

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구  
국제사무국

(43) 국제공개일

2024년 1월 4일 (04.01.2024)



(10) 국제공개번호

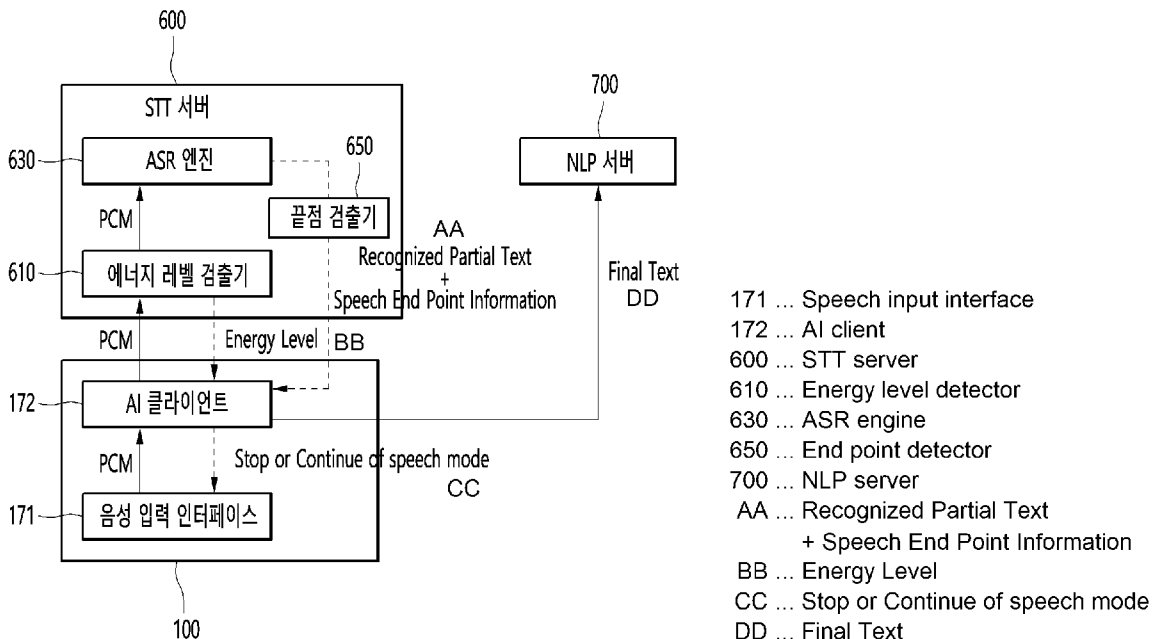
WO 2024/005226 A1

- (51) 국제특허분류: *G10L 15/22* (2006.01)      *G10L 25/87* (2013.01)  
*G10L 15/30* (2013.01)      *G10L 15/18* (2006.01)  
*G10L 15/26* (2006.01)      *G10L 19/10* (2006.01)
- (21) 국제출원번호: PCT/KR2022/009357
- (22) 국제출원일: 2022년 6월 29일 (29.06.2022)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (71) 출원인: 엘지전자 주식회사 (**LG ELECTRONICS INC.**) [KR/KR]: 07336 서울특별시 영등포구 여의대로 128, Seoul (KR).
- (72) 발명자: 허승현 (**HEO, Seunghyun**): 06772 서울특별시 서초구 양재대로11길 19 LG전자 특허센터, Seoul (KR).
- (74) 대리인: 허용록 (**HAW, Yong Noke**): 06252 서울특별시 강남구 역삼로 114 현죽빌딩 6층, 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, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, 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, SC, 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

(54) Title: DISPLAY DEVICE

(54) 발명의 명칭: 디스플레이 장치

70



(57) Abstract: The present disclosure relates to a display device capable of accurately recognizing an end point of a speech input of a user, and the display device may comprise a network interface which communicates with a first server and a second server, and a controller which: acquires a speech input of a user; transmits, to the first server, a speech signal corresponding to the acquired speech input; receives, from the first server, the energy level of the speech signal, text corresponding to the speech input, and speech end point information for the speech input; and determines whether an utterance of the user has ended on the basis of the energy level and the speech end point information.

WO 2024/005226 A1

(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML,  
MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

**(57) 요약서:** 본 개시는 사용자의 음성 입력의 끝점을 정확히 인식할 수 있는 디스플레이 장치에 관한 것으로, 디스플레이 장치는 제1 서버 및 제2 서버와 통신하는 네트워크 인터페이스 및 사용자의 음성 입력을 획득하고, 획득된 음성 입력에 대응하는 음성 신호를 상기 제1 서버에 전송하고, 상기 제1 서버로부터 상기 음성 신호의 에너지 레벨, 상기 음성 입력에 대응하는 텍스트 및 상기 음성 입력에 대한 발화 끝점 정보를 수신하고, 상기 에너지 레벨 및 상기 발화 끝점 정보에 기초하여 상기 사용자의 발화가 종료되었는지 여부를 판단하는 컨트롤러를 포함할 수 있다.

## 명세서

### 발명의 명칭: 디스플레이 장치

#### 기술분야

- [1] 본 개시는 디스플레이 장치에 관한 것으로, 보다 상세하게는 음성 인식 서비스를 제공하는 디스플레이 장치에 관한 것이다.

#### 배경기술

- [2] 유선 또는 무선 통신망을 이용한 디지털 TV 서비스가 보편화되고 있다. 디지털 TV 서비스는 기존의 아날로그 방송 서비스에서는 제공할 수 없었던 다양한 서비스를 제공할 수 있다.
- [3] 예를 들어, 디지털 TV 서비스의 종류인 IPTV(Internet Protocol Television), smart TV 서비스의 경우 사용자로 하여금 시청 프로그램의 종류, 시청 시간 등을 능동적으로 선택할 수 있도록 하는 양방향성을 제공한다. IPTV, smart TV 서비스는 이러한 양방향성을 기반으로 하여 다양한 부가 서비스, 예컨대 인터넷 검색, 홈쇼핑, 온라인 게임 등을 제공할 수도 있다.
- [4] 최근의 TV는 사용자의 음성 인식에 기반한 음성 인식 서비스를 제공하고 있다. TV의 리모컨에 구비된 마이크 버튼을 통해 음성 입력을 하는 경우에는 버튼의 Press/Release를 통해 사용자 입력의 시작과 끝을 알 수 있다.
- [5] 하지만, 원거리 음성명령과 같이 기동어(Wake-up word)를 사용하는 경우나, 가상 키보드의 음성버튼을 활용하는 등 사용자 발화의 끝을 알 수 없는 경우에는 신호의 크기정보(예를 들어, Amplitude, Energy strength)를 활용하여 음성입력의 끝을 인지한다.
- [6] 하지만, 종래의 기술은 주변에 노이즈나 다른 소리가 있는 경우 발화의 끝을 사용자의 의도와 같이 인식하지 못하고, 계속 음성 입력을 받는 상태를 유지한다.
- [7] TV는 사용자가 음성 입력 끝마쳤음에도 이를 모르고 음성입력을 계속 받는 상태로 유지하여, 원하지 않는 음성 인식 결과를 출력한다.
- [8] 이에 따라, 사용자는 음성 인식 서비스를 제공받는 과정에서 상당한 불편함을 느낄 수 있다.

#### 발명의 상세한 설명

##### 기술적 과제

- [9] 본 개시는 사용자의 음성 입력의 끝점을 정확히 인식할 수 있는 디스플레이 장치의 제공을 목적으로 한다.
- [10] 본 개시는 음성 입력의 에너지 레벨 및 음성 입력의 끝점 정보를 활용하여 정확한 음성 인식 서비스를 제공하는 것을 목적으로 한다.

##### 과제 해결 수단

- [11] 본 개시의 실시 예에 따른 디스플레이 장치는 제1 서버 및 제2 서버와 통신하는

네트워크 인터페이스 및 사용자의 음성 입력을 획득하고, 획득된 음성 입력에 상응하는 음성 신호를 상기 제1 서버에 전송하고, 상기 제1 서버로부터 상기 음성 신호의 에너지 레벨, 상기 음성 입력에 상응하는 텍스트 및 상기 음성 입력에 대한 발화 끝점 정보를 수신하고, 상기 에너지 레벨 및 상기 발화 끝점 정보에 기초하여 상기 사용자의 발화가 종료되었는지 여부를 판단하는 컨트롤러를 포함할 수 있다.

- [12] 본 개시의 또 다른 실시 예에 따른 디스플레이 장치는 제1 서버 및 제2 서버와 통신하는 네트워크 인터페이스 및 사용자의 음성 입력을 획득하고, 획득된 음성 입력에 상응하는 음성 신호를 상기 제1 서버에 전송하고, 상기 제1 서버로부터 상기 음성 신호의 에너지 레벨 및 상기 음성 입력에 상응하는 텍스트를 수신하고, 상기 텍스트를 상기 제2 서버에 전송하고, 상기 제2 서버로부터 상기 음성 입력에 대한 발화 끝점 정보를 수신하고, 상기 에너지 레벨 및 상기 발화 끝점 정보에 기초하여 상기 사용자의 발화가 종료되었는지 여부를 판단하는 컨트롤러를 포함할 수 있다.

### 발명의 효과

- [13] 본 개시의 실시 예에 따르면, 음성 입력에 대한 텍스트 분석 및 에너지 레벨을 이용하여 발화의 끝을 정확히 인식할 수 있다. 이에 따라 불필요한 음성 인식이 수행되는 것이 방지될 수 있다.
- [14] 본 개시의 실시 예에 따르면, 사용자의 음성 입력에 대한 에너지 레벨, 발화 끝점 정보 및 분석 결과 정보를 사용하여 사용자의 발화가 종료되었는지가 정확히 판단될 수 있다. 이에 따라 사용자가 음성 입력을 종료하였지만, 노이즈로 인한 음성 인식이 이루어지는 것이 방지될 수 있고, 음성 입력의 종료에 맞게 정확한 음성 인식 서비스가 제공될 수 있다.

### 도면의 간단한 설명

- [15] 도 1은 본 발명의 일 실시 예에 따른 디스플레이 장치의 구성을 블록도로 도시한 것이다.
- [16] 도 2은 본 발명의 일 실시 예에 따른 원격제어장치의 블록도이다.
- [17] 도 3은 본 발명의 일 실시 예에 따른 원격제어장치의 실제 구성 예를 보여준다.
- [18] 도 4는 본 발명의 실시 예에 따라 원격 제어 장치를 활용하는 예를 보여준다.
- [19] 도 5는 종래 기술에 따른 음성 입력의 끝점 인식을 위한 과정을 설명하는 도면이다.
- [20] 도 6은 본 개시의 일 실시 예에 따른 시스템의 동작 방법을 설명하는 래더 다이어그램이다.
- [21] 도 7은 본 개시의 일 실시 예에 따른 발화의 끝점을 검출하여 의도 분석을 수행하는 과정을 설명하는 도면이다.
- [22] 도 8은 본 개시의 일 실시 예에 따라 끝점 검출기가 발화 끝점 정보를 생성하는 과정을 설명하는 도면이다.

