



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2023년07월19일  
(11) 등록번호 10-2557092  
(24) 등록일자 2023년07월14일

(51) 국제특허분류(Int. Cl.)  
G06F 40/40 (2020.01) G06F 3/14 (2006.01)  
G06F 3/16 (2018.01) G10L 15/00 (2006.01)  
G10L 15/04 (2006.01) G10L 15/22 (2006.01)  
G10L 15/26 (2006.01) G10L 21/0208 (2013.01)  
H04R 1/32 (2006.01)

(52) CPC특허분류  
G06F 40/40 (2020.01)  
G06F 3/14 (2020.08)  
(21) 출원번호 10-2022-0150167  
(22) 출원일자 2022년11월11일  
심사청구일자 2022년11월11일

(56) 선행기술조사문헌  
JP2006023953 A\*  
(뒷면에 계속)

(73) 특허권자  
주식회사 디앤티랩스  
서울특별시 용산구 한강대로40길 18, 4층 404호  
(한강로2가)

(72) 발명자  
한범수  
서울특별시 성북구 북악산로 713 산장빌라, 5동  
301호

김석찬  
경기도 동두천시 동광로53번길 63, 가동 202호  
(뒷면에 계속)

(74) 대리인  
특허법인해안

전체 청구항 수 : 총 4 항

심사관 : 김영신

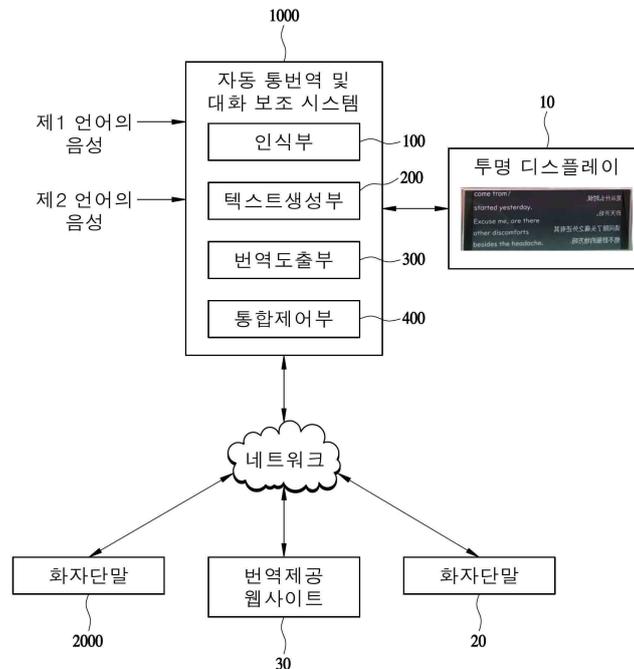
(54) 발명의 명칭 투명 디스플레이를 이용한 자동 통번역 및 대화 보조 시스템

(57) 요약

본 발명은 자동 통번역 및 대화 보조 시스템에 관한 것으로서, 양방향으로 인식된 대화를 통역 및 번역하여 기설정된 한쌍의 언어로 구성된 대화텍스트를 제공할 수 있는 투명 디스플레이를 이용한 자동 통번역 및 대화 보조 시스템에 관한 것이다. 이를 위해, 자동 통번역 및 대화 보조 시스템은 투명 디스플레이를 이용한 자동 통번역

(뒷면에 계속)

대표도 - 도1a



및 대화 보조 시스템은 일방향에서 발화되는 제1 언어의 음성과 타방향에서 발화되는 제2 언어의 음성을 화자별로 구분하여 개별적으로 인식하는 인식부, 상기 제1 및 제2 언어의 음성을 딥러닝의 기반의 인공지능을 통해 분석함에 따라 시간별 문장텍스트를 생성하고 이를 토대로 대화창 리스트를 생성하는 텍스트생성부, 기설정된 규칙 기반, 패턴기반, 통계기반 및 인공지능 기반 중 어느 하나의 번역기가 임베딩되고, 이를 이용하여 상기 대화창 리스트를 기설정된 언어들로 개별적으로 번역함에 따라 제1 및 제2 번역텍스트를 도출하는 번역도출부 및 상기 대화창 리스트로부터 확인되는 시간 및 화자정보에 기초하여, 상기 제1 및 제2 번역텍스트로부터 한쌍의 출력용 자막을 순차적으로 선택하여 투명 디스플레이를 통해 서로 다른 출력방향으로 출력시키는 통합제어부를 포함한다.

(52) CPC특허분류

- G06F 3/16 (2019.02)
- G10L 15/005 (2013.01)
- G10L 15/04 (2013.01)
- G10L 15/22 (2013.01)
- G10L 15/26 (2013.01)
- G10L 17/02 (2013.01)
- G10L 21/0208 (2013.01)
- H04R 1/326 (2013.01)
- Y10S 715/978 (2013.01)

(56) 선행기술조사문헌

- JP2017054065 A\*
- KR1020120072244 A\*
- KR1020170112713 A\*
- KR1020190059381 A\*
- KR1020200032441 A\*
- KR1020220070709 A\*
- \*는 심사관에 의하여 인용된 문헌

(72) 발명자

**최은수**

서울특별시 송파구 올림픽로 135, 255동 2401호 (잠실동, 리센츠)

**김운**

서울특별시 서대문구 북아현로22다길 17, 101동 202호 (북아현동, 에코타운)

## 명세서

### 청구범위

#### 청구항 1

일방향에서 발화되는 제1 언어의 음성과 타방향에서 발화되는 제2 언어의 음성을 화자별로 구분하여 개별적으로 인식하는 인식부;

상기 제1 및 제2 언어의 음성을 딥러닝의 기반의 인공지능을 통해 분석함에 따라 시간별 문장텍스트를 생성하고 이를 토대로 대화창 리스트를 생성하는 텍스트생성부;

기설정된 규칙기반, 패턴기반, 통계기반 및 인공지능 기반 중 어느 하나의 번역기가 임베딩되고, 이를 이용하여 상기 대화창 리스트를 기설정된 언어들로 개별적으로 번역함에 따라 제1 및 제2 번역텍스트를 도출하는 번역도출부; 및

상기 대화창 리스트로부터 확인되는 시간 및 화자정보에 기초하여, 상기 제1 및 제2 번역텍스트로부터 한쌍의 출력용자막을 순차적으로 선택하여 투명 디스플레이를 통해 서로 다른 출력방향으로 출력시키는 통합제어부를 포함하고,

