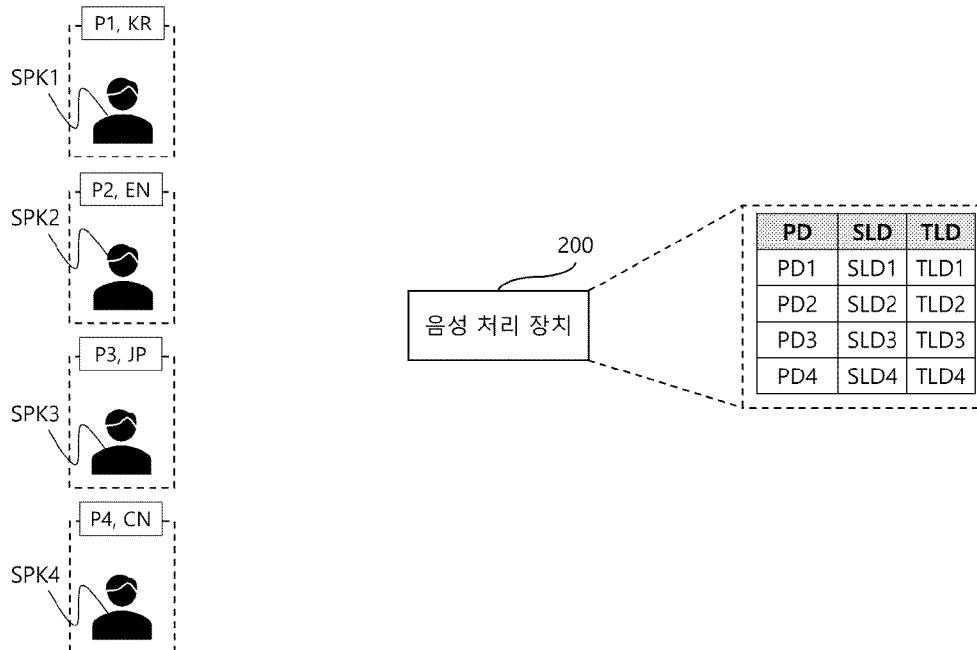
[도6]



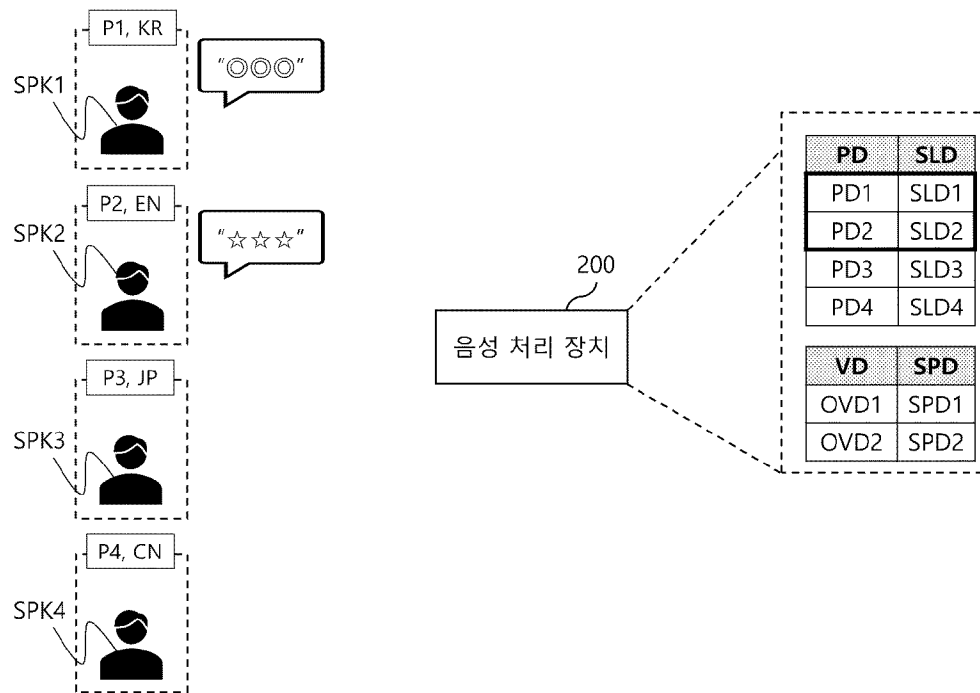
[도7]