- [23] 도 9는 본 개시의 또 다른 실시 예에 따른 시스템의 동작 방법을 설명하는 래더 다이어그램이다.
- [24] 도 10은 본 개시의 또 다른 실시 예에 따른 발화의 끝점을 검출하여 의도 분석을 수행하는 과정을 설명하는 도면이다.
- 발명의 실시를 위한 최선의 형태**
- [25] 이하, 본 발명과 관련된 실시 예에 대하여 도면을 참조하여 보다 상세하게 설명한다. 이하의 설명에서 사용되는 구성요소에 대한 접미사 "모듈" 및 "부"는 명세서 작성의 용이함만이 고려되어 부여되거나 혼용되는 것으로서, 그 자체로서 구별되는 의미 또는 역할을 갖는 것은 아니다.
- [26] 본 발명의 실시예에 따른 디스플레이 장치는, 예를 들어 방송 수신 기능에 컴퓨터 지원 기능을 추가한 지능형 디스플레이 장치로서, 방송 수신 기능에 충실하면서도 인터넷 기능 등이 추가되어, 수기 방식의 입력 장치, 터치 스크린 또는 공간 리모콘 등 보다 사용에 편리한 인터페이스를 갖출 수 있다. 그리고, 유선 또는 무선 인터넷 기능의 지원으로 인터넷 및 컴퓨터에 접속되어, 이메일, 웹브라우저,뱅킹 또는 게임 등의 기능도 수행가능하다. 이러한 다양한 기능을 위해 표준화된 범용 OS가 사용될 수 있다.
- [27] 따라서, 본 발명에서 기술되는 디스플레이 장치는, 예를 들어 범용의 OS 커널 상에, 다양한 애플리케이션이 자유롭게 추가되거나 삭제 가능하므로, 사용자 친화적인 다양한 기능이 수행될 수 있다. 상기 디스플레이 장치는, 보다 구체적으로 예를 들면, 네트워크 TV, HBBTV, 스마트 TV, LED TV, OLED TV 등이 될 수 있으며, 경우에 따라 스마트폰에도 적용 가능하다.
- [28] 도 1은 본 발명의 일 실시예에 따른 디스플레이 장치의 구성을 블록도로 도시한 것이다.
- [29] 도 1을 참조하면, 디스플레이 장치(100)는 방송 수신부(130), 외부장치 인터페이스(135), 메모리(140), 사용자입력 인터페이스(150), 컨트롤러(170), 무선 통신 인터페이스(173), 디스플레이(180), 스피커(185), 전원 공급 회로(190)를 포함할 수 있다.
- [30] 방송 수신부(130)는 튜너(131), 복조기(132) 및 네트워크 인터페이스(133)를 포함할 수 있다.
- [31] 튜너(131)는 채널 선국 명령에 따라 특정 방송 채널을 선국할 수 있다. 튜너(131)는 선국된 특정 방송 채널에 대한 방송 신호를 수신할 수 있다.
- [32] 복조기(132)는 수신한 방송 신호를 비디오 신호, 오디오 신호, 방송 프로그램과 관련된 데이터 신호로 분리할 수 있고, 분리된 비디오 신호, 오디오 신호 및 데이터 신호를 출력이 가능한 형태로 복원할 수 있다.
- [33] 외부장치 인터페이스(135)는 인접하는 외부 장치 내의 애플리케이션 또는 애플리케이션 목록을 수신하여, 컨트롤러(170) 또는 메모리(140)로 전달할 수 있다.

- [34] 외부장치 인터페이스(135)는 디스플레이 장치(100)와 외부 장치 간의 연결 경로를 제공할 수 있다. 외부장치 인터페이스(135)는 디스플레이 장치(100)에 무선 또는 유선으로 연결된 외부장치로부터 출력된 영상, 오디오 중 하나 이상을 수신하여, 컨트롤러(170)로 전달할 수 있다. 외부장치 인터페이스(135)는 복수의 외부 입력 단자들을 포함할 수 있다. 복수의 외부 입력 단자들은 RGB 단자, 하나 이상의 HDMI(High Definition Multimedia Interface) 단자, 컴포넌트(Component) 단자를 포함할 수 있다.
- [35] 외부장치 인터페이스(135)를 통해 입력된 외부장치의 영상 신호는 디스플레이(180)를 통해 출력될 수 있다. 외부장치 인터페이스(135)를 통해 입력된 외부장치의 음성 신호는 스피커(185)를 통해 출력될 수 있다.
- [36] 외부장치 인터페이스(135)에 연결 가능한 외부 장치는 셋톱박스, 블루레이 플레이어, DVD 플레이어, 게임기, 사운드 바, 스마트폰, PC, USB 메모리, 홈 씨어터 중 어느 하나일 수 있으나, 이는 예시에 불과하다.
- [37] 네트워크 인터페이스(133)는 디스플레이 장치(100)를 인터넷망을 포함하는 유/무선 네트워크와 연결하기 위한 인터페이스를 제공할 수 있다. 네트워크 인터페이스(133)는 접속된 네트워크 또는 접속된 네트워크에 링크된 다른 네트워크를 통해, 다른 사용자 또는 다른 전자 기기와 데이터를 송신 또는 수신할 수 있다.
- [38] 또한, 디스플레이 장치(100)에 미리 등록된 다른 사용자 또는 다른 전자 기기 중 선택된 사용자 또는 선택된 전자기기에, 디스플레이 장치(100)에 저장된 일부의 콘텐츠 데이터를 송신할 수 있다.
- [39] 네트워크 인터페이스(133)는 접속된 네트워크 또는 접속된 네트워크에 링크된 다른 네트워크를 통해, 소정 웹 페이지에 접속할 수 있다. 즉, 네트워크를 통해 소정 웹 페이지에 접속하여, 해당 서버와 데이터를 송신 또는 수신할 수 있다.
- [40] 그리고, 네트워크 인터페이스(133)는 콘텐츠 제공자 또는 네트워크 운영자가 제공하는 콘텐츠 또는 데이터들을 수신할 수 있다. 즉, 네트워크 인터페이스(133)는 네트워크를 통하여 콘텐츠 제공자 또는 네트워크 제공자로부터 제공되는 영화, 광고, 게임, VOD, 방송 신호 등의 콘텐츠 및 그와 관련된 정보를 수신할 수 있다.
- [41] 또한, 네트워크 인터페이스(133)는 네트워크 운영자가 제공하는 펌웨어의 업데이트 정보 및 업데이트 파일을 수신할 수 있으며, 인터넷 또는 콘텐츠 제공자 또는 네트워크 운영자에게 데이터들을 송신할 수 있다.
- [42] 네트워크 인터페이스(133)는 네트워크를 통해, 공중에 공개(open)된 애플리케이션들 중 원하는 애플리케이션을 선택하여 수신할 수 있다.
- [43] 메모리(140)는 컨트롤러(170) 내의 각 신호 처리 및 제어를 위한 프로그램이 저장하고, 신호 처리된 영상, 음성 또는 데이터신호를 저장할 수 있다.
- [44] 또한, 메모리(140)는 외부장치 인터페이스(135) 또는 네트워크 인터페이스(133)로부터 입력되는 영상, 음성, 또는 데이터 신호의 임시 저장을

- 위한 기능을 수행할 수도 있으며, 채널 기억 기능을 통하여 소정 이미지에 관한 정보를 저장할 수도 있다.
- [45] 메모리(140)는 외부장치 인터페이스(135) 또는 네트워크 인터페이스(133)로부터 입력되는 애플리케이션 또는 애플리케이션 목록을 저장할 수 있다.
- [46] 디스플레이 장치(100)는 메모리(140) 내에 저장되어 있는 콘텐츠 파일(동영상 파일, 정지영상 파일, 음악 파일, 문서 파일, 애플리케이션 파일 등)을 재생하여 사용자에게 제공할 수 있다.
- [47] 사용자입력 인터페이스(150)는 사용자가 입력한 신호를 컨트롤러(170)로 전달하거나, 컨트롤러(170)로부터의 신호를 사용자에게 전달할 수 있다. 예를 들어, 사용자입력 인터페이스(150)는 블루투스(Bluetooth), WB(Ultra Wideband), 지그비(ZigBee) 방식, RF(Radio Frequency) 통신 방식 또는 적외선(IR) 통신 방식 등 다양한 통신 방식에 따라, 원격제어장치(200)로부터 전원 온/오프, 채널 선택, 화면 설정 등의 제어 신호를 수신하여 처리하거나, 컨트롤러(170)로부터의 제어 신호를 원격제어장치(200)로 송신하도록 처리할 수 있다.
- [48] 또한, 사용자입력 인터페이스(150)는, 전원키, 채널키, 볼륨키, 설정키 등의 로컬키(미도시)에서 입력되는 제어 신호를 컨트롤러(170)에 전달할 수 있다.
- [49] 컨트롤러(170)에서 영상 처리된 영상 신호는 디스플레이(180)로 입력되어 해당 영상 신호에 대응하는 영상으로 표시될 수 있다. 또한, 컨트롤러(170)에서 영상 처리된 영상 신호는 외부장치 인터페이스(135)를 통하여 외부 출력장치로 입력될 수 있다.
- [50] 컨트롤러(170)에서 처리된 음성 신호는 스피커(185)로 오디오 출력될 수 있다. 또한, 컨트롤러(170)에서 처리된 음성 신호는 외부장치 인터페이스(135)를 통하여 외부 출력장치로 입력될 수 있다.
- [51] 그 외, 컨트롤러(170)는, 디스플레이 장치(100) 내의 전반적인 동작을 제어할 수 있다.
- [52] 또한, 컨트롤러(170)는 사용자입력 인터페이스(150)를 통하여 입력된 사용자 명령 또는 내부 프로그램에 의하여 디스플레이 장치(100)를 제어할 수 있으며, 네트워크에 접속하여 사용자가 원하는 애플리케이션 또는 애플리케이션 목록을 디스플레이 장치(100) 내로 다운받을 수 있도록 할 수 있다.
- [53] 컨트롤러(170)는 사용자가 선택한 채널 정보 등이 처리한 영상 또는 음성신호와 함께 디스플레이(180) 또는 스피커(185)를 통하여 출력될 수 있도록 한다.
- [54] 또한, 컨트롤러(170)는 사용자입력 인터페이스(150)를 통하여 수신한 외부장치 영상 재생 명령에 따라, 외부장치 인터페이스(135)를 통하여 입력되는 외부 장치, 예를 들어, 카메라 또는 캠코더로부터의, 영상 신호 또는 음성 신호가 디스플레이(180) 또는 스피커(185)를 통해 출력될 수 있도록 한다.
- [55] 한편, 컨트롤러(170)는 영상을 표시하도록 디스플레이(180)를 제어할 수

있으며, 예를 들어 튜너(131)를 통해 입력되는 방송 영상, 또는 외부장치 인터페이스(135)를 통해 입력되는 외부 입력 영상, 또는 네트워크 인터페이스부를 통해 입력되는 영상, 또는 메모리(140)에 저장된 영상이 디스플레이(180)에서 표시되도록 제어할 수 있다. 이 경우, 디스플레이(180)에 표시되는 영상은 정지 영상 또는 동영상일 수 있으며, 2D 영상 또는 3D 영상일 수 있다.

- [56] 또한, 컨트롤러(170)는 디스플레이 장치(100) 내에 저장된 콘텐츠, 또는 수신된 방송 콘텐츠, 외부로부터 입력되는 외부 입력 콘텐츠가 재생되도록 제어할 수 있으며, 상기 콘텐츠는 방송 영상, 외부 입력 영상, 오디오 파일, 정지 영상, 접속된 웹 화면, 및 문서 파일 등 다양한 형태일 수 있다.
- [57] 무선 통신 인터페이스(173)는 유선 또는 무선 통신을 통해 외부 기기와 통신을 수행할 수 있다. 무선 통신 인터페이스(173)는 외부 기기와 근거리 통신(Short range communication)을 수행할 수 있다. 이를 위해, 무선 통신 인터페이스(173)는 블루투스(Bluetooth™), RFID(Radio Frequency Identification), 적외선 통신(Infrared Data Association; IrDA), UWB(Ultra Wideband), ZigBee, NFC(Near Field Communication), Wi-Fi(Wireless-Fidelity), Wi-Fi Direct, Wireless USB(Wireless Universal Serial Bus) 기술 중 적어도 하나를 이용하여, 근거리 통신을 지원할 수 있다. 이러한, 무선 통신 인터페이스(173)는 근거리 무선 통신망(Wireless Area Networks)을 통해 디스플레이 장치(100)와 무선 통신 시스템 사이, 디스플레이 장치(100)와 다른 디스플레이 장치(100) 사이, 또는 디스플레이 장치(100)와 디스플레이 장치(100, 또는 외부서버)가 위치한 네트워크 사이의 무선 통신을 지원할 수 있다. 근거리 무선 통신망은 근거리 무선 개인 통신망(Wireless Personal Area Networks)일 수 있다.
- [58] 여기에서, 다른 디스플레이 장치(100)는 본 발명에 따른 디스플레이 장치(100)와 데이터를 상호 교환하는 것이 가능한(또는 연동 가능한) 웨어러블 디바이스(wearable device, 예를 들어, 스마트워치(smartwatch), 스마트 글래스(smart glass), HMD(head mounted display)), 스마트폰과 같은 이동 단말기가 될 수 있다. 무선 통신 인터페이스(173)는 디스플레이 장치(100) 주변에, 통신 가능한 웨어러블 디바이스를 감지(또는 인식)할 수 있다.
- [59] 나아가, 컨트롤러(170)는 감지된 웨어러블 디바이스가 본 발명에 따른 디스플레이 장치(100)와 통신하도록 인증된 디바이스인 경우, 디스플레이 장치(100)에서 처리되는 데이터의 적어도 일부를, 무선 통신 인터페이스(173)를 통해 웨어러블 디바이스로 송신할 수 있다. 따라서, 웨어러블 디바이스의 사용자는, 디스플레이 장치(100)에서 처리되는 데이터를, 웨어러블 디바이스를 통해 이용할 수 있다.
- [60] 디스플레이(180)는 컨트롤러(170)에서 처리된 영상 신호, 데이터 신호, OSD 신호 또는 외부장치 인터페이스(135)에서 수신되는 영상 신호, 데이터 신호 등을 각각 R,G,B 신호로 변환하여 구동 신호를 생성할 수 있다.