상기 한쌍의 출력용자막은 서로 의미가 동일하게 번역되고 상기 제1 언어로 표시된 제1 출력용자막과 상기 제2 언어로 표시된 제2 출력용자막을 포함하며,

상기 인식부는 특정방향에서 발화되는 발화음성을 선택적으로 인식하기 위한 적어도 하나의 지향성마이크;

상기 발화음성을 대신할 입력텍스트를 입력받기 위한 적어도 하나의 입력장치;

상기 특정방향을 향해 촬영하여 한쌍의 화자영상을 획득하는 적어도 하나의 카메라;

상기 한쌍의 화자영상 중 어느 하나로부터 기설정된 표정 및 제스처 객체를 검출하는 검출모듈;

상기 제스처 객체에 대응되는 수화용 텍스트와 상기 기설정된 표정에 따라 긴급안내 텍스트를 메모리로부터 추출하는 텍스트추출모듈; 및

상기 발화음성, 상기 입력텍스트 및 상기 수화용 텍스트에 발화시간과 화자방향을 표시한 식별코드를 부여하여 상기 메모리에 저장하는 식별모듈을 포함하고,

상기 통합제어부는 상기 인식부를 통해 인식된 상기 제1 및 제2 언어의 각 음성으로부터 분석되는 음성특성과 기설정된 기준샘플을 비교하는 비교부; 및

상기 비교부를 통해 확인된 비교 결과에 기초하여, 상기 제1 및 제2 언어의 각 음성을 기준샘플의 음성특징에 따라 보정하는 음성보정부를 포함하고,

상기 음성보정부는 상기 제1 및 제2 언어의 음성신호를 주파수 분석함에 따라 획득된 오디오 시각정보를 상기 한쌍의 출력용자막에 통합하여 상기 투명 디스플레이를 통해 출력하며,

상기 오디오 시각정보는 화자음성의 최적 볼륨 크기로 발화되도록 유도하기 위하여, 소리의 성분과 강도를 주파수 분석함에 따라 시각화되는 오디오 스펙트럼이고,

상기 음성보정부는 음성의 가청 주파수 이외의 대역의 볼륨크기에 기초하여, 상기 제1 및 제2 언어의 음성신호에 대한 기설정된 가청주파수 대역을 증폭시키도록 제어하고,

상기 통합제어부는 상기 투명 디스플레이에 배치된 거리측정센서를 통해 측정된 화자별 거리정보에 기초하여, 상기 투명 디스플레이에 대한 상기 한쌍의 출력용자막의 출력면적을 서로 다른 면적비율로 조절하며,

상기 통합제어부는 상기 한쌍의 출력용자막을 출력시키기 이전에 간이 시력 검사표를 출력함에 따라 피드백받는 응답신호에 기초하여 화자시력을 추정하고, 이를 기초로 한쌍의 출력용자막에 대한 각 글씨체 확대비율을 자동으로 조절하고,

상기 통합제어부는 상기 투명 디스플레이에 배치된 아이트래커 모듈을 통해 측정된 안구운동 데이터에

기초하여, 상기 한쌍의 출력용자막 이후에 출력되는 다음자막의 페이지 전환속도를 조절하며,

상기 통합제어부는 상기 적어도 하나의 입력장치를 통해 입력받는 SNS 계정정보를 등록함에 따라, 네트워크를 통해 연결된 통합관리서버를 이용하여 SNS 계정정보에 대응되는 화자단말에 상기 대화창 리스트를 이미지 파일로 전송하며,

상기 통합제어부는 상기 투명 디스플레이에 배치된 환경감지센서들을 통해 감지된 재난환경상태에 기초하여, 상기 투명 디스플레이를 통해 재난경보를 출력하고,

상기 환경감지센서들은 연기 감지센서, 일산화탄소 감지센서, 온도 감지센서를 포함하는, 투명 디스플레이를 이용한 자동 통번역 및 대화 보조 시스템.

**청구항 2**

제1항에 있어서,

상기 번역기는 복수의 분야별 맞춤형 번역기들을 포함하고,

상기 복수의 분야별 맞춤형 번역기들은 직업, 연령, 지역, 관심분야, 종교에 따라 불필요 단어들 제거된 번역기들인, 투명 디스플레이를 이용한 자동 통번역 및 대화 보조 시스템.

**청구항 3**

삭제

**청구항 4**

제1항에 있어서,

상기 텍스트생성부는 상기 메모리에 기수집된 어휘단어사전을 학습하여 딥러닝 기반의 음성인식모델을 모델링하는 제1 모델링부;

상기 메모리에 기수집된 발음규칙을 학습하여 규칙 및 통계 기반의 음향모델을 모델링하는 제2 모델링부;

상기 제1 및 제2 언어의 음성을 음성인식 특징벡터에 따라 입력신호로 변환처리하는 신호처리부; 및

상기 입력신호를 상기 음성인식모델과 상기 음향모델에 적용함에 따라 문자정보인 화자별 텍스트로 출력하고, 상기 식별코드에 따라 상기 입력텍스트 및 상기 수화용 텍스트를 화자별 텍스트에 병합하는 디코더를 포함하는, 투명 디스플레이를 이용한 자동 통번역 및 대화 보조 시스템.

**청구항 5**

삭제

**청구항 6**

제1항에 있어서,

상기 음성보정부는 상기 인식부를 통해 인식되는 묵음신호로부터 확인된 소음신호에 따라 노이즈 캔슬링 신호를 생성하고, 이를 토대로 상기 제1 및 제2 언어의 음성신호에 대한 노이즈를 필터링하는 필터링부; 및

상기 묵음신호가 일정시간 동안 지속되는 경우, 상기 투명 디스플레이를 통해 상기 한쌍의 출력용자막에 대한 출력을 일시중지시키고 분위기 전환을 위한 힐링콘텐츠 영상을 대체하여 출력시키는 콘텐츠제공부를 포함하고,

상기 묵음신호는 상기 제1 및 제2 언어의 음성에 해당하는 주파수대역의 불륨이 기설정된 시간동안 인식받지 않는 동안에 상기 인식부를 통해 인식된 특정주파수 대역의 반복적인 소음신호를 포함하는, 투명 디스플레이를 이용한 자동 통번역 및 대화 보조 시스템.