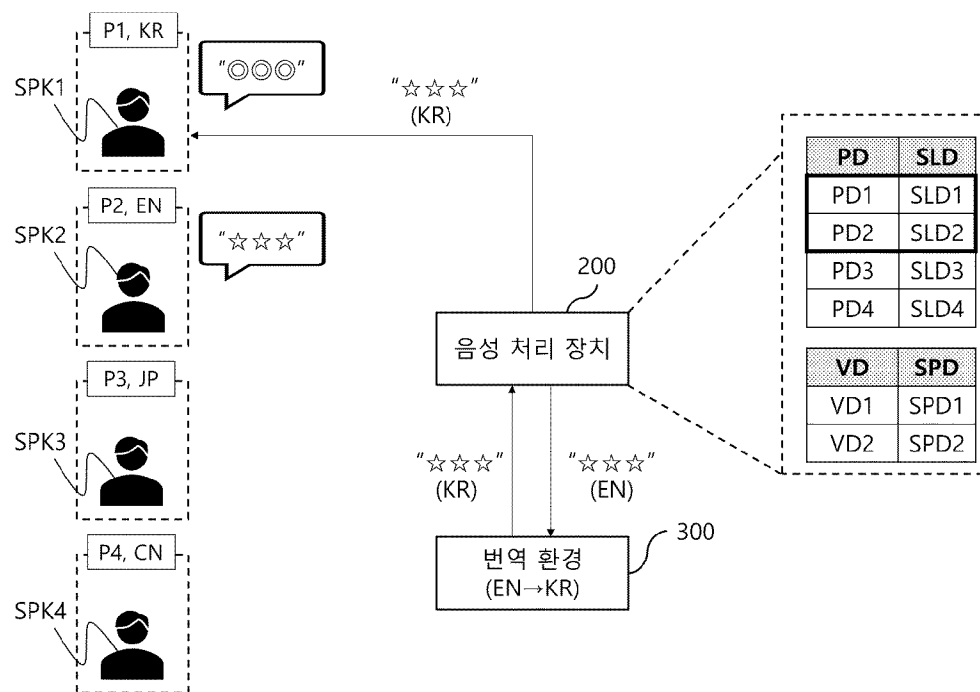
[도8]



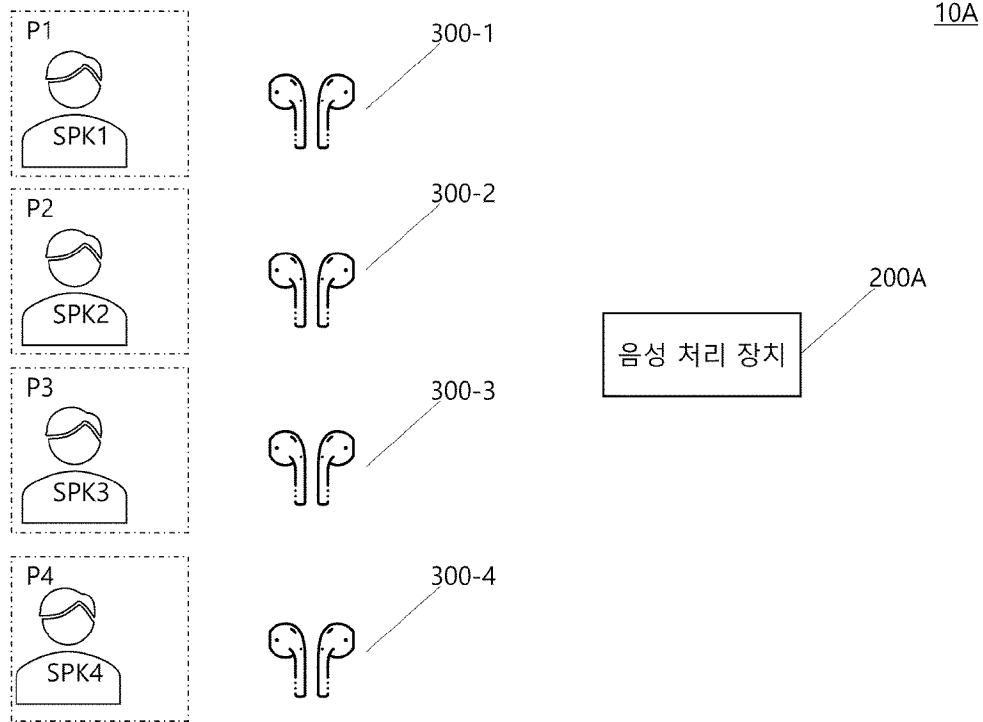
[도9]