- [61] 한편, 도 1에 도시된 디스플레이 장치(100)는 본 발명의 일실시예에 불과하므로, 도시된 구성요소들 중 일부는 실제 구현되는 디스플레이 장치(100)의 사양에 따라 통합, 추가, 또는 생략될 수 있다.
- [62] 즉, 필요에 따라 2 이상의 구성요소가 하나의 구성요소로 합쳐지거나, 혹은 하나의 구성요소가 2 이상의 구성요소로 세분되어 구성될 수 있다. 또한, 각 블록에서 수행하는 기능은 본 발명의 실시예를 설명하기 위한 것이며, 그 구체적인 동작이나 장치는 본 발명의 권리범위를 제한하지 아니한다.
- [63] 본 발명의 또 다른 실시예에 따르면, 디스플레이 장치(100)는 도 1에 도시된 바와 달리, 튜너(131)와 복조기(132)를 구비하지 않고 네트워크 인터페이스(133) 또는 외부장치 인터페이스(135)를 통해서 영상을 수신하여 재생할 수도 있다.
- [64] 예를 들어, 디스플레이 장치(100)는 방송 신호 또는 다양한 네트워크 서비스에 따른 콘텐츠들을 수신하기 위한 등과 같은 셋탑 박스 등과 같은 영상 처리 장치와 상기 영상 처리 장치로부터 입력되는 콘텐츠를 재생하는 콘텐츠 재생 장치로 분리되어 구현될 수 있다.
- [65] 이 경우, 이하에서 설명할 본 발명의 실시예에 따른 디스플레이 장치의 동작 방법은 도 1을 참조하여 설명한 바와 같은 디스플레이 장치(100)뿐 아니라, 상기 분리된 셋탑 박스 등과 같은 영상 처리 장치 또는 디스플레이(180) 및 오디오출력부(185)를 구비하는 콘텐츠 재생 장치 중 어느 하나에 의해 수행될 수도 있다.
- [66] 다음으로, 도 2 내지 도 3을 참조하여, 본 발명의 일 실시 예에 따른 원격제어장치에 대해 설명한다.
- [67] 도 2은 본 발명의 일 실시 예에 따른 원격제어장치의 블록도이고, 도 3은 본 발명의 일 실시 예에 따른 원격제어장치(200)의 실제 구성 예를 보여준다.
- [68] 먼저, 도 2를 참조하면, 원격제어장치(200)는 지문인식기(210), 무선통신회로(220), 사용자 입력 인터페이스(230), 센서(240), 출력 인터페이스(250), 전원공급회로(260), 메모리(270), 컨트롤러(280), 마이크론(290)을 포함할 수 있다.
- [69] 도 2을 참조하면, 무선통신회로(220)는 전술하여 설명한 본 발명의 실시 예들에 따른 디스플레이 장치 중 임의의 어느 하나와 신호를 송수신한다.
- [70] 원격제어장치(200)는 RF 통신규격에 따라 디스플레이 장치(100)와 신호를 송수신할 수 있는 RF 회로(221)을 구비하며, IR 통신규격에 따라 디스플레이 장치(100)와 신호를 송수신할 수 있는 IR 회로(223)을 구비할 수 있다. 또한, 원격제어장치(200)는 블루투스 통신규격에 따라 디스플레이 장치(100)와 신호를 송수신할 수 있는 블루투스 회로(225)를 구비할 수 있다. 또한, 원격제어장치(200)는 NFC(Near Field Communication) 통신 규격에 따라 디스플레이 장치(100)와 신호를 송수신할 수 있는 NFC 회로(227)을 구비하며, WLAN(Wireless LAN) 통신 규격에 따라 디스플레이 장치(100)와 신호를 송수신할 수 있는 WLAN 회로(229)을 구비할 수 있다.

- [71] 또한, 원격제어장치(200)는 디스플레이 장치(100)로 원격제어장치(200)의 움직임 등에 관한 정보가 담긴 신호를 무선통신회로(220)를 통해 전송한다.
- [72] 한편, 원격제어장치(200)는 디스플레이 장치(100)가 전송한 신호를 RF 회로(221)을 통하여 수신할 수 있으며, 필요에 따라 IR 회로(223)을 통하여 디스플레이 장치(100)로 전원 온/오프, 채널 변경, 볼륨 변경 등에 관한 명령을 전송할 수 있다.
- [73] 사용자 입력 인터페이스(230)는 키패드, 버튼, 터치 패드, 또는 터치 스크린 등으로 구성될 수 있다. 사용자는 사용자 입력 인터페이스(230)를 조작하여 원격제어장치(200)으로 디스플레이 장치(100)와 관련된 명령을 입력할 수 있다. 사용자 입력 인터페이스(230)가 하드키 버튼을 구비할 경우 사용자는 하드키 버튼의 푸쉬 동작을 통하여 원격제어장치(200)으로 디스플레이 장치(100)와 관련된 명령을 입력할 수 있다. 이에 대해서는 도 3을 참조하여 설명한다.
- [74] 도 3을 참조하면, 원격제어장치(200)는 복수의 버튼을 포함할 수 있다. 복수의 버튼은 지문 인식 버튼(212), 전원 버튼(231), 홈 버튼(232), 라이브 버튼(233), 외부 입력 버튼(234), 음량 조절 버튼(235), 음성 인식 버튼(236), 채널 변경 버튼(237), 확인 버튼(238) 및 뒤로 가기 버튼(239)을 포함할 수 있다.
- [75] 지문 인식 버튼(212)은 사용자의 지문을 인식하기 위한 버튼일 수 있다. 일 실시예로, 지문 인식 버튼(212)은 푸쉬 동작이 가능하며, 푸쉬 동작 및 지문 인식 동작을 수신할 수도 있다.
- [76] 전원 버튼(231)은 디스플레이 장치(100)의 전원을 온/오프 하기 위한 버튼일 수 있다.
- [77] 홈 버튼(232)은 디스플레이 장치(100)의 홈 화면으로 이동하기 위한 버튼일 수 있다.
- [78] 라이브 버튼(233)은 실시간 방송 프로그램을 디스플레이 하기 위한 버튼일 수 있다.
- [79] 외부 입력 버튼(234)은 디스플레이 장치(100)에 연결된 외부 입력을 수신하기 위한 버튼일 수 있다.
- [80] 음량 조절 버튼(235)은 디스플레이 장치(100)가 출력하는 음량의 크기를 조절하기 위한 버튼일 수 있다.
- [81] 음성 인식 버튼(236)은 사용자의 음성을 수신하고, 수신된 음성을 인식하기 위한 버튼일 수 있다.
- [82] 채널 변경 버튼(237)은 특정 방송 채널의 방송 신호를 수신하기 위한 버튼일 수 있다.
- [83] 확인 버튼(238)은 특정 기능을 선택하기 위한 버튼일 수 있고, 뒤로 가기 버튼(239)은 이전 화면으로 되돌아가기 위한 버튼일 수 있다.
- [84] 다시 도 2를 설명한다.
- [85] 사용자 입력 인터페이스(230)가 터치스크린을 구비할 경우 사용자는 터치스크린의 소프트웨어를 터치하여 원격제어장치(200)로 디스플레이

장치(100)와 관련된 명령을 입력할 수 있다. 또한, 사용자 입력 인터페이스(230)는 스크롤 키나, 조그 키 등 사용자가 조작할 수 있는 다양한 종류의 입력수단을 구비할 수 있으며 본 실시 예는 본 발명의 권리범위를 제한하지 아니한다.

- [86] 센서(240)는 자이로 센서(241) 또는 가속도 센서(243)를 구비할 수 있으며, 자이로 센서(241)는 원격제어장치(200)의 움직임에 관한 정보를 센싱할 수 있다.
- [87] 예를 들어, 자이로 센서(241)는 원격제어장치(200)의 동작에 관한 정보를 x,y,z 축을 기준으로 센싱할 수 있으며, 가속도 센서(243)는 원격제어장치(200)의 이동속도 등에 관한 정보를 센싱할 수 있다. 한편, 원격제어장치(200)는 거리측정센서를 더 구비할 수 있어, 디스플레이 장치(100)의 디스플레이(180)와의 거리를 센싱할 수 있다.
- [88] 출력 인터페이스(250)는 사용자 입력 인터페이스(230)의 조작에 대응하거나 디스플레이 장치(100)에서 전송한 신호에 대응하는 영상 또는 음성 신호를 출력할 수 있다.
- [89] 사용자는 출력 인터페이스(250)를 사용자 입력 인터페이스(230)의 조작 여부 또는 디스플레이 장치(100)의 제어 여부를 인지할 수 있다.
- [90] 예를 들어, 출력 인터페이스(250)는 사용자 입력 인터페이스(230)가 조작되거나 무선 통신부(225)를 통하여 디스플레이 장치(100)와 신호가 송수신되면 점등되는 LED(251), 진동을 발생하는 진동기(253), 음향을 출력하는 스피커(255), 또는 영상을 출력하는 디스플레이(257)를 구비할 수 있다.
- [91] 또한, 전원공급회로(260)는 원격제어장치(200)으로 전원을 공급하며, 원격제어장치(200)이 소정 시간 동안 움직이지 않은 경우 전원 공급을 중단함으로써 전원 낭비를 줄일 수 있다.
- [92] 전원공급회로(260)는 원격제어장치(200)에 구비된 소정 키가 조작된 경우에 전원 공급을 재개할 수 있다.
- [93] 메모리(270)는 원격제어장치(200)의 제어 또는 동작에 필요한 여러 종류의 프로그램, 애플리케이션 데이터 등이 저장될 수 있다.
- [94] 원격제어장치(200)가 디스플레이 장치(100)와 RF 회로(221)를 통하여 무선으로 신호를 송수신할 경우, 원격제어장치(200)과 디스플레이 장치(100)는 소정 주파수 대역을 통하여 신호를 송수신한다.
- [95] 원격제어장치(200)의 컨트롤러(280)는 원격제어장치(200)과 페어링된 디스플레이 장치(100)와 신호를 무선으로 송수신할 수 있는 주파수 대역 등에 관한 정보를 메모리(270)에 저장하고 참조할 수 있다.
- [96] 컨트롤러(280)는 원격제어장치(200)의 제어에 관련된 제반사항을 제어한다. 컨트롤러(280)는 사용자 입력 인터페이스(230)의 소정 키 조작에 대응하는 신호 또는 센서(240)에서 센싱한 원격제어장치(200)의 움직임에 대응하는 신호를 무선 통신부(225)를 통하여 디스플레이 장치(100)로 전송할 수 있다.
- [97] 또한, 원격제어장치(200)의 마이크로폰(290)은 음성을 획득할 수 있다.