**청구항 7**

삭제

**청구항 8**

삭제

**청구항 9**

삭제

**발명의 설명**

**기술 분야**

[0001] 본 발명은 자동 통번역 및 대화 보조 시스템에 관한 것으로서, 양방향으로 인식된 대화를 통역 및 번역하여 기 설정된 한쌍의 언어로 구성된 대화텍스트를 제공할 수 있는 투명 디스플레이를 이용한 자동 통번역 및 대화 보조 시스템에 관한 것이다.

**배경 기술**

[0002] 세계적으로 소통의 기회가 넓어진만큼 번역의 필요성이 날이 갈수록 중요해지고 있다.

[0003] 특히, 팬데믹 이후 침체되었던 외국인 관광 및 이주에 따라 각국의 관공서나 은행, 호텔 등 기관에서의 외국인 응대가 많아짐에 따라 소통의 어려움을 호소하는 경우가 많다.

[0004] 현재 출시되는 자동 통역 및 번역장치는 외국인과 대화를 할 경우, 자국어를 외국어로 번역하는 장치와 외국어를 자국어로 번역해 주는 장치가 독립적으로 작동하여 외국인의 대화를 자국어로, 자신의 대화를 외국어로 독립적으로 번역을 하여 상대방의 대화 내용에는 상관없이 번역 대상 문장을 독립적으로 번역을 한다.

[0005] 이에 따라, 본 발명에서는 오프라인에서 양방향으로 동시에 소통이 가능한 투명 디스플레이를 이용한 자동 통번역 및 대화 보조 시스템을 제공하고자 한다.

**선행기술문헌**

**특허문헌**

[0006] (특허문헌 0001) (0001) 한국공개특허 10-2002-0076044호(2002.10.09)

**발명의 내용**

**해결하려는 과제**

[0007] 본 발명은 상기와 같은 문제점을 해결하기 위한 것으로서, 본 발명의 목적은 양방향으로 인식된 대화를 통역 및 번역하여 기설정된 한쌍의 언어로 구성된 대화텍스트를 실시간 출력할 수 있는 자동 통번역 및 대화 보조 시스템을 제공하기 위한 것이다.

[0008] 또한, 화자의 투명 디스플레이를 통해 한쌍의 출력용자막 이후에 출력되는 다음자막의 페이지 전환속도를 조절할 수 있는 자동 통번역 및 대화 보조 시스템을 제공하기 위한 것이다.

[0009] 본 발명의 상기 및 다른 목적과 이점은 바람직한 실시예를 설명한 하기의 설명으로부터 분명해질 것이다.

**과제의 해결 수단**

[0010] 상기와 같은 목적을 달성하기 위한 본 발명의 일실시예에 따른 투명 디스플레이를 이용한 자동 통번역 및 대화 보조 시스템은 일방향에서 발화되는 제1 언어의 음성과 타방향에서 발화되는 제2 언어의 음성을 화자별로 구분하여 개별적으로 인식하는 인식부, 상기 제1 및 제2 언어의 음성을 딥러닝의 기반의 인공지능을 통해 분석함에 따라 시간별 문장텍스트를 생성하고 이를 토대로 대화창 리스트를 생성하는 텍스트생성부, 기설정된 규칙기반, 패턴기반, 통계기반 및 인공지능 기반 중 어느 하나의 번역기가 임베딩되고, 이를 이용하여 상기 대화창 리스트를 기설정된 언어들로 개별적으로 번역함에 따라 제1 및 제2 번역텍스트를 도출하는 번역도출부 및 상기 대화창 리스트로부터 확인되는 시간 및 화자정보에 기초하여, 상기 제1 및 제2 번역텍스트로부터 한쌍의 출력용자막을

순차적으로 선택하여 투명 디스플레이를 통해 서로 다른 출력방향으로 출력시키는 통합제어부를 포함하고, 상기 한쌍의 출력용자막은 서로 의미가 동일하게 번역되고 상기 제1 언어로 표시된 제1 출력용자막과 상기 제2 언어로 표시된 제2 출력용자막을 포함한다.