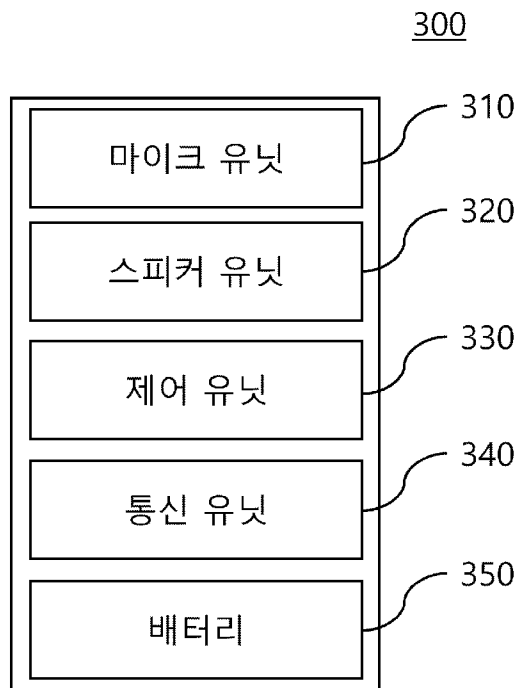
[도10]



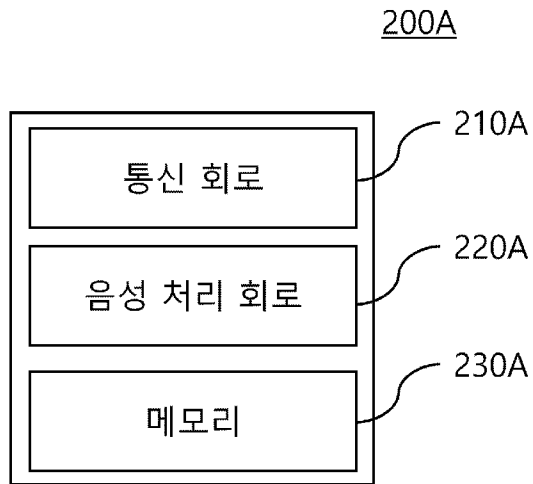
[도11]



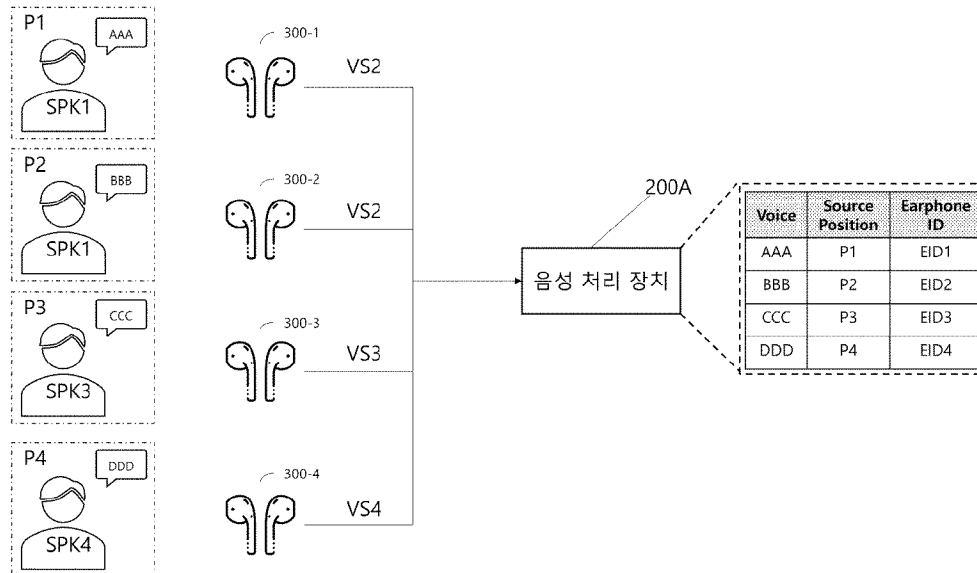
[도12]