- [98] 마이크론(290)은 복수 개로 구비될 수 있다.
- [99] 다음으로 도 4를 설명한다.
- [100] 도 4는 본 발명의 실시 예에 따라 원격 제어 장치를 활용하는 예를 보여준다.
- [101] 도 4의 (a)는 원격 제어 장치(200)에 대응하는 포인터(205)가 디스플레이(180)에 표시되는 것을 예시한다.
- [102] 사용자는 원격 제어 장치(200)를 상하, 좌우로 움직이거나 회전할 수 있다. 디스플레이 장치(100)의 디스플레이(180)에 표시된 포인터(205)는 원격 제어 장치(200)의 움직임에 대응한다. 이러한 원격 제어 장치(200)는, 도면과 같이, 3D 공간 상의 움직임에 따라 해당 포인터(205)가 이동되어 표시되므로, 공간 리모콘이라 명명할 수 있다.
- [103] 도 4의 (b)는 사용자가 원격 제어 장치(200)를 왼쪽으로 이동하면, 디스플레이 장치(100)의 디스플레이(180)에 표시된 포인터(205)도 이에 대응하여 왼쪽으로 이동하는 것을 예시한다.
- [104] 원격 제어 장치(200)의 센서를 통하여 감지된 원격 제어 장치(200)의 움직임에 관한 정보는 디스플레이 장치(100)로 전송된다. 디스플레이 장치(100)는 원격 제어 장치(200)의 움직임에 관한 정보로부터 포인터(205)의 좌표를 산출할 수 있다. 디스플레이 장치(100)는 산출한 좌표에 대응하도록 포인터(205)를 표시할 수 있다.
- [105] 도 4의 (c)는, 원격 제어 장치(200) 내의 특정 버튼을 누른 상태에서, 사용자가 원격 제어 장치(200)를 디스플레이(180)에서 멀어지도록 이동하는 경우를 예시한다. 이에 의해, 포인터(205)에 대응하는 디스플레이(180) 내의 선택 영역이 좁아져 확대 표시될 수 있다.
- [106] 이와 반대로, 사용자가 원격 제어 장치(200)를 디스플레이(180)에 가까워지도록 이동하는 경우, 포인터(205)에 대응하는 디스플레이(180) 내의 선택 영역이 좁아져 축소 표시될 수 있다.
- [107] 한편, 원격 제어 장치(200)가 디스플레이(180)에서 멀어지는 경우, 선택 영역이 좁아져지고, 원격 제어 장치(200)가 디스플레이(180)에 가까워지는 경우, 선택 영역이 좁아질 수도 있다.
- [108] 또한, 원격 제어 장치(200) 내의 특정 버튼을 누른 상태에서는 상하, 좌우 이동의 인식이 배제될 수 있다. 즉, 원격 제어 장치(200)가 디스플레이(180)에서 멀어지거나 접근하도록 이동하는 경우, 상, 하, 좌, 우 이동은 인식되지 않고, 앞뒤 이동만 인식되도록 할 수 있다. 원격 제어 장치(200) 내의 특정 버튼을 누르지 않은 상태에서는, 원격 제어 장치(200)의 상, 하, 좌, 우 이동에 따라 포인터(205)만 이동하게 된다.
- [109] 한편, 포인터(205)의 이동속도나 이동방향은 원격 제어 장치(200)의 이동속도나 이동방향에 대응할 수 있다.
- [110] 한편, 본 명세서에서의 포인터는, 원격 제어 장치(200)의 동작에 대응하여, 디스플레이(180)에 표시되는 오브젝트를 의미한다. 따라서, 포인터(205)로

도면에 도시된 화살표 형상 외에 다양한 형상의 오브젝트가 가능하다. 예를 들어, 점, 커서, 프롭포트, 두꺼운 외곽선 등을 포함하는 개념일 수 있다. 그리고, 포인터(205)가 디스플레이(180) 상의 가로축과 세로축 중 어느 한 지점(point)에 대응하여 표시되는 것은 물론, 선(line), 면(surface) 등 복수 지점에 대응하여 표시되는 것도 가능하다.

- [111] 도 5는 종래 기술에 따른 음성 입력의 끝점 인식을 위한 과정을 설명하는 도면이다.
- [112] 도 5를 참조하면, 디스플레이 장치(100)의 음성 입력 인터페이스(101)는 음성 입력을 획득하고, 획득된 음성 입력에 상응하는 음성 신호를 펄스 폭 변조(Pulse code Modulation, PCM) 신호로 변환할 수 있다.
- [113] 음성 입력 인터페이스(101)는 변환된 PCM 신호를 인공지능(Artificial Intelligence, AI) 클라이언트(103)에 전달할 수 있다.
- [114] AI 클라이언트(103)는 PCM 신호를 STT(Speech To Text) 서버(50)에 전송할 수 있다.
- [115] STT(Speech To Text) 서버(50)는 에너지 레벨 검출기(51) 및 ASR(Automatic Speech Recognition, 자동 음성 인식) 엔진(53)을 포함할 수 있다.
- [116] 에너지 레벨 검출기(51)는 PCM 신호의 에너지 레벨을 검출할 수 있다.
- [117] 에너지 레벨 검출기(51)는 AI 클라이언트(103)로부터 수신된 PCM 신호를 ASR 엔진(53)에 전달할 수 있다.
- [118] ASR 엔진(53)은 디지털 형태의 PCM 신호를 텍스트로 변환할 수 있다.
- [119] STT 서버(50)는 PCM 신호의 에너지 레벨 및 변환된 텍스트를 AI 클라이언트(103)에 전달할 수 있다.
- [120] AI 클라이언트(103)는 PCM 신호의 에너지 레벨이 기 설정된 레벨 미만인 경우, 사용자의 음성 입력이 종료된 것으로 판단할 수 있다. 즉, AI 클라이언트(103)는 PCM 신호의 에너지 레벨이 기 설정된 레벨 미만인 경우, 발화 끝점이 검출된 것으로 판단할 수 있다.
- [121] AI 클라이언트(103)는 발화 끝점이 검출된 것으로 판단한 경우, 텍스트를 NLP(Natural Language Processing, 자연어 처리) 서버(60)에 전송할 수 있다.
- [122] NLP 서버(60)는 수신된 텍스트에 기반하여 텍스트의 의도 분석 결과를 포함하는 분석 결과 정보를 획득하고, 획득된 분석 결과 정보를 AI 클라이언트(103)에 전송할 수 있다.
- [123] 이와 같이, 종래 기술에 따르면, 음성 입력에 상응하는 PCM 신호의 에너지 레벨을 이용하여 발화 끝점을 검출하였다.
- [124] 그러나, 실제 사용자의 발화가 종료되었음에도 주위의 잡음이나 노이즈가 인식됨에 따라 음성 인식의 수행이 지속될 수 있다.
- [125] 이에 따라, 사용자는 원하지 않은 음성 인식의 수행 결과를 제공받을 수 있어, 상당한 불편함이 발생되고 있다.
- [126] 이하의 실시 예에서는 사용자의 음성 입력에 대한 발화 끝점을 보다 정확히

- 검출할 수 있는 방법을 제안하고자 한다.
- [127] 도 6은 본 개시의 일 실시 예에 따른 시스템의 동작 방법을 설명하는 래더 다이어그램이다.
- [128] 도 6을 참조하면, 시스템은 디스플레이 장치(100), STT 서버(600) 및 NLP 서버(700)를 포함할 수 있다.
- [129] 디스플레이 장치(100)의 컨트롤러(170)는 사용자가 발화한 음성 입력을 획득한다(S601).
- [130] 일 실시 예에서 컨트롤러(170)는 원격 제어 장치(200)로부터 음성 입력을 수신할 수 있다. 컨트롤러(170)는 사용자 입력 인터페이스(150)를 통해 원격 제어 장치(200)로부터 음성 입력을 수신할 수 있다. 사용자는 원격 제어 장치(200)에 음성을 발화할 수 있다.
- [131] 또 다른 실시 예에서 컨트롤러(170)는 디스플레이 장치(100)의 외부장치 인터페이스(135)와 연결된 마이크로부터 음성 입력을 수신할 수 있다.
- [132] 또 다른 실시 예에서 컨트롤러(170)는 디스플레이 장치(100)에 구비된 마이크를 통해 음성 입력을 수신할 수 있다.
- [133] 디스플레이 장치(100)의 컨트롤러(170)는 획득된 음성 입력에 상응하는 음성 신호를 PCM 신호로 변환한다(S603).
- [134] 컨트롤러(170)는 아날로그 형태의 음성 신호를 디지털 형태의 펄스 코드 변조 신호로 변환할 수 있다. 이를 위해 컨트롤러(170)는 PCM 변환기를 더 구비할 수 있다.
- [135] 디스플레이 장치(100)의 컨트롤러(170)는 변환된 PCM 신호를 네트워크 인터페이스(133)를 통해 STT 서버(600)에 전송한다(S605).
- [136] 컨트롤러(170)는 네트워크 인터페이스(133)를 통해 PCM 신호를 실시간 또는 주기적으로 전송할 수 있다. 음성 입력이 연속적으로 수신될 수 있기 때문이다.
- [137] STT 서버(600)는 수신된 PCM 신호에 기반하여 PCM 신호의 에너지 레벨, 음성 입력에 상응하는 텍스트 및 발화 끝점 정보를 획득한다(S607).
- [138] PCM 신호의 에너지 레벨은 PCM 신호의 크기를 나타내며, 단위는 dB로 표현될 수 있다.
- [139] STT 서버(600)는 STT 엔진(또는 ASR 엔진)을 통해 PCM 신호를 텍스트 데이터로 변환하여 음성 입력에 상응하는 텍스트를 획득할 수 있다.
- [140] 일 실시 예에서 발화 끝점 정보는 변환된 텍스트의 신뢰도 스코어일 수 있다.
- [141] STT 서버(600)는 후술할 끝점 검출기를 통해 텍스트의 신뢰도 스코어를 계산할 수 있다. 끝점 검출기는 텍스트의 각 단어에 품사를 부여하고, 부여 결과에 따른 텍스트 패턴을 분석하여 텍스트의 끝점을 검출할 수 있다. 이에 대해서는 후술한다.
- [142] 신뢰도 스코어는 음성 입력의 발화가 종료되었는지 여부를 판단하는 기준이 될 수 있다.
- [143] 신뢰도 스코어는 0 내지 1 사이의 값을 가질 수 있다. 신뢰도 스코어가 클수록

발화의 끝점이 검출될 확률이 커지고, 신뢰도 스코어가 작을수록 발화의 끝점이 검출될 확률이 작아질 수 있다.