- [0011] 실시예에 있어서, 상기 번역기는 복수의 분야별 맞춤형 번역기들을 포함하고, 상기 복수의 분야별 맞춤형 번역기들은 직업, 연령, 지역, 관심분야, 종교에 따라 불필요 단어들 이 제거되고 최적화된 번역기들이다.
- [0012] 실시예에 있어서, 상기 인식부는 특정방향에서 발화되는 발화음성을 선택적으로 인식하기 위한 적어도 하나의 지향성마이크, 상기 발화음을 대신할 입력텍스트를 입력받기 위한 적어도 하나의 입력장치, 상기 특정방향을 향해 촬영하여 한쌍의 화자영상을 획득하는 적어도 하나의 카메라, 상기 한쌍의 화자영상 중 어느 하나로부터 기설정된 표정 및 제스처 객체를 검출하는 검출모듈, 상기 제스처 객체에 대응되는 수화용 텍스트와 상기 기설정된 표정에 따라 긴급안내 텍스트를 메모리로부터 추출하는 텍스트추출모듈 및 상기 발화음성, 상기 입력텍스트 및 상기 수화용 텍스트에 발화시간과 화자방향을 표시한 식별코드를 부여하여 상기 메모리에 저장하는 식별모듈을 포함한다.
- [0013] 실시예에 있어서, 상기 텍스트생성부는 상기 메모리에 기수집된 어휘단어사전을 학습하여 딥러닝 기반의 음성인식모델을 모델링하는 제1 모델링부, 상기 메모리에 기수집된 발음규칙을 학습하여 규칙 및 통계 기반의 음향모델을 모델링하는 제2 모델링부, 상기 제1 및 제2 언어의 음성을 음성인식 특징벡터에 따라 입력신호로 변환처리하는 신호처리부 및 상기 입력신호를 상기 음성인식모델과 상기 음향모델에 적용함에 따라 문자정보인 화자별 텍스트로 출력하고, 상기 식별코드에 따라 상기 입력텍스트 및 상기 수화용 텍스트를 화자별 텍스트에 병합하는 디코더를 포함한다.
- [0014] 실시예에 있어서, 상기 통합제어부는 상기 인식부를 통해 인식된 상기 제1 및 제2 언어의 각 음성으로부터 분석되는 음성특성과 기설정된 기준샘플을 비교하는 비교부 및 상기 비교부를 통해 확인된 비교 결과에 기초하여, 상기 제1 및 제2 언어의 각 음성을 기준샘플의 음성특징에 따라 보정하는 음성보정부를 포함한다.
- [0015] 실시예에 있어서, 상기 음성보정부는 상기 인식부를 통해 인식되는 묵음신호로부터 확인된 소음신호에 따라 노이즈 캔슬링 신호를 생성하고, 이를 토대로 상기 제1 및 제2 언어의 음성신호에 대한 노이즈를 필터링하는 필터링부 및 상기 묵음신호가 일정시간 동안 지속되는 경우, 상기 투명 디스플레이를 통해 상기 한쌍의 출력용자막에 대한 출력을 일시중지시키고 분위기 전환을 위한 힐링콘텐츠 영상을 대체하여 출력시키는 콘텐츠제공부를 포함하고, 상기 묵음신호는 상기 제1 및 제2 언어의 음성에 해당하는 주파수대역의 볼륨이 기설정된 시간동안 인식받지 않는 동안에 상기 인식부를 통해 인식된 특정주파수 대역의 반복적인 소음신호를 포함한다.
- [0016] 실시예에 있어서, 상기 통합제어부는 상기 투명 디스플레이에 배치된 거리측정센서를 통해 측정된 화자별 거리 정보에 기초하여, 상기 투명 디스플레이에 대한 상기 한쌍의 출력용자막의 출력면적을 서로 다른 면적비율로 조절한다.
- [0017] 실시예에 있어서, 상기 통합제어부는 상기 한쌍의 출력용자막을 출력시키기 이전에 간이 시력 검사표를 출력함에 따라 피드백받는 응답신호에 기초하여 화자시력을 추정하고, 이를 기초로 한쌍의 출력용자막에 대한 각 글씨체 확대비율을 자동으로 조절한다.
- [0018] 실시예에 있어서, 상기 통합제어부는 상기 투명 디스플레이에 배치된 아이트래커 모듈을 통해 측정된 안구운동 데이터에 기초하여, 상기 한쌍의 출력용자막 이후에 출력되는 다음자막의 페이지 전환속도를 조절한다.

**발명의 효과**

- [0019] 본 발명의 실시예에 따르면, 양방향으로 인식된 대화를 통역 및 번역하여 기설정된 한쌍의 언어로 구성된 대화 텍스트를 실시간으로 출력함으로써, 통번역에 따른 소통을 보다 빠르게 진행할 수 있다.
- [0020] 또한, 화자의 투명 디스플레이를 통해 한쌍의 출력용자막 이후에 출력되는 다음자막의 페이지 전환속도를 조절함으로써, 소통을 보다 원활하게 지원할 수 있다.

**도면의 간단한 설명**

- [0021] 도 1a는 본 발명의 실시예에 따른 자동 통번역 및 대화 보조 시스템(1000)을 개략적으로 나타내는 도이다.  
 도 1b는 도 1a의 자동 통번역 및 대화 보조 시스템(1000)에 대한 실시예를 보여주는 사진들이다.

도 2는 도 1의 인식부(100)를 구체적으로 나타내는 블록도이다.

도 3은 도 1의 텍스트생성부(200)의 실시예에 따른 블록도이다.

도 4는 도 1의 통합제어부(400)의 실시예에 따른 블록도이다.

**발명을 실시하기 위한 구체적인 내용**

- [0022] 이하, 본 발명의 실시예와 도면을 참조하여 본 발명을 상세히 설명한다. 이들 실시예는 오로지 본 발명을 보다 구체적으로 설명하기 위해 예시적으로 제시한 것일 뿐, 본 발명의 범위가 이들 실시예에 의해 제한되지 않는다는 것은 당업계에서 통상의 지식을 가지는 자에 있어서 자명할 것이다.
- [0023] 또한, 달리 정의하지 않는 한, 본 명세서에서 사용되는 모든 기술적 및 과학적 용어는 본 발명이 속하는 기술 분야의 숙련자에 의해 통상적으로 이해되는 바와 동일한 의미를 가지며, 상충되는 경우에는, 정의를 포함하는 본 명세서의 기재가 우선할 것이다.
- [0024] 도면에서 제안된 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다. 그리고, 어떤 부분이 어떤 구성 요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성 요소를 제외하는 것이 아니라 다른 구성 요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에서 기술한 "부"란, 특정 기능을 수행하는 하나의 단위 또는 블록을 의미한다.
- [0025] 각 단계들에 있어 식별부호(제1, 제2, 등)는 설명의 편의를 위하여 사용되는 것으로 식별부호는 각 단계들의 순서를 설명하는 것이 아니며, 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않는 이상 명기된 순서와 다르게 실시될 수 있다. 즉, 각 단계들은 명기된 순서와 동일하게 실시될 수도 있고 실질적으로 동시에 실시될 수도 있으며 반대의 순서대로 실시될 수도 있다. 도 1a는 본 발명의 실시예에 따른 자동 통번역 및 대화 보조 시스템(1000)을 개략적으로 나타내는 도이고, 도 1b는 도 1a의 자동 통번역 및 대화 보조 시스템(1000)에 대한 실시예를 보여주는 사진들이다.
- [0026] 도 1a와 도 1b를 참조하여 설명하면, 자동 통번역 및 대화 보조 시스템(1000)은 인식부(100), 텍스트생성부(200), 번역도출부(300) 및 통합제어부(400)를 포함할 수 있다.
- [0027] 먼저, 인식부(100)는 일방향에서 발화되는 제1 언어의 음성과 타방향에서 발화되는 제2 언어의 음성을 화자별로 구분하여 개별적으로 인식할 수 있다.
- [0028] 다음으로, 텍스트생성부(200)는 인식부(100)를 통해 순차적으로 인식되는 제1 및 제2 언어의 음성을 딥러닝 기반의 인공지능모델을 통해 분석함에 따라, 시간별 문장텍스트를 생성하고 이를 토대로 대화창 리스트를 생성할 수 있다.
- [0029] 여기서, 대화창 리스트는 일방향에 위치한 화자와 타방향에 위치한 화자 간의 시간별 문장텍스트들이 서로 구분되는 채팅창일 수 있다.
- [0030] 이때, 딥러닝 기반의 인공지능모델은 인공 신경 회로망(Artificial Neural Network), SVM(Support Vector Machine), 의사 결정 트리(Decision Tree) 및 랜덤 포레스트(Random Forest) 중 어느 하나의 알고리즘일 수 있다. 예를 들면, 인공 신경 회로망은 주로 딥러닝에서 사용되어 지고, 기계학습과 생물학의 신경망에서 영감을 얻은 통계학적 학습 알고리즘으로서, 특징 추출 신경망과 분류 신경망을 포함하는 컨볼루션 신경망일 수 있다.
- [0031] 다음으로, 번역도출부(300)는 기설정된 규칙기반(Rule-Based Machine Translation: RBMT), 패턴기반, 통계기반(Statistical Machine Translation: SMT), 및 인공지능 기반(Neuronal Machine Translation: NMT) 중 어느 하나의 번역기가 임베딩되고, 이를 이용하여 대화창 리스트에서 검출되는 화자별 문장텍스트를 기설정된 언어로 개별적으로 번역함에 따라 제1 및 제2 번역텍스트를 도출할 수 있다.
- [0032] 여기서, 번역기는 오프라인 상에서 제1 및 제2 번역텍스트를 도출하기 위하여, 기설정된 규칙기반(Rule-Based Machine Translation: RBMT), 패턴기반, 통계기반(Statistical Machine Translation: SMT) 및 인공지능 기반(Neuronal Machine Translation: NMT)의 번역 기술에서 필요로 하는 형태소 사전 데이터, 문법 사전 데이터, 문장 패턴 사전 데이터, 구(phase) 사전 데이터 및 병렬 코퍼스 데이터를 메모리(미도시)에 사전에 등록할 수 있다.
- [0033] 이러한 번역기는 복수의 분야별 맞춤형 번역기들을 포함할 수 있다. 구체적으로, 복수의 분야별 맞춤형 번역기