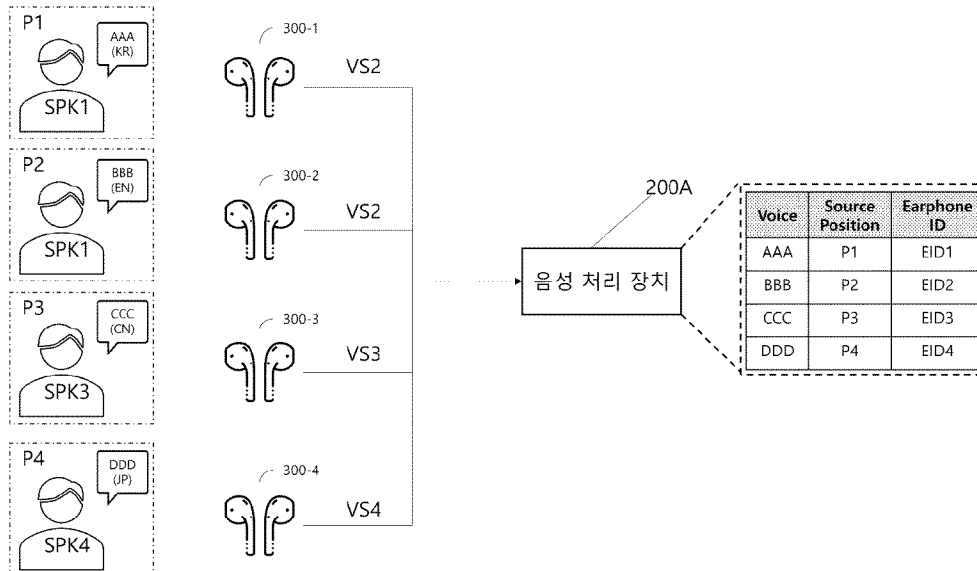
[도13]



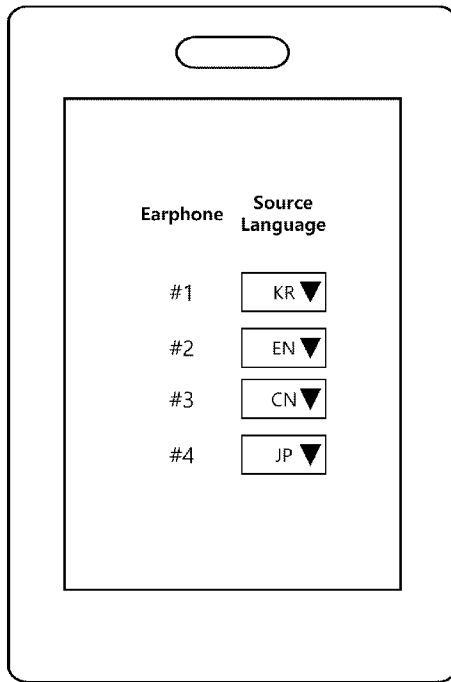
[도14]