- [144] 또 다른 실시 예에서 발화 끝점 정보는 변환된 텍스트에 상응하는 음성 입력이 발화 끝점인지 여부를 나타내는 정보를 포함할 수 있다. 예를 들어, 발화 끝점 정보는 발화의 끝점임을 나타내는 값(End) 또는 발화가 종료되지 않았음을 나타내는 값(Continuous)을 포함할 수 있다.
- [145] STT 서버(600)는 PCM 신호의 에너지 레벨, 음성 입력에 상응하는 텍스트 및 발화 끝점 정보를 디스플레이 장치(100)에 전송한다(S609).
- [146] 디스플레이 장치(100)의 컨트롤러(170)는 PCM 신호의 에너지 레벨 및 발화 끝점 정보에 기반하여 사용자의 발화가 종료되었는지를 판단한다(S611).
- [147] 일 실시 예에서 컨트롤러(170)는 발화 끝점 정보에 포함된 신뢰도 스코어가 기 설정된 스코어 이상이고, PCM 신호의 에너지 레벨이 기 설정된 제1 레벨 미만인 경우, 사용자의 발화가 종료된 것으로 판단할 수 있다. 여기서, 기 설정된 스코어는 0.8이고, 기 설정된 제1 레벨은 12dB일 수 있으나, 예시에 불과하다.
- [148] 신뢰도 스코어는 음성 입력의 발화 끝점일 확률을 나타내는 스코어일 수 있다.
- [149] 또 다른 실시 예에서 컨트롤러(170)는 발화 끝점 정보에 음성 입력의 발화 끝점임을 나타내는 값(End)이 포함되고, PCM 신호의 에너지 레벨이 기 설정된 제1 레벨 미만인 경우, 사용자의 발화가 종료된 것으로 판단할 수 있다.
- [150] 일 실시 예에서 컨트롤러(170)는 발화 끝점 정보에 포함된 신뢰도 스코어가 기 설정된 스코어 이상이고, PCM 신호의 에너지 레벨이 기 설정된 제2 레벨 이상인 경우, 사용자의 발화가 종료되지 않은 것으로 판단할 수 있다. 여기서, 기 설정된 스코어는 0.8이고, 제2 레벨은 50dB 일 수 있다.
- [151] 즉, 컨트롤러(170)는 발화 끝점 정보에 포함된 신뢰도 스코어가 기 설정된 스코어 이상이라도 PCM 신호의 에너지 레벨이 매우 크면 발화가 종료되지 않은 것으로 판단할 수 있다.
- [152] 또 다른 실시 예에서 컨트롤러(170)는 발화 끝점 정보에 음성 입력의 발화 끝점임을 나타내는 값(End)이 포함되고, PCM 신호의 에너지 레벨이 기 설정된 제2 레벨 이상인 경우, 사용자의 발화가 종료되지 않은 것으로 판단할 수 있다.
- [153] 즉, 컨트롤러(170)는 발화 끝점 정보에 발화 종료를 나타내는 값이 포함되더라도 PCM 신호의 에너지 레벨이 매우 크면 발화가 종료되지 않은 것으로 판단할 수 있다.
- [154] 디스플레이 장치(100)의 컨트롤러(170)는 사용자의 발화가 종료된 것으로 판단된 경우, 발화 모드를 종료한다(S613).
- [155] 일 실시 예에서 발화 모드는 음성 인식 모드로 명명될 수 있다.
- [156] 발화 모드는 사용자가 발화한 음성을 수신하여 수신된 음성에 대한 분석 결과를 출력하는 모드일 수 있다.
- [157] 발화 모드의 종료는 음성 입력에 대한 음성 인식 결과를 출력하기까지 추가로 입력되는 음성 입력을 무시하는 것을 나타낼 수 있다.

- [158] 즉, 컨트롤러(170)는 음성 입력을 수신하여도 이에 대한 처리를 수행하지 않을 수 있다.
- [159] 발화 모드의 종료는 음성 인식을 비활성화하는 동작일 수 있다.
- [160] 디스플레이 장치(100)의 컨트롤러(170)는 텍스트 및 텍스트의 의도 분석 요청을 NLP 서버(700)에 전송한다(S615).
- [161] NLP 서버(700)는 디스플레이 장치(100)로부터 수신된 텍스트에 기반하여 텍스트의 의도를 분석할 수 있다. NLP 서버(700)는 의도 분석 결과를 나타내는 분석 결과 정보를 생성할 수 있다.
- [162] NLP 서버(700)는 텍스트 데이터에 대해, 형태소 분석 단계, 구문 분석 단계, 화행 분석 단계, 대화 처리 단계를 순차적으로, 수행하여, 분석 결과 정보를 생성할 수 있다.
- [163] 형태소 분석 단계는 사용자가 발화한 음성에 대응하는 텍스트 데이터를 의미를 지닌 가장 작은 단위인 형태소 단위로 분류하고, 분류된 각 형태소가 어떤 품사를 가지는지를 결정하는 단계이다.
- [164] 구문 분석 단계는 형태소 분석 단계의 결과를 이용하여, 텍스트 데이터를 명사구, 동사구, 형용사 구 등으로 구분하고, 구분된 각 구들 사이에, 어떠한 관계가 존재하는지를 결정하는 단계이다.
- [165] 구문 분석 단계를 통해, 사용자가 발화한 음성의 주어, 목적어, 수식어들이 결정될 수 있다.
- [166] 화행 분석 단계는 구문 분석 단계의 결과를 이용하여, 사용자가 발화한 음성에 대한 의도를 분석하는 단계이다. 구체적으로, 화행 분석 단계는 사용자가 질문을 하는 것인지, 요청을 하는 것인지, 단순한 감정 표현을 하는 것인지와 같은 문장의 의도를 결정하는 단계이다.
- [167] 대화 처리 단계는 화행 분석 단계의 결과를 이용하여, 사용자의 발화에 대해 대답을 할지, 호응을 할지, 추가 정보를 문의하는 질문을 할지를 판단하는 단계이다.
- [168] NLP 서버(700)는 대화 처리 단계 후, 사용자가 발화한 의도에 대한 답변, 호응, 추가 정보 문의 중 하나 이상을 포함하는 분석 결과 정보를 생성할 수 있다.
- [169] NLP 서버(700)는 사용자의 발화 의도에 맞는 정보를 검색하기 위해, 검색 서버(미도시)에 검색 요청을 전송하고, 검색 요청에 상응하는 검색 정보를 수신할 수 있다.
- [170] 사용자의 발화 의도가 콘텐츠의 검색인 경우, 검색 정보는 검색된 콘텐츠에 대한 정보를 포함할 수 있고, 분석 결과 정보는 검색 정보를 포함할 수 있다.
- [171] 디스플레이 장치(100)의 컨트롤러(170)는 네트워크 인터페이스(133)를 통해 NLP 서버(700)로부터 텍스트의 의도 분석을 나타내는 분석 결과 정보를 수신한다(S617).
- [172] 디스플레이 장치(100)의 컨트롤러(170)는 수신된 분석 결과 정보를 출력한다(S619).

- [173] 일 실시 예에서 컨트롤러(170)는 분석 결과 정보를 디스플레이(180)를 통해 표시할 수 있다.
- [174] 또 다른 실시 예에서 컨트롤러(170)는 분석 결과 정보를 스피커(185)를 통해 오디오 형태로 출력할 수 있다.
- [175] 이와 같이, 본 개시의 실시 예에 따르면, 음성 입력에 대한 텍스트 분석 및 에너지 레벨을 이용하여 발화의 끝을 정확히 인식할 수 있다. 이에 따라 불필요한 음성 인식이 수행되는 것이 방지될 수 있다.
- [176] 다음으로, 도 7을 설명한다.
- [177] 도 7은 본 개시의 일 실시 예에 따른 발화의 끝점을 검출하여 의도 분석을 수행하는 과정을 설명하는 도면이다.
- [178] 음성 인식 시스템(70)은 디스플레이 장치(100), STT 서버(600) 및 NLP 서버(700)를 포함할 수 있다.
- [179] STT 서버(600) 및 NLP 서버(700)는 하나의 서버로 구성될 수도 있다.
- [180] 도 7을 참조하면, 디스플레이 장치(100)는 음성 입력 인터페이스(171) 및 AI 클라이언트(172)를 포함할 수 있다.
- [181] 음성 입력 인터페이스(171) 및 AI 클라이언트(172)는 컨트롤러(170)에 포함될 수 있다.
- [182] AI 클라이언트(172)는 네트워크 인터페이스(133)에 포함될 수도 있다.
- [183] 음성 입력 인터페이스(171)는 원격 제어 장치(200) 또는 디스플레이 장치(100)에 구비된 마이크를 통해 음성 입력을 수신할 수 있다.
- [184] 음성 입력 인터페이스(171)는 수신된 음성 입력에 상응하는 음성 신호를 PCM 신호로 변환할 수 있다.
- [185] AI 클라이언트(172)는 변환된 PCM 신호를 STT 서버(600)에 전송할 수 있다.
- [186] STT 서버(600)는 에너지 레벨 검출기(610), ASR 엔진(630) 및 끝점 검출기(650)를 포함할 수 있다.
- [187] 에너지 레벨 검출기(610)는 AI 클라이언트(172)로부터 수신된 음성 입력에 대응하는 PCM 신호의 에너지 레벨을 검출할 수 있다. 에너지 레벨의 단위는 dB 일 수 있다.
- [188] 에너지 레벨 검출기(610)는 검출된 에너지 레벨을 AI 클라이언트(172)에 전송할 수 있다.
- [189] ASR 엔진(630)은 PCM 신호를 텍스트 데이터로 변환할 수 있다. ASR 엔진(630)은 변환된 텍스트 데이터를 AI 클라이언트(172)에 전송할 수 있다.
- [190] 끝점 검출기(650)는 ASR 엔진(630)에서 출력된 텍스트 데이터에 기초하여 발화 끝점 정보를 생성할 수 있다.
- [191] 일 실시 예에서 발화 끝점 정보는 음성 입력의 발화가 끝점인지 여부를 나타내는 정보를 포함할 수 있다. 이 경우, 발화 끝점 정보는 발화의 끝점임을 식별하는 값(End), 발화의 끝점이 아님을 식별하는 값(Continue) 중 어느 하나의 값을 포함할 수 있다.

- [192] 또 다른 실시 예에서 발화 끝점 정보는 음성 입력의 발화의 끝점에 대한 신뢰도 스코어를 포함할 수 있다. 신뢰도 스코어는 발화의 끝점에 가까울수록 그 값이 커지고, 발화의 끝점에 멀어질수록 그 값이 작아질 수 있다.
- [193] 신뢰도 스코어는 0 내지 1 사이의 값을 가질 수 있다.
- [194] 끝점 검출기(650)는 텍스트 데이터의 품사 분석, 패턴 분석, 기호 분석을 통해 발화 끝점 정보를 생성할 수 있다.
- [195] 끝점 검출기(650)가 발화 끝점 정보를 생성하는 과정을 이하의 도면을 참조하여 설명한다.
- [196] 도 8은 본 개시의 일 실시 예에 따라 끝점 검출기가 발화 끝점 정보를 생성하는 과정을 설명하는 도면이다.
- [197] 도 8을 참조하면, 끝점 검출기(650)는 음성 입력에 상응하는 텍스트의 품사 분석(S801), 패턴 분석(S803) 및 기호 분석(S805)을 순차적으로 수행할 수 있다.
- [198] 예를 들어, 끝점 검출기(650)는 <볼륨 올려>라는 텍스트에 대한 발화 끝점 정보를 생성함을 가정한다.
- [199] 품사 분석(S801)은 텍스트를 구성하는 각 단어에 품사를 붙이는 과정일 수 있다. 끝점 검출기(650)는 <볼륨>이라는 단어에 명사, <올리>라는 어간에 동사를 매칭시킬 수 있다.
- [200] 또한, 끝점 검출기(650)는 <ㄱ>라는 어미도 분류할 수 있다.
- [201] 패턴 분석(S803)은 품사 분석 후 텍스트의 패턴을 분석하여 텍스트의 끝을 인식하는 과정일 수 있다. 예를 들어, 끝점 검출기(650)는 <올리>라는 동사의 어미<ㄱ>가 해당 텍스트의 끝인지 여부를 판단할 수 있다.
- [202] 일 실시 예에서 끝점 검출기(650)는 <올리>라는 동사에 뒤따르는 어미<ㄱ>가 텍스트의 끝을 나타내는 경우, 발화의 끝을 식별하는 값을 생성할 수 있다.
- [203] 또 다른 실시 예에서 끝점 검출기(650)는 <올리>라는 동사에 뒤따르는 어미<ㄱ>가 텍스트의 끝을 나타낼 확률이 큰 경우, 발화 끝점에 대한 신뢰도 스코어를 큰 값으로 산출할 수 있다.
- [204] 기호 분석(S805)은 언어별 기호(마침표, 물음표, 느낌표) 등을 참고하여 텍스트의 끝을 판단하는 과정일 수 있다.
- [205] 기호 분석(S805)은 optional한 과정일 수 있다.
- [206] 끝점 검출기(650)는 <올려>라는 단어를 인식하고, 해당 단어에 기 매칭되어 있는 기호인 느낌표를 인식하여 발화의 끝점을 인식할 수 있다.
- [207] 다시, 도 7을 설명한다.
- [208] STT 서버(600)는 디스플레이 장치(100)와 통신을 수행하기 위한 통신 인터페이스(미도시)를 더 구비할 수 있다. STT 서버(600)의 통신 인터페이스는 PCM 신호의 에너지 레벨, 변환된 텍스트 데이터 및 발화 끝점 정보를 AI 클라이언트(172)에 전송할 수 있다.
- [209] 일 실시 예에서 AI 클라이언트(172)는 PCM 신호의 에너지 레벨이 기 설정된 제1 레벨 미만이고, 발화 끝점 정보에 발화 끝점을 나타내는 값이 포함된 경우,