들은 직업, 연령, 지역, 관심분야, 종교에 따라 불필요 단어들 이 제거된 번역기들일 수 있다. 예를 들면, 이용자의 직업이 의사인 경우, 번역도출부(300)는 의학 용어로 최적화된 번역기를 이용하고, 이용자의 연령이 유아인 경우, 신조어 용어로 최적화된 번역기를 이용할 수 있다.

- [0034] 일 실시예에 따라, 대화창 리스트에서 검출되는 화자별 문장텍스트를 도메인 특화 기반의 번역기술을 통해 기설정된 언어로 개별적으로 번역함에 따라, 제1 및 제2 번역텍스트를 획득할 수도 있다.
- [0035] 여기서, 도메인 특화 기반의 번역기술은 transformer 기반의 인코딩 및 디코딩 알고리즘을 통해 화자별 문장텍스트를 기설정된 언어로 기계번역할 수 있다.
- [0036] 다음으로, 통합제어부(400)는 대화창 리스트로부터 확인되는 시간 및 화자정보에 기초하여, 제1 및 제2 번역텍스트로부터 한쌍의 출력용자막을 선택하여 투명 디스플레이(10)를 통해 서로 다른 출력방향으로 출력시킬 수 있다.
- [0037] 여기서, 한쌍의 출력용자막은 서로 의미가 대응되도록 번역되는 문장텍스트로써, 일측에 위치한 화자에게 제공하기 위하여 제1 언어로 표시된 제1 출력용자막과 타측에 위치한 화자에게 제공하기 위하여 제2 언어로 표시된 제2 출력용자막을 포함할 수 있다.
- [0038] 일 실시예에 따라, 통합제어부(400)는 투명 디스플레이(10)에 배치된 거리측정센서(미도시)를 통해 측정된 화자별 거리정보에 기초하여, 투명 디스플레이(10)에 대한 한쌍의 출력용자막의 각 출력면적을 서로 다른 면적비율로 조절할 수 있다.
- [0039] 예를 들면, 일측에 위치한 화자의 거리정보가 80cm 이고 타측에 위치한 화자의 거리정보가 40cm로 일정거리 차이를 가지는 경우, 통합제어부(400)는 한쌍의 출력용자막 중 제1 출력용자막의 출력면적을 증가시키고 제2 출력용자막의 출력면적을 증가시킬 수 있다.
- [0040] 다른 실시예에 따라, 통합제어부(400)는 투명 디스플레이(10)에 배치된 레이저센서(미도시)를 통해 측정된 화자별 호흡 및 심박신호에 따라 분석되는 화자별 심리상태에 기초하여, 한쌍의 출력용자막을 컬러테라피 색상으로 보정처리할 수 있다.
- [0041] 예를 들면, 심리상태가 무기력한 경우 컬러테라피 색상은 빨강이고, 심리상태가 슬픔과 상실감인 경우 컬러테라피 색상은 주황이며, 심리상태가 불면증, 피로감 및 두통인 경우 컬러테라피 색상은 파랑이고, 심리상태가 불안정한 경우 컬러테라피 색상은 초록이고, 심리상태가 따분하거나 지겨움인 경우 컬러테라피 색상은 노랑일 수 있다.
- [0042] 또 다른 실시예에 따라, 통합제어부(400)는 한쌍의 출력용자막을 출력하기 이전에 간이 시력 검사표를 출력함에 따라 피드백받는 응답신호에 기초하여 화자시력을 추정하고, 이를 기초로 한쌍의 출력용자막에 대한 각 글씨체 확대비율을 자동으로 조절할 수 있다.
- [0043] 또 다른 실시예에 따라, 통합제어부(400)는 투명 디스플레이(10)에 배치된 아이트래커 모듈(미도시)을 통해 측정된 안구운동 데이터에 기초하여, 투명 디스플레이(10)를 통해 한쌍의 출력용자막 이후에 출력되는 다음자막의 페이지 전환속도를 조절할 수 있다.
- [0044] 예를 들면, 안구운동 데이터가 기설정된 속도 이상인 경우, 통합제어부(400)는 투명 디스플레이(10)를 통해 한쌍의 출력용자막 이후에 출력되는 다음자막의 페이지 전환속도를 증가시킬 수 있다. 또한, 안구운동 데이터가 기설정된 속도 미만인 경우, 통합제어부(400)는 투명 디스플레이(10)를 통해 한쌍의 출력용자막 이후에 출력되는 다음자막의 페이지 전환속도를 감소시킬 수 있다.
- [0045] 또 다른 실시예에 따라, 통합제어부(400)는 하기 도 2에서 설명될 적어도 하나의 입력장치(120)를 통해 입력받는 SNS 계정정보를 등록함에 따라, 네트워크를 통해 연결된 통합관리서버(2000)를 이용하여 SNS 계정정보에 대응되는 화자단말(20)에 대화창 리스트를 이미지 파일로 전송할 수 있다.
- [0046] 또 다른 실시예에 따라, 통합제어부(400)는 메모리(미도시)에 수집된 제1 및 제2 번역텍스트에 기초하여, 제1 번역텍스트를 투명 디스플레이(10)를 통해 출력함에 따라 제2 번역텍스트를 음성, 수화 및 텍스트 중 어느 하나의 방식으로 요청하는 학습서비스 모드를 제공할 수 있다.
- [0047] 또 다른 실시예에 따라, 통합제어부(400)는 투명 디스플레이(10)에 배치된 환경감지센서들(미도시)을 통해 감지된 재난환경상태에 기초하여, 투명 디스플레이(10)를 통해 재난정보를 출력할 수 있다. 여기서, 환경감지센서들(미도시)은 연기 감지센서, 일산화탄소 감지센서, 온도 감지센서를 포함할 수 있다.