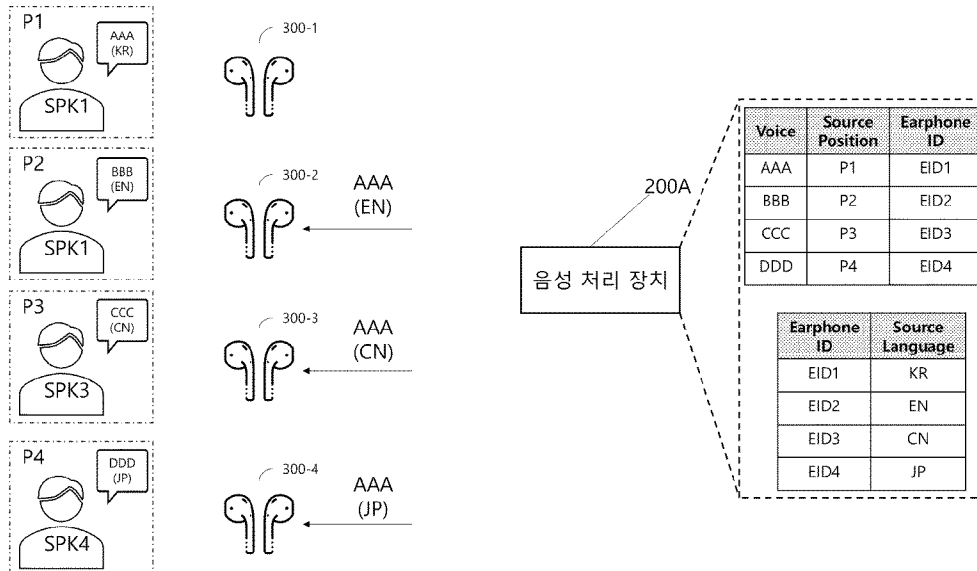
[도15]



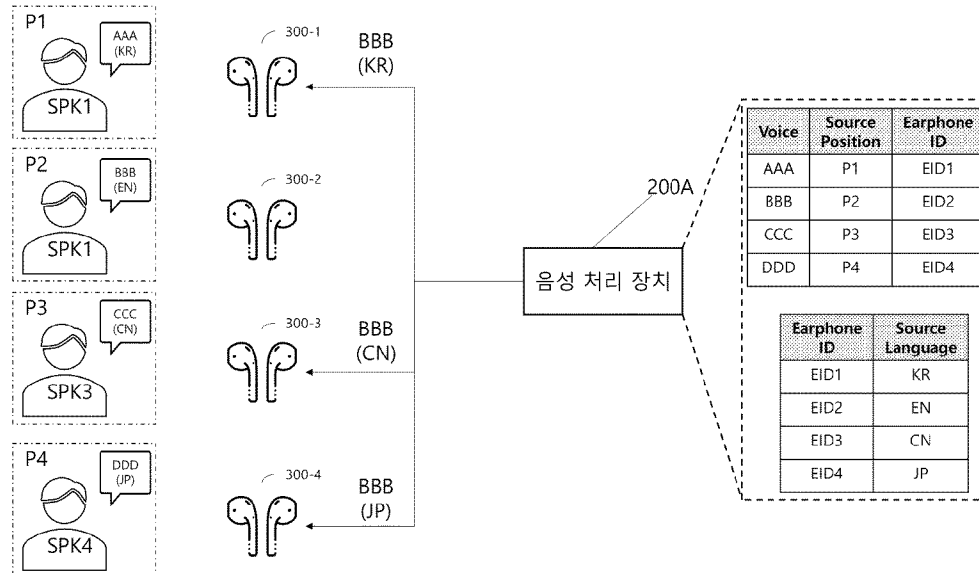
[도16]



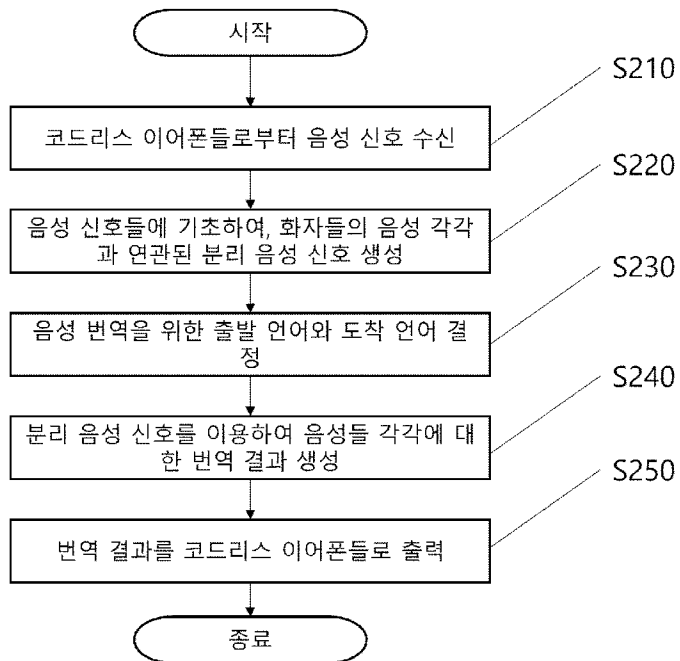
[도17]