- 발화 끝점이 검출된 것으로 판단할 수 있다.
- [210] 또 다른 실시 예에서 AI 클라이언트(172)는 PCM 신호의 에너지 레벨이 일정 시간 동안 기 설정된 제1 레벨 미만이고, 발화 끝점 정보에 발화 끝점을 나타내는 값이 포함된 경우, 발화 끝점이 검출된 것으로 판단할 수 있다.
- [211] 또 다른 실시 예에서 일 실시 예에서 AI 클라이언트(172)는 PCM 신호의 에너지 레벨이 기 설정된 제1 레벨 미만이고, 발화 끝점 정보에 포함된 신뢰도 스코어가 기 설정된 스코어 이상인 경우, 발화 끝점이 검출된 것으로 판단할 수 있다.
- [212] 또 다른 실시 예에서 일 실시 예에서 AI 클라이언트(172)는 PCM 신호의 에너지 레벨이 일정 시간 동안 기 설정된 제1 레벨 미만이고, 발화 끝점 정보에 포함된 신뢰도 스코어가 기 설정된 스코어 이상인 경우, 발화 끝점이 검출된 것으로 판단할 수 있다.
- [213] AI 클라이언트(172)는 발화 끝점이 검출된 것으로 판단한 경우, 음성 입력 인터페이스(171)에 발화 모드의 종료를 위한 명령을 전달할 수 있다.
- [214] 음성 입력 인터페이스(171)는 수신된 명령에 따라 발화 모드를 종료할 수 있다. 음성 입력 인터페이스(171)는 기동어가 재 인식될 때까지, 수신된 음성 입력을 무시할 수 있다.
- [215] 또한, AI 클라이언트(172)는 발화 끝점이 검출된 것으로 판단된 경우, STT 서버(600)로부터 수신된 텍스트를 NLP 서버(700)에 전송할 수 있다.
- [216] AI 클라이언트(172)는 NLP 서버(700)로부터 텍스트에 대한 의도 분석 결과를 나타내는 분석 결과 정보를 수신하고, 수신된 분석 결과 정보를 출력할 수 있다.
- [217] 도 9는 본 개시의 또 다른 실시 예에 따른 시스템의 동작 방법을 설명하는 래더 다이어그램이다.
- [218] 도 6의 실시 예와 다른 점은 끝점 검출기가 STT 서버(600)가 아닌 NLP 서버(700)에 포함되는 것이다.
- [219] 도 9의 각 스텝에서 도 6의 실시 예와 중복되는 부분의 자세한 설명은 생략한다.
- [220] 디스플레이 장치(100)의 컨트롤러(170)는 사용자가 발화한 음성 입력을 획득한다(S901).
- [221] 디스플레이 장치(100)의 컨트롤러(170)는 획득된 음성 입력에 상응하는 음성 신호를 PCM 신호로 변환한다(S903).
- [222] 디스플레이 장치(100)의 컨트롤러(170)는 변환된 PCM 신호를 네트워크 인터페이스(133)를 통해 STT 서버(600)에 전송한다(S905).
- [223] 컨트롤러(170)는 네트워크 인터페이스(133)를 통해 PCM 신호를 실시간 또는 주기적으로 전송할 수 있다. 음성 입력은 시간의 흐름에 따라 연속적으로 수신될 수 있기 때문이다.
- [224] STT 서버(600)는 수신된 PCM 신호에 기반하여 PCM 신호의 에너지 레벨 및 음성 입력에 상응하는 텍스트를 획득한다(S907).
- [225] STT 서버(600)는 PCM 신호의 에너지 레벨 및 음성 입력에 상응하는 텍스트를 디스플레이 장치(100)에 전송한다(S909).

- [226] 디스플레이 장치(100)의 컨트롤러(170)는 수신된 텍스트를 NLP 서버(700)로 전송한다(S911).
- [227] 일 실시 예에서 컨트롤러(170)는 텍스트의 의도 분석 및 발화 끝점 인식을 문의하기 위해 네트워크 인터페이스(133)를 통해 텍스트를 NLP 서버(700)에 전송할 수 있다.
- [228] NLP 서버(700)는 디스플레이 장치(100)로부터 수신된 텍스트에 기반하여 발화 끝점 정보 및 분석 결과 정보를 생성하여 디스플레이 장치(100)에 전송한다(S913).
- [229] NLP 서버(700)는 끝점 검출기를 통해 수신된 텍스트에 대한 발화 끝점 정보를 생성할 수 있다.
- [230] 일 실시 예에서 NLP 서버(700)는 텍스트의 의도 분석을 수행하여 분석 결과 정보를 생성할 수 있다.
- [231] 또 다른 실시 예에서 NLP 서버(700)는 발화 끝점 정보에 음성 입력의 발화의 끝점임을 나타내는 값(End)이 포함되어 있는 경우 또는 신뢰도 스코어가 기 설정된 스코어 이상인 경우에만 분석 결과 정보를 디스플레이 장치(100)에 전송할 수 있다. 발화가 끝나지 않은 경우 분석 결과 정보는 부정확할 수 있기 때문이다.
- [232] 또 다른 실시 예에서 NLP 서버(700)는 발화 끝점 정보에 음성 입력의 발화의 끝점임을 나타내는 값(End)이 포함되어 있는 경우 또는 신뢰도 스코어가 기 설정된 스코어 이상인 경우에만 분석 결과 정보를 생성할 수 있다.
- [233] 디스플레이 장치(100)의 컨트롤러(170)는 PCM 신호의 에너지 레벨 및 발화 끝점 정보에 기반하여 사용자의 발화가 종료되었는지를 판단한다(S915).
- [234] 일 실시 예에서 컨트롤러(170)는 발화 끝점 정보에 포함된 신뢰도 스코어가 기 설정된 스코어 이상이고, PCM 신호의 에너지 레벨이 기 설정된 제1 레벨 미만인 경우, 사용자의 발화가 종료된 것으로 판단할 수 있다. 여기서, 기 설정된 스코어는 0.8이고, 기 설정된 레벨은 12dB일 수 있으나, 예시에 불과하다.
- [235] 또 다른 실시 예에서 컨트롤러(170)는 발화 끝점 정보에 음성 입력의 발화 끝점임을 나타내는 값(End)이 포함되고, PCM 신호의 에너지 레벨이 기 설정된 제1 레벨 미만인 경우, 사용자의 발화가 종료된 것으로 판단할 수 있다.
- [236] 디스플레이 장치(100)의 컨트롤러(170)는 사용자의 발화가 종료된 것으로 판단된 경우, 발화 모드를 종료한다(S917).
- [237] 디스플레이 장치(100)의 컨트롤러(170)는 수신된 분석 결과 정보를 출력한다(S919).
- [238] 도 10은 본 개시의 또 다른 실시 예에 따른 발화의 끝점을 검출하여 의도 분석을 수행하는 과정을 설명하는 도면이다.
- [239] 음성 인식 시스템(1000)은 디스플레이 장치(100), STT 서버(600) 및 NLP 서버(700)를 포함할 수 있다.
- [240] STT 서버(600) 및 NLP 서버(700)는 하나의 서버로 구성될 수도 있다.

- [241] 도 10을 참조하면, 디스플레이 장치(100)는 음성 입력 인터페이스(171) 및 AI 클라이언트(172)를 포함할 수 있다.
- [242] 음성 입력 인터페이스(171) 및 AI 클라이언트(172)는 컨트롤러(170)에 포함될 수 있다.
- [243] AI 클라이언트(172)는 네트워크 인터페이스(133)에 포함될 수도 있다.
- [244] 음성 입력 인터페이스(171)는 원격 제어 장치(200) 또는 디스플레이 장치(100)에 구비된 마이크를 통해 음성 입력을 수신할 수 있다.
- [245] 음성 입력 인터페이스(171)는 수신된 음성 입력에 상응하는 음성 신호를 PCM 신호로 변환할 수 있다.
- [246] AI 클라이언트(172)는 변환된 PCM 신호를 STT 서버(600)에 전송할 수 있다.
- [247] STT 서버(600)는 에너지 레벨 검출기(610) 및 ASR 엔진(630)를 포함할 수 있다.
- [248] 에너지 레벨 검출기(610)는 AI 클라이언트(172)로부터 수신된 음성 입력에 대응하는 PCM 신호의 에너지 레벨을 검출할 수 있다. 에너지 레벨의 단위는 dB 일 수 있다.
- [249] 에너지 레벨 검출기(610)는 검출된 에너지 레벨을 AI 클라이언트(172)에 전송할 수 있다.
- [250] ASR 엔진(630)은 PCM 신호를 텍스트 데이터로 변환할 수 있다. ASR 엔진(630)은 변환된 텍스트 데이터를 AI 클라이언트(172)에 전송할 수 있다.
- [251] 도 7의 실시 예와 다른 점은 끝점 검출기가 STT 서버(600)가 아닌 NLP 서버(700)에 구비되는 점이다.
- [252] NLP 서버(700)는 AI 클라이언트(172)로부터 수신된 텍스트에 기초하여 발화 끝점 정보 및 분석 결과 정보를 생성할 수 있다.
- [253] NLP 서버(700)는 도 8에서 설명된 끝점 검출기(650)와 동일한 기능을 수행하는 끝점 검출기(710)를 포함할 수 있다.
- [254] 일 실시 예에서 발화 끝점 정보는 음성 입력의 발화가 끝점인지 여부를 나타내는 정보를 포함할 수 있다. 이 경우, 발화 끝점 정보는 발화의 끝점임을 식별하는 값(End), 발화의 끝점이 아님을 식별하는 값(Continue) 중 어느 하나의 값을 포함할 수 있다.
- [255] 또 다른 실시 예에서 발화 끝점 정보는 음성 입력의 발화의 끝점에 대한 신뢰도 스코어를 포함할 수 있다. 신뢰도 스코어는 발화의 끝점에 가까울수록 그 값이 커지고, 발화의 끝점에 멀어질수록 그 값이 작아질 수 있다.
- [256] 신뢰도 스코어는 0 내지 1 사이의 값을 가질 수 있다.
- [257] 끝점 검출기(710)는 도 8에 도시된 바와 같이, 텍스트 데이터의 품사 분석, 패턴 분석, 기호 분석을 통해 발화 끝점 정보를 생성할 수 있다. 추가적으로, 끝점 검출기(710)는 텍스트에 대한 의도 분석 결과 정보를 참고하여 발화 끝점 정보를 생성할 수도 있다.
- [258] 예를 들어, 끝점 검출기(710)는 텍스트 데이터의 품사 분석, 패턴 분석, 기호 분석에 기반하여 생성된 신뢰도 스코어가 0.7이고, 의도 분석 결과 정보를 통해

사용자의 의도가 추출된 경우, 신뢰도 스코어에 0.1을 가산할 수 있다.

[259] 이에 따라 신뢰도 스코어는 0.8이 될 수 있다.

[260] 일 실시 예에서 AI 클라이언트(172)는 STT 서버(600)로부터 수신된 에너지 레벨이 기 설정된 제1 레벨 미만이고, NLP 서버(700)로부터 수신된 발화 끝점 정보에 발화 끝점을 나타내는 값이 포함된 경우, 사용자의 발화가 종료된 것으로 판단할 수 있다.

[261] 또 다른 실시 예에서 AI 클라이언트(172)는 에너지 레벨이 기 설정된 제1 레벨 미만이고, 발화 끝점 정보에 포함된 신뢰도 스코어가 기 설정된 스코어 이상인 경우, 사용자의 발화가 종료된 것으로 판단할 수 있다.

[262] 또 다른 실시 예에서 AI 클라이언트(172)는 발화 끝점 정보에 발화 끝점을 나타내지 않는 값(Continue)이 포함되어 있더라도, 에너지 레벨이 기 설정된 제3 레벨 미만이고, 분석 결과 정보를 수신한 경우, 사용자의 발화가 종료된 것으로 판단할 수 있다. 제3 레벨은 6dB 일 수 있다.