- [0048] 이하, 구체적인 실시예와 비교예를 통하여 본 발명의 구성 및 그에 따른 효과를 보다 상세히 설명하고자 한다. 그러나, 본 실시예는 본 발명을 보다 구체적으로 설명하기 위한 것이며, 본 발명의 범위가 이들 실시예에 한정되는 것은 아니다.
- [0049] 도 2는 도 1의 인식부(100)를 구체적으로 나타내는 블록도이다.
- [0050] 도 1a 내지 도 2를 참조하면, 인식부(100)는 적어도 하나의 지향성마이크(110), 적어도 하나의 입력장치(120), 적어도 하나의 카메라(130), 검출모듈(140), 텍스트추출모듈(150) 및 식별모듈(160)을 포함할 수 있다.
- [0051] 먼저, 적어도 하나의 지향성마이크(110)는 투명 디스플레이(10)에 배치되어, 투명 디스플레이(10)로부터 일정거리 이격된 특정방향에서 발화되는 발화음성을 선택적으로 인식할 수 있다.
- [0052] 다음으로, 적어도 하나의 입력장치(120)는 투명 디스플레이(10)로부터 일정거리 이격되도록 배치되어, 발화음성을 대신할 입력텍스트를 입력받을 수 있다.
- [0053] 다음으로, 적어도 하나의 카메라(130)는 특정방향을 향해 촬영하여 한쌍의 화자영상을 획득할 수 있다.
- [0054] 예를 들면, 적어도 하나의 지향성마이크(110)와 적어도 하나의 카메라(130)는 투명 디스플레이(10)를 기준으로 일면과 타면에 각각 개별적으로 배치되고, 적어도 하나의 입력장치(120)는 양면을 기준으로 일정거리 이격된 위치에 배치될 수 있다.
- [0055] 다음으로, 검출모듈(140)은 적어도 하나의 카메라(130)를 통해 촬영된 한쌍의 화자영상 중 어느 하나로부터 기설정된 표정 및 제스처 객체를 검출할 수 있다.
- [0056] 다음으로, 텍스트추출모듈(150)은 제스처 객체에 대응되는 수화용 텍스트와 기설정된 표정에 따라 기설정된 긴급안내 텍스트를 메모리(미도시)로부터 추출하여 인식할 수 있다.
- [0057] 다음으로, 식별모듈(160)은 적어도 하나의 지향성마이크(110)를 통해 인식된 발화음성, 적어도 하나의 입력장치(120)를 통해 입력받는 입력텍스트 및 텍스트추출모듈(150)을 통해 추출된 수화용 텍스트에 대화시간과 화자방향을 표시한 식별코드를 부여하여 메모리(미도시)에 저장 및 분류할 수 있다.
- [0058] 이러한 식별모듈(160)은 제1 및 제2 언어의 음성을 식별코드에 따라 분류하여 텍스트생성부(200)에 문장단위별로 딥러닝 기반의 인공지능모델에 순차적으로 입력할 수 있다.
- [0059] 도 3은 도 1의 텍스트생성부(200)의 실시예에 따른 블록도이다.
- [0060] 도 1a 내지 도 3을 참조하여 설명하면, 텍스트생성부(200)는 제1 및 제2 모델링부(210, 220), 잡음처리부(230) 및 디코더(240)를 포함할 수 있다.
- [0061] 먼저, 제1 모델링부(210)는 메모리(미도시)에 기수집된 어휘단어사전을 학습하여 딥러닝 기반의 음성인식모델을 모델링할 수 있다. 여기서, 딥러닝 기반의 음성인식모델은 기수집된 어휘단어사전을 학습하는 구글의 답마인드와 IBM의 TrueNorth 중 어느 하나를 통해 도출되는 알고리즘일 수 있다.
- [0062] 다음으로, 제2 모델링부(220)는 메모리(미도시)에 기수집된 발음규칙을 학습하여 규칙 및 통계 기반의 음향모델을 모델링할 수 있다. 여기서, 규칙 및 통계 기반의 음향모델은 기수집된 발음규칙을 학습하는 IBM의 Watson 프로그램을 통해 도출되는 알고리즘일 수 있다.
- [0063] 다음으로, 신호처리부(230)는 식별모듈(160)을 통해 문장단위별로 입력받는 제1 및 제2 언어의 음성을 음성인식 특징벡터에 따라 입력신호로 변환처리할 수 있다.
- [0064] 여기서, 음성인식 특징벡터는 MFCC(mel frequency cepstral coefficient), LPC(linear predictive coefficients), PLP(perceptual linear prediction) 등과 같은 다양한 특징 벡터들 중의 어느 하나를 포함할 수 있다.
- [0065] 이때, 입력신호는 딥러닝 기반의 음성인식모델과 규칙 및 통계 기반의 음향모델에 입력되는 인자를 의미할 수 있다.
- [0066] 다음으로, 디코더(240)는 신호처리부(230)를 통해 변환처리된 입력신호를 딥러닝 기반의 음성인식모델과 규칙 및 통계 기반의 음향모델에 적용함에 따라 문자정보인 화자별 텍스트로 출력할 수 있다.
- [0067] 이때, 디코더(240)는 메모리(미도시)에서 확인된 식별코드에 따라, 화자별 텍스트에 입력텍스트 및 수화용 텍스

트를 화자별 텍스트에 순차적으로 리스트 및 병합시킬 수 있다.