[도18]



[도19]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2021/010939

A. CLASSIFICATION OF SUBJECT MATTER

H04R 3/00(2006.01)i; **G10L 17/02**(2013.01)i; **G06F 40/40**(2020.01)i; **G01S 3/808**(2006.01)i; **G10L 15/00**(2006.01)i;
G10L 17/22(2013.01)i; **H04W 4/02**(2009.01)i; **G06F 40/58**(2020.01)i; **H04R 1/40**(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04R 3/00(2006.01); G10L 13/00(2006.01); G10L 15/08(2006.01); G10L 15/18(2006.01); G10L 15/22(2006.01);
 G10L 15/26(2006.01); G10L 21/0208(2013.01); G10L 21/0216(2013.01); G10L 25/90(2013.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models: IPC as above
 Japanese utility models and applications for utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & keywords: 번역(translation), 화자(speaker), 위치(position), 언어(language), 마이크로폰(microphone)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2019-174786 A (PANASONIC CORP.) 10 October 2019 (2019-10-10) See paragraphs [0010], [0031], [0036], [0056], [0073], [0095], [0096], [0149], [0159] and [0161].	1-7
X	KR 10-1989127 B1 (NAVER LABS CORPORATION) 30 September 2019 (2019-09-30) See paragraphs [0005]-[0023] and claims 1-16.	1
X	KR 10-2019-0005617 A (YOON, Seong Yong) 16 January 2019 (2019-01-16) See paragraphs [0021]-[0029] and claims 1-4.	1
A	US 2015-0154957 A1 (HONDA MOTOR CO., LTD.) 04 June 2015 (2015-06-04) See paragraphs [0006]-[0015] and [0132]-[0140].	1-7
A	WO 2019-208860 A1 (SYSTRAN INTERNATIONAL) 31 October 2019 (2019-10-31) See paragraphs [0064]-[0086].	1-7



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
 “D” document cited by the applicant in the international application
 “E” earlier application or patent but published on or after the international filing date
 “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 “O” document referring to an oral disclosure, use, exhibition or other means
 “P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

20 December 2021

Date of mailing of the international search report

21 December 2021

Name and mailing address of the ISA/KR

**Korean Intellectual Property Office
 Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208**

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

The invention of group 1: claims 1-7 pertain to a voice processing device, which uses inputted voice data to translate corresponding voice data on the basis of the position of a speaker,

The invention of group 2: claims 8-15 pertain to a voice processing system which uses a wireless earphone to isolate the voice of a speaker from among compounded voice signals and translate same in a multi-party translation system.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: **claims 1-7**

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2021/010939

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
JP	2019-174786	A	10 October 2019	CN	110322881	A	11 October 2019
				US	1182567	B2	23 November 2021
				US	2019-0303443	A1	03 October 2019
KR	10-1989127	B1	30 September 2019	KR	10-2018-0131155	A	10 December 2018
KR	10-2019-0005617	A	16 January 2019	KR	10-1983310	B1	29 May 2019
US	2015-0154957	A1	04 June 2015	JP	2015-106014	A	08 June 2015
				JP	6148163	B2	14 June 2017
				US	9691387	B2	27 June 2017
WO	2019-208860	A1	31 October 2019	CN	112055876	A	08 December 2020
				EP	3779971	A1	17 February 2021
				JP	2021-529337	A	28 October 2021
				KR	10-2020-0125735	A	04 November 2020
				US	2021-0232776	A1	29 July 2021