[263] 또 다른 실시 예에서 AI 클라이언트(172)는 발화 끝점 정보에 발화 끝점을 나타내는 값(Continue)이 포함되어 있더라도, 에너지 레벨이 기 설정된 제2 레벨 이상인 경우, 사용자의 발화가 종료되지 않은 것으로 판단할 수 있다. 제2 레벨은 50dB 일 수 있다.

[264] 또 다른 실시 예에서 AI 클라이언트(172)는 발화 끝점 정보에 발화 끝점을 나타내는 값(Continue)이 포함되어 있고, 분석 결과 정보를 수신한 경우, 에너지 레벨이 기 설정된 제2 레벨 이상이더라도, 사용자의 발화가 종료된 것으로 판단할 수 있다.

[265] 이와 같이, 본 개시의 실시 예에 따르면, 사용자의 음성 입력에 대한 에너지 레벨, 발화 끝점 정보 및 분석 결과 정보를 사용하여 사용자의 발화가 종료되었는지가 정확히 판단될 수 있다.

[266] 이에 따라 사용자가 음성 입력을 종료하였지만, 노이즈로 인한 음성 인식이 이루어지는 것이 방지될 수 있고, 음성 입력의 종료에 맞게 정확한 음성 인식 서비스가 제공될 수 있다.

[267] 본 발명의 일 실시 예에 의하면, 전술한 방법은, 프로그램이 기록된 매체에 프로세서가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 프로세서가 읽을 수 있는 매체의 예로는, ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광 데이터 저장장치 등이 있다.

[268] 상기와 같이 설명된 디스플레이 장치는 상기 설명된 실시 예들의 구성과 방법이 한정되게 적용될 수 있는 것이 아니라, 상기 실시 예들은 다양한 변형이 이루어질 수 있도록 각 실시 예들의 전부 또는 일부가 선택적으로 조합되어 구성될 수도 있다.

## 청구범위

- [청구항 1] 디스플레이 장치에 있어서,  
제1 서버 및 제2 서버와 통신하는 네트워크 인터페이스; 및  
사용자의 음성 입력을 획득하고, 획득된 음성 입력에 상응하는 음성  
신호를 상기 제1 서버에 전송하고, 상기 제1 서버로부터 상기 음성 신호의  
에너지 레벨, 상기 음성 입력에 상응하는 텍스트 및 상기 음성 입력에  
대한 발화 끝점 정보를 수신하고, 상기 에너지 레벨 및 상기 발화 끝점  
정보에 기초하여 상기 사용자의 발화가 종료되었는지 여부를 판단하는  
컨트롤러를 포함하는  
디스플레이 장치.
- [청구항 2] 제1항에 있어서,  
상기 컨트롤러는  
상기 에너지 레벨이 기 설정된 제1 레벨 미만이고, 상기 발화 끝점 정보가  
상기 발화의 끝점임을 나타내는 값을 포함하는 경우, 상기 발화가 종료된  
것으로 판단하는  
디스플레이 장치.
- [청구항 3] 제1항에 있어서,  
상기 컨트롤러는  
상기 에너지 레벨이 기 설정된 제1 레벨 미만이고, 상기 발화 끝점 정보에  
포함된 발화 끝점을 나타내는 신뢰도 스코어가 기 설정된 스코어 이상인  
경우, 상기 발화가 종료된 것으로 판단하는  
디스플레이 장치.
- [청구항 4] 제3항에 있어서,  
상기 컨트롤러는  
상기 에너지 레벨이 상기 제1 레벨보다 큰 기 설정된 제2 레벨 이상이고,  
상기 발화 끝점 정보에 포함된 발화 끝점을 나타내는 신뢰도 스코어가 기  
설정된 스코어 이상인 경우, 상기 발화가 종료되지 않은 것으로 판단하는  
디스플레이 장치.
- [청구항 5] 제1항에 있어서,  
상기 컨트롤러는  
상기 사용자의 발화가 종료된 것으로 판단한 경우, 상기 제2 서버에 상기  
텍스트를 전송하고, 상기 제2 서버로부터 상기 텍스트의 의도 분석  
결과를 나타내는 분석 결과 정보를 수신하고, 수신된 분석 결과 정보를  
출력하는  
디스플레이 장치.
- [청구항 6] 제1항에 있어서,  
상기 컨트롤러는

상기 사용자의 발화가 종료된 것으로 판단한 경우, 상기 음성 입력에 대한 음성 인식 결과를 출력하기까지 추가로 입력되는 음성 입력을 무시하는 디스플레이 장치.

- [청구항 7] 제1항에 있어서,  
상기 컨트롤러는  
상기 음성 신호를 펄스 코드 변조(Pulse Code Modulation, PCM) 신호로 변환하고, 변환된 PCM 신호를 상기 네트워크 인터페이스를 통해 상기 제1 서버에 전송하는  
디스플레이 장치.
- [청구항 8] 제1항에 있어서,  
상기 제1 서버는 음성을 텍스트로 변환하는 STT(Speech To Text) 서버이고, 제2 서버는 NLP(Natural Language Processing) 서버인  
디스플레이 장치.
- [청구항 9] 디스플레이 장치에 있어서,  
제1 서버 및 제2 서버와 통신하는 네트워크 인터페이스; 및  
사용자의 음성 입력을 획득하고, 획득된 음성 입력에 상응하는 음성 신호를 상기 제1 서버에 전송하고, 상기 제1 서버로부터 상기 음성 신호의 에너지 레벨 및 상기 음성 입력에 상응하는 텍스트를 수신하고, 상기 텍스트를 상기 제2 서버에 전송하고, 상기 제2 서버로부터 상기 음성 입력에 대한 발화 끝점 정보를 수신하고, 상기 에너지 레벨 및 상기 발화 끝점 정보에 기초하여 상기 사용자의 발화가 종료되었는지 여부를 판단하는 컨트롤러를 포함하는  
디스플레이 장치.
- [청구항 10] 제9항에 있어서,  
상기 컨트롤러는  
상기 에너지 레벨이 기 설정된 제1 레벨 미만이고, 상기 발화 끝점 정보가 상기 발화의 끝점임을 나타내는 값을 포함하는 경우, 상기 발화가 종료된 것으로 판단하는  
디스플레이 장치.
- [청구항 11] 제9항에 있어서,  
상기 컨트롤러는  
상기 에너지 레벨이 기 설정된 제1 레벨 미만이고, 상기 발화 끝점 정보에 포함된 발화 끝점을 나타내는 신뢰도 스코어가 기 설정된 스코어 이상인 경우, 상기 발화가 종료된 것으로 판단하는  
디스플레이 장치.
- [청구항 12] 제11항에 있어서,  
상기 컨트롤러는  
상기 에너지 레벨이 상기 제1 레벨보다 큰 기 설정된 제2 레벨 이상이고,

상기 발화 끝점 정보에 포함된 발화 끝점을 나타내는 신뢰도 스코어가 기 설정된 스코어 이상인 경우, 상기 발화가 종료되지 않은 것으로 판단하는 디스플레이 장치.

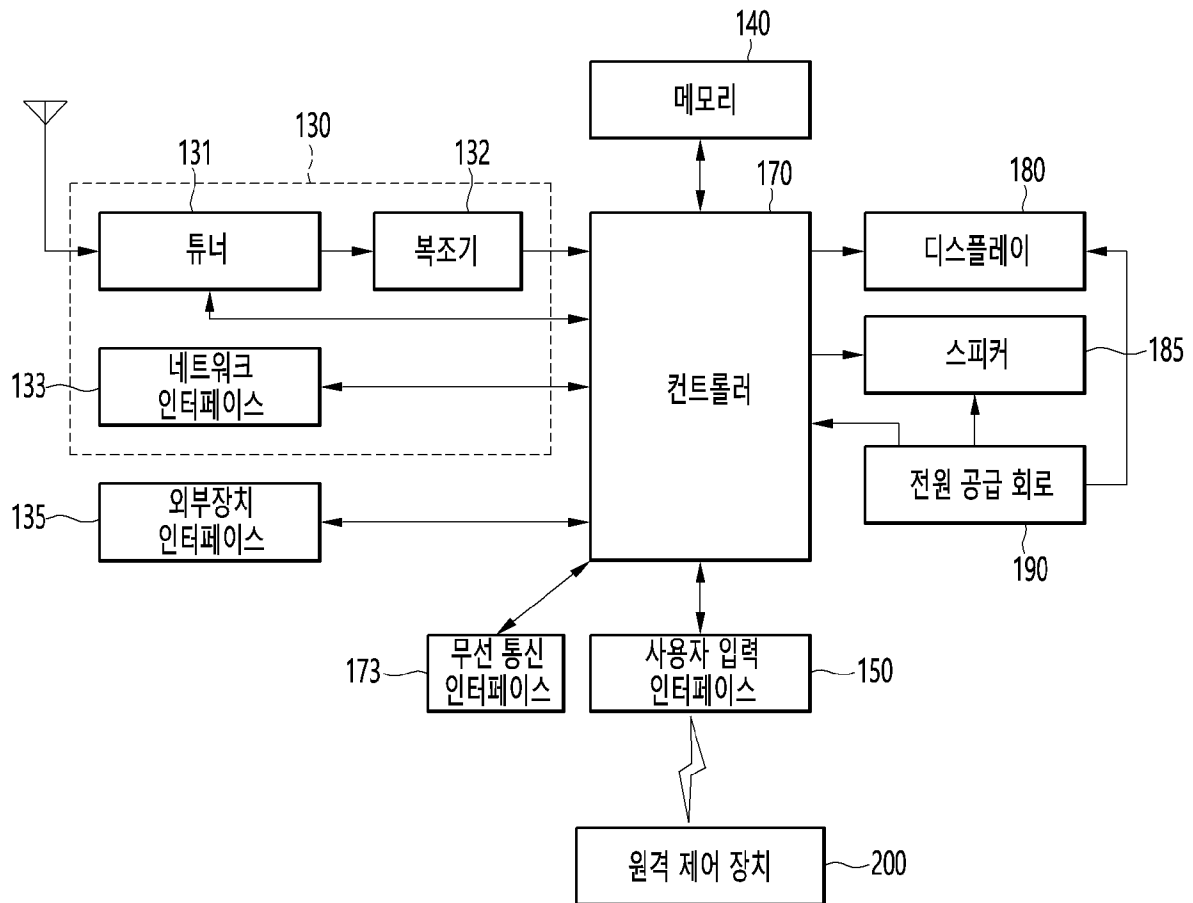
[청구항 13] 제9항에 있어서,  
상기 컨트롤러는  
상기 네트워크 인터페이스를 통해 상기 제2 서버로부터 상기 음성 입력에 대한 의도 분석을 나타내는 분석 결과 정보를 더 수신하는  
디스플레이 장치.

[청구항 14] 제9항에 있어서,  
상기 컨트롤러는  
상기 컨트롤러는  
상기 사용자의 발화가 종료된 것으로 판단한 경우, 상기 음성 입력에 대한 음성 인식 결과를 출력하기까지 추가로 입력되는 음성 입력을 무시하는  
디스플레이 장치.