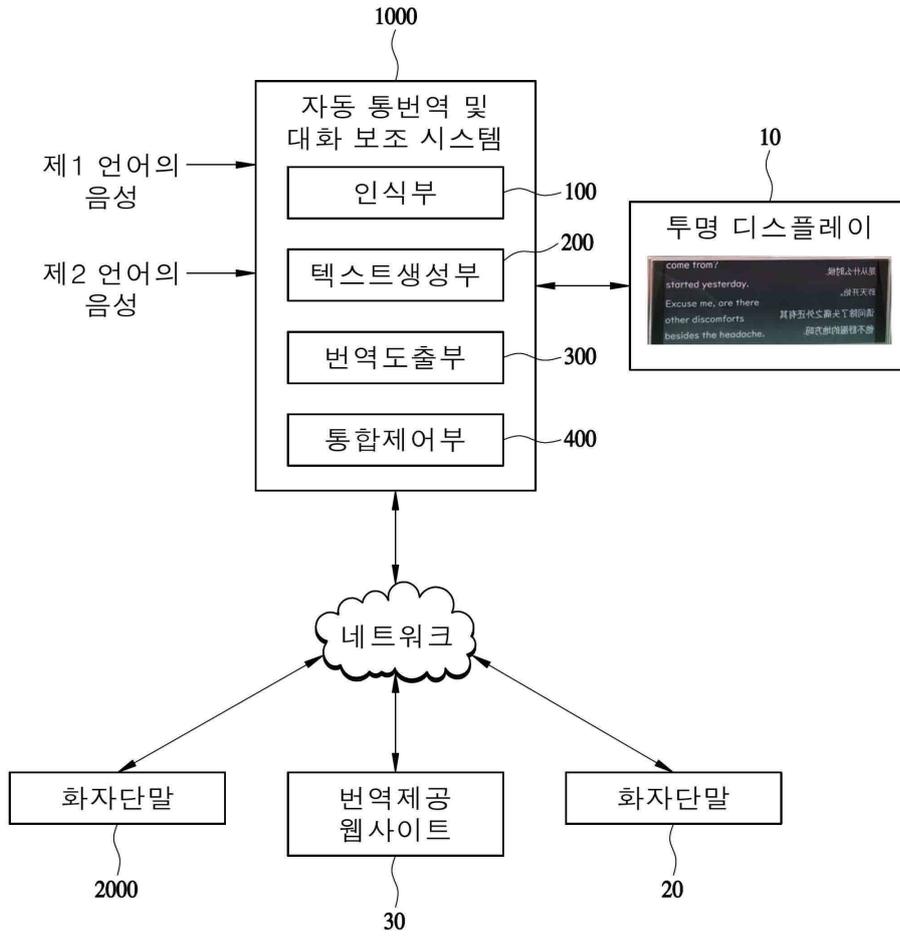
- [0068] 도 4는 도 1의 통합제어부(400)의 실시예에 따른 블록도이다.
- [0069] 도 1a 내지 도 4를 참조하면, 통합제어부(400)는 비교부(410)와 음성보정부(420)를 포함할 수 있다.
- [0070] 먼저, 비교부(410)는 인식부(100)를 통해 인식된 제1 및 제2 언어의 각 음성으로부터 분석되는 음성특성과 기설정된 기준샘플을 비교할 수 있다.
- [0071] 여기서, 기준샘플은 음성을 정확하게 인식할 수 있게 하는 샘플음성을 의미하고, 또한, 음성특성은 음절당 발생 속도, 발생 중 묵음시간 및 음성의 포먼트주파수 중 어느 하나를 포함할 수 있다.
- [0072] 이때, 음성보정부(420)는 비교부(410)를 통해 확인된 비교 결과에 기초하여, 제1 및 제2 언어의 각 음성을 기준샘플의 음성특징에 따라 보정하여 인식부(100)로 피드백 제공할 수 있다.
- [0073] 예를 들면, 음절 당 발생 속도가 기준샘플에 해당하는 임계 범위를 벗어날 경우, 음성보정부(220)는 제1 및 제2 언어의 음성의 발생 속도를 기준샘플에 따른 발생 속도로 보정할 수 있다.
- [0074] 일 실시예에 따라, 음성보정부(420)는 인식부(100)를 통해 인식된 제1 및 제2 언어의 음성신호를 주파수 분석함에 따라 획득된 오디오 시각정보를 한쌍의 출력용자막에 통합하여 투명 디스플레이(10)를 통해 출력함으로써, 화자음성의 최적 볼륨 크기로 발화되도록 유도할 수 있다.
- [0075] 다른 실시예에 따라, 음성보정부(420)는 인식부(100)를 통해 인식된 음성의 가청 주파수 이외의 대역 예컨대, 1000,000hz 이상의 초음파 및 7hz 이하의 초저주파의 볼륨크기에 기초하여, 제1 및 제2 언어의 음성신호에 대한 기설정된 가청주파수 대역을 증폭시키도록 제어할 수 있다.
- [0076] 이러한 음성보정부(420)는 필터링부(421)과 콘텐츠제공부(422)를 더 포함할 수 있다. 구체적으로, 필터링부(421)는 인식부(100)를 통해 인식되는 묵음신호로부터 확인된 소음신호에 따라 노이즈 캔슬링 신호를 생성하고, 이를 토대로 제1 및 제2 언어의 음성신호에 대한 노이즈를 필터링할 수 있다. 여기서, 묵음신호는 제1 및 제2 언어의 음성에 해당하는 주파수대역의 볼륨이 기설정된 시간동안 인식받지 않는 동안에 인식부(100)를 통해 인식된 특정주파수 대역의 반복적인 소음신호를 포함할 수 있다. 이때, 콘텐츠제공부(422)는 묵음신호가 일정시간 동안 지속되는 경우, 투명 디스플레이(10)를 통해 한쌍의 출력용자막에 대한 출력을 일시중지시키고 분위기 전환을 위한 힐링콘텐츠를 대체하여 출력시킬 수 있다.
- [0077] 본 명세서에서는 본 발명자들이 수행한 다양한 실시예 가운데 몇 개의 예만을 들어 설명하는 것이나 본 발명의 기술적 사상은 이에 한정하거나 제한되지 않고, 당업자에 의해 변형되어 다양하게 실시될 수 있음은 물론이다.