A. 발명이 속하는 기술분류(국제특허분류(IPC)) H04R 3/00(2006.01)i; G10L 17/02(2013.01)i; G06F 40/40(2020.01)i; G01S 3/808(2006.01)i; G10L 15/00(2006.01)i; G10L 17/22(2013.01)i; H04W 4/02(2009.01)i; G06F 40/58(2020.01)i; H04R 1/40(2006.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) H04R 3/00(2006.01); G10L 13/00(2006.01); G10L 15/08(2006.01); G10L 15/18(2006.01); G10L 15/22(2006.01); G10L 15/26(2006.01); G10L 21/0208(2013.01); G10L 21/0216(2013.01); G10L 25/90(2013.01) 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 번역(translation), 화자(speaker), 위치(position), 언어(language), 마이크 로폰(microphone)		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X	JP 2019-174786 A (PANASONIC CORP.) 2019.10.10 단락 [0010], [0031], [0036], [0056], [0073], [0095], [0096], [0149], [0159], [0161]	1-7
X	KR 10-1989127 B1 (네이버랩스 주식회사) 2019.09.30 단락 [0005]-[0023] 및 청구항 1-16	1
X	KR 10-2019-0005617 A (윤성용) 2019.01.16 단락 [0021]-[0029] 및 청구항 1-4	1
A	US 2015-0154957 A1 (HONDA MOTOR CO., LTD.) 2015.06.04 단락 [0006]-[0015], [0132]-[0140]	1-7
A	WO 2019-208860 A1 (SYSTRAN INTERNATIONAL) 2019.10.31 단락 [0064]-[0086]	1-7
☐ 추가 문헌이 C(계속)에 기재되어 있습니다. ☒ 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “D” 본 국제출원에서 출원인이 인용한 문헌 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다. “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. “&” 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일 2021년12월20일(20.12.2021)		국제조사보고서 발송일 2021년12월21일(21.12.2021)
ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578		심사관 박혜련 전화번호 +82-42-481-3463

제3기재란 발명의 단일성이 결여된 경우의 의견(첫 번째 용지의 3의 계속)

본 국제조사기관은 본 국제출원에 다음과 같이 다수의 발명이 있다고 봅니다.

제1군 발명: 청구항 1-7은 입력받은 음성 데이터를 이용해 화자의 위치를 기반으로 해당 음성 데이터를 번역하는 음성 처리 장치에 관한 것이고,

제2군 발명: 청구항 8-15는 무선 이어폰을 이용하여 다자간 번역 시스템에서 복합 음성 신호 중 화자의 음성을 분리하여 번역하는 음성 처리 시스템에 관한 것입니다.

1. ☐ 출원인이 모든 추가수수료를 기간 내에 납부하였으므로, 본 국제조사보고서는 모든 조사 가능한 청구항을 대상으로 합니다.
2. ☐ 추가수수료 납부를 요구하지 않고도 모든 조사 가능한 청구항을 조사할 수 있었으므로, 본 기관은 추가수수료 납부를 요구하지 아니하였습니다.
3. ☐ 출원인이 추가수수료의 일부만을 기간 내에 납부하였으므로, 본 국제조사보고서는 수수료가 납부된 청구항만을 대상으로 합니다. 구체적인 청구항은 아래와 같습니다.
4. ☒ 출원인이 기간 내에 추가수수료를 납부하지 아니하였습니다. 따라서 본 국제조사보고서는 청구범위에 처음 기재된 발명에 한정되어 있으며, 해당 청구항은 아래와 같습니다. **청구항 1-7**

- 이의신청에 관한 기재 ☐ 출원인의 이의신청 및 이의신청료 납부(해당하는 경우)와 함께 추가수수료가 납부되었습니다.
- ☐ 출원인의 이의신청과 함께 추가수수료가 납부되었으나 이의신청료가 보정요구서에 명시된 기간 내에 납부되지 아니하였습니다.
- ☐ 이의신청 없이 추가수수료가 납부되었습니다.

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
JP 2019-174786 A	2019/10/10	CN 110322881 A	2019/10/11
		US 1182567 B2	2021/11/23
		US 2019-0303443 A1	2019/10/03
KR 10-1989127 B1	2019/09/30	KR 10-2018-0131155 A	2018/12/10
KR 10-2019-0005617 A	2019/01/16	KR 10-1983310 B1	2019/05/29
US 2015-0154957 A1	2015/06/04	JP 2015-106014 A	2015/06/08
		JP 6148163 B2	2017/06/14
		US 9691387 B2	2017/06/27
WO 2019-208860 A1	2019/10/31	CN 112055876 A	2020/12/08
		EP 3779971 A1	2021/02/17
		JP 2021-529337 A	2021/10/28
		KR 10-2020-0125735 A	2020/11/04
		US 2021-0232776 A1	2021/07/29