**부호의 설명**

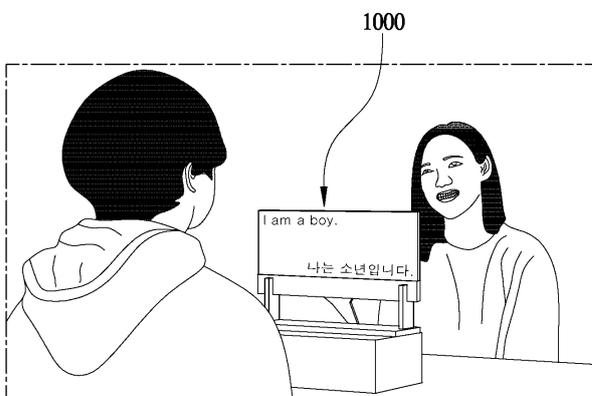
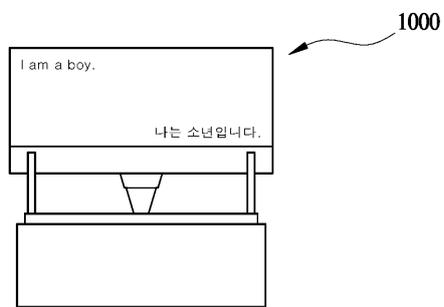
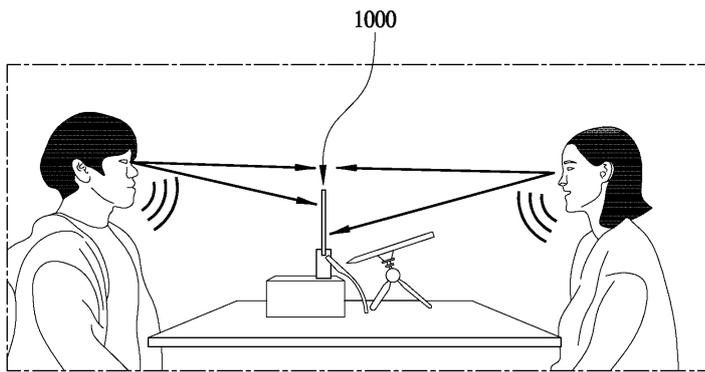
- [0078] 10: 투명 디스플레이
- 100: 인식부
- 200: 텍스트생성부
- 300: 번역도출부
- 400: 통합제어부
- 1000: 자동 통번역 및 대화 보조 시스템

도면

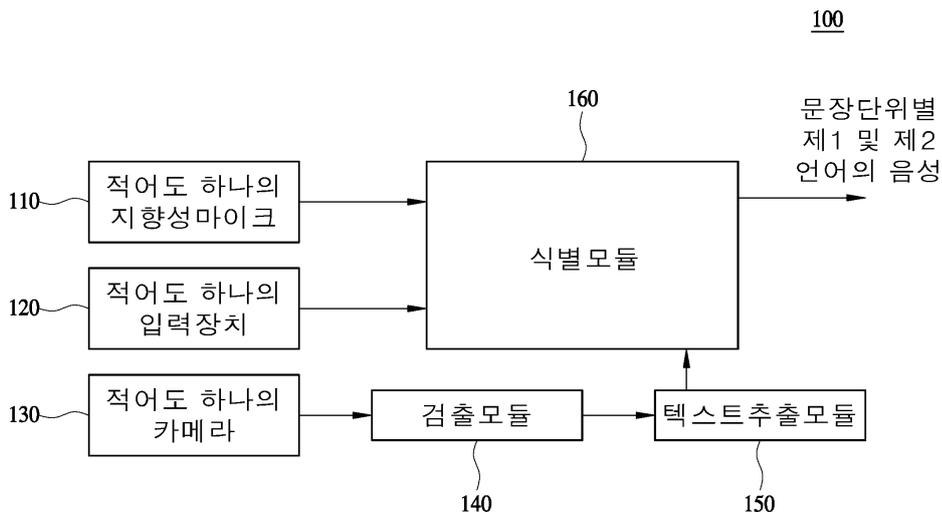
도면1a



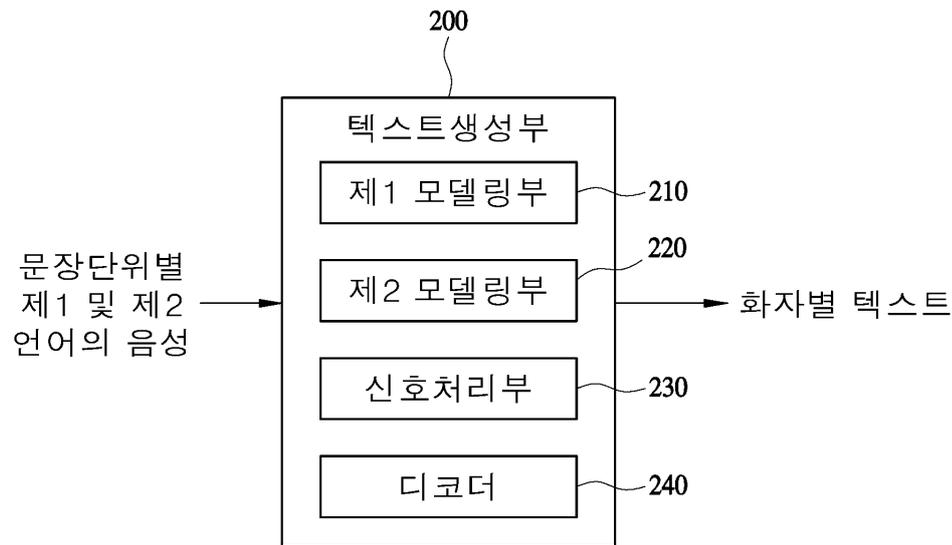
도면1b



도면2



도면3



도면4