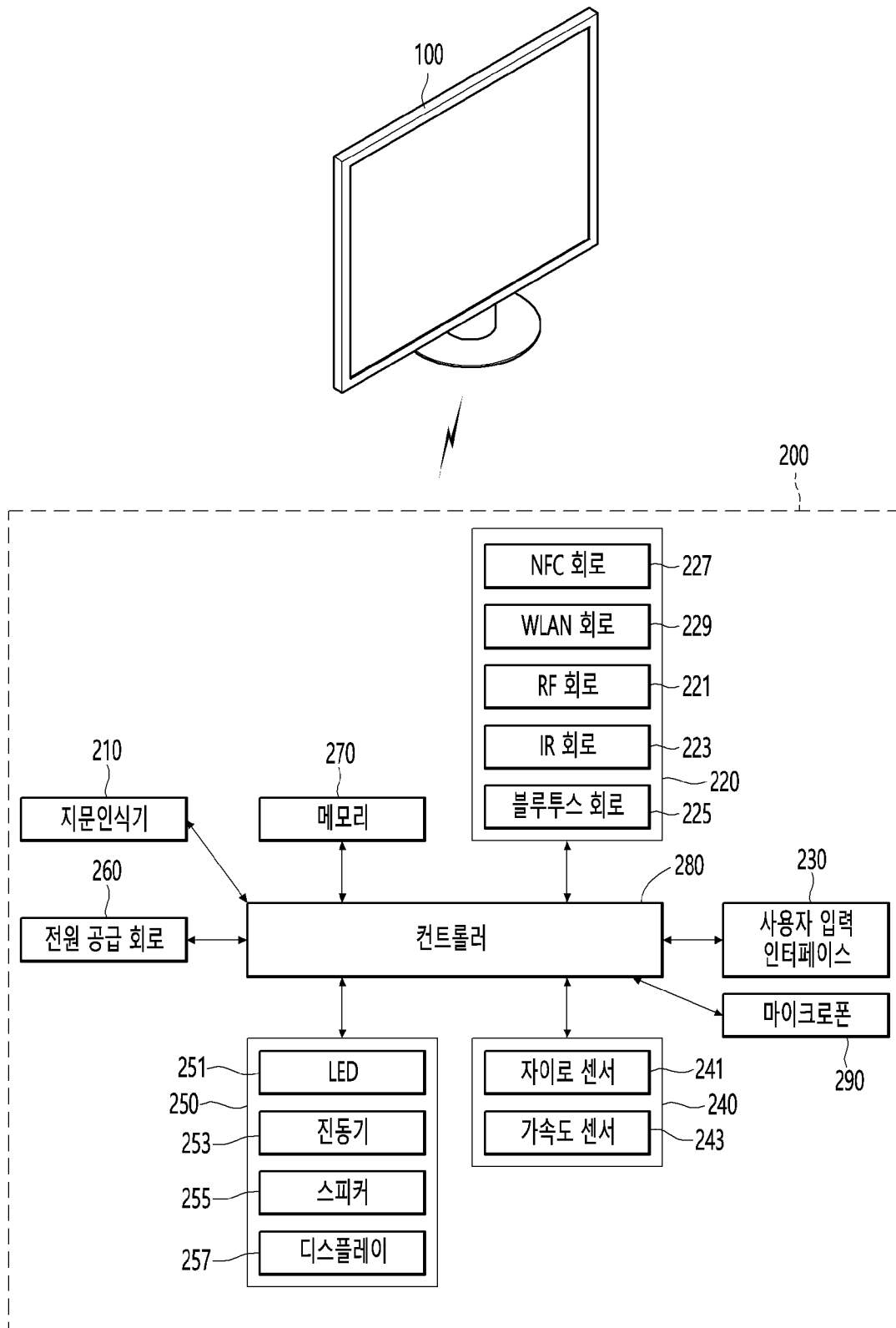
[청구항 15] 제9항에 있어서,  
상기 제1 서버는 음성을 텍스트로 변환하는 STT(Speech To Text) 서버이고, 제2 서버는 NLP(Natural Language Processing) 서버인  
디스플레이 장치.

[도 1]

100

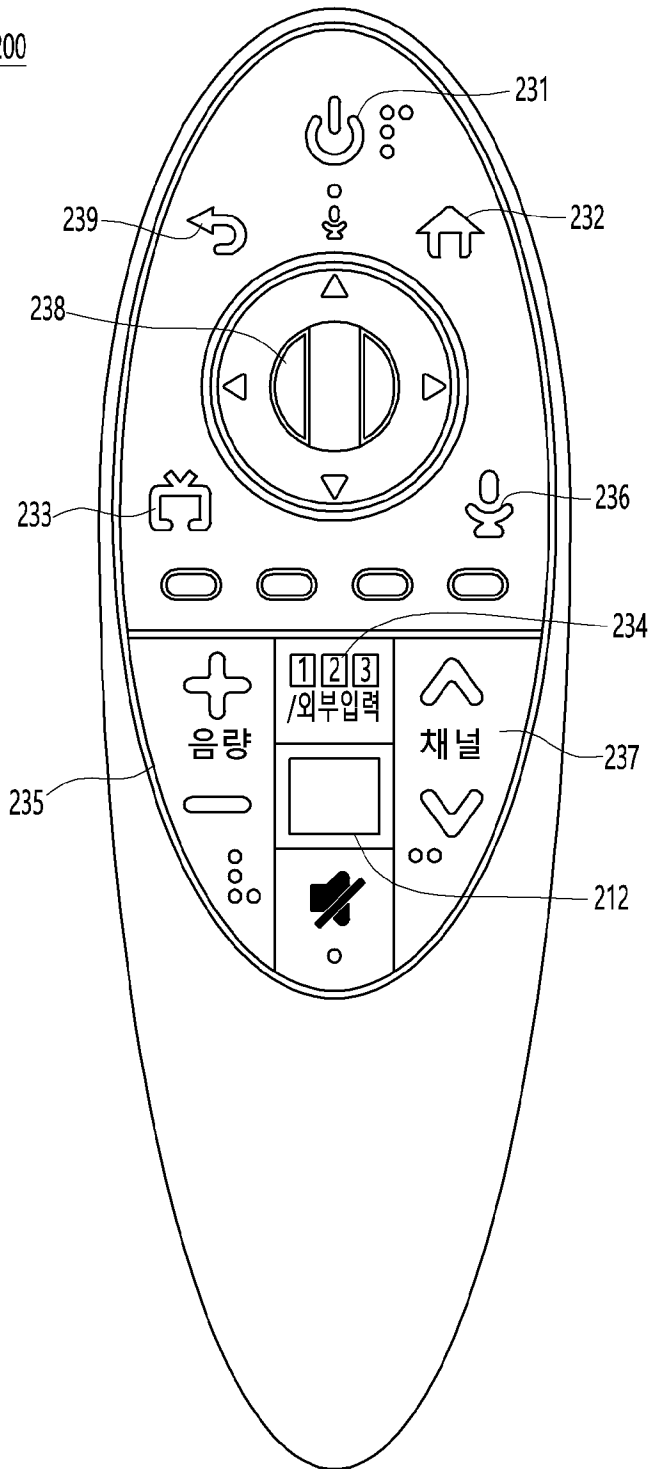


[도2]

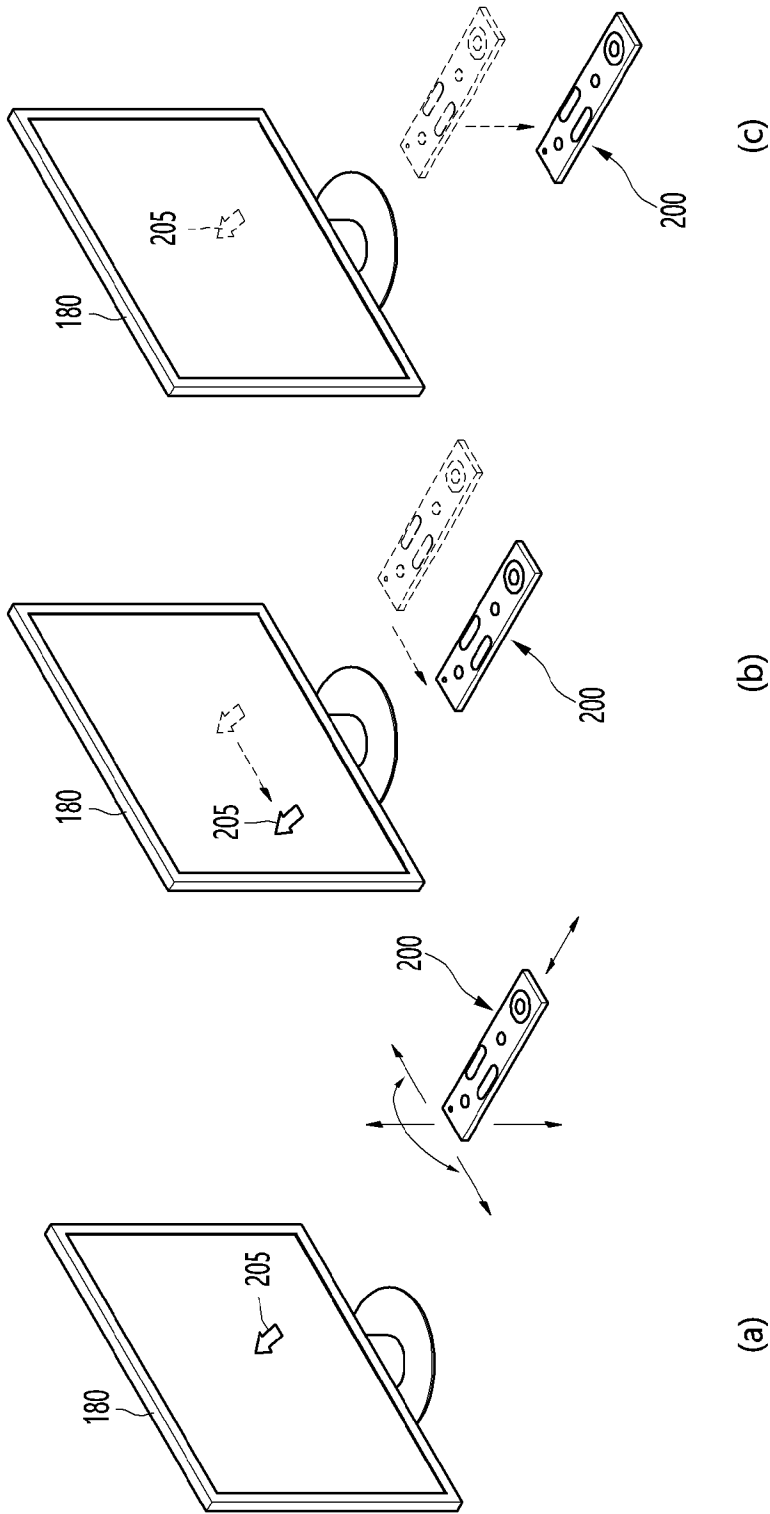


[도3]

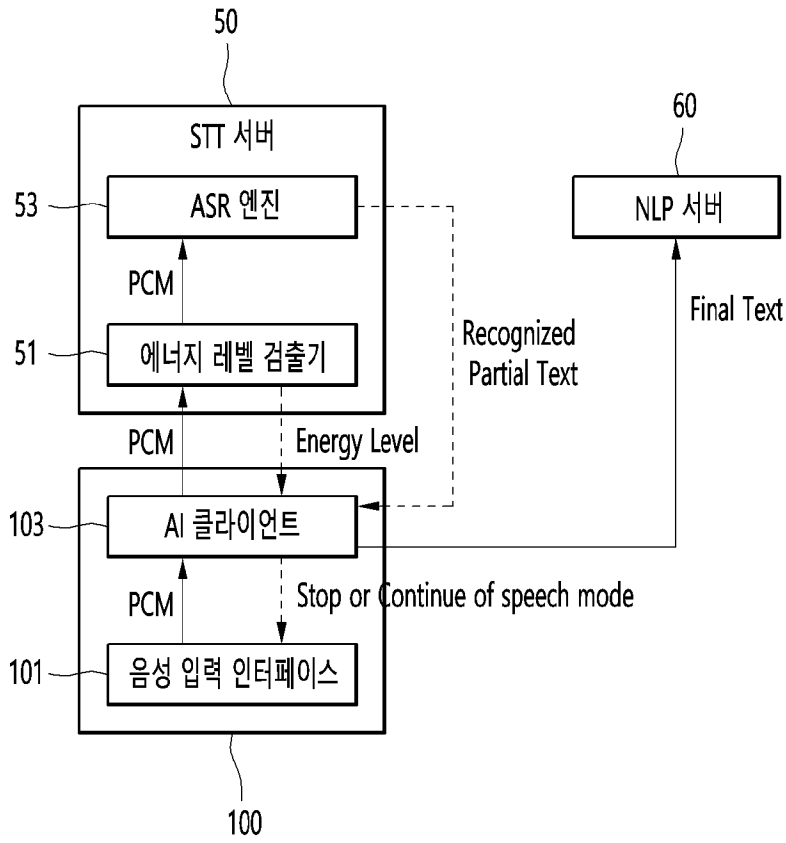
200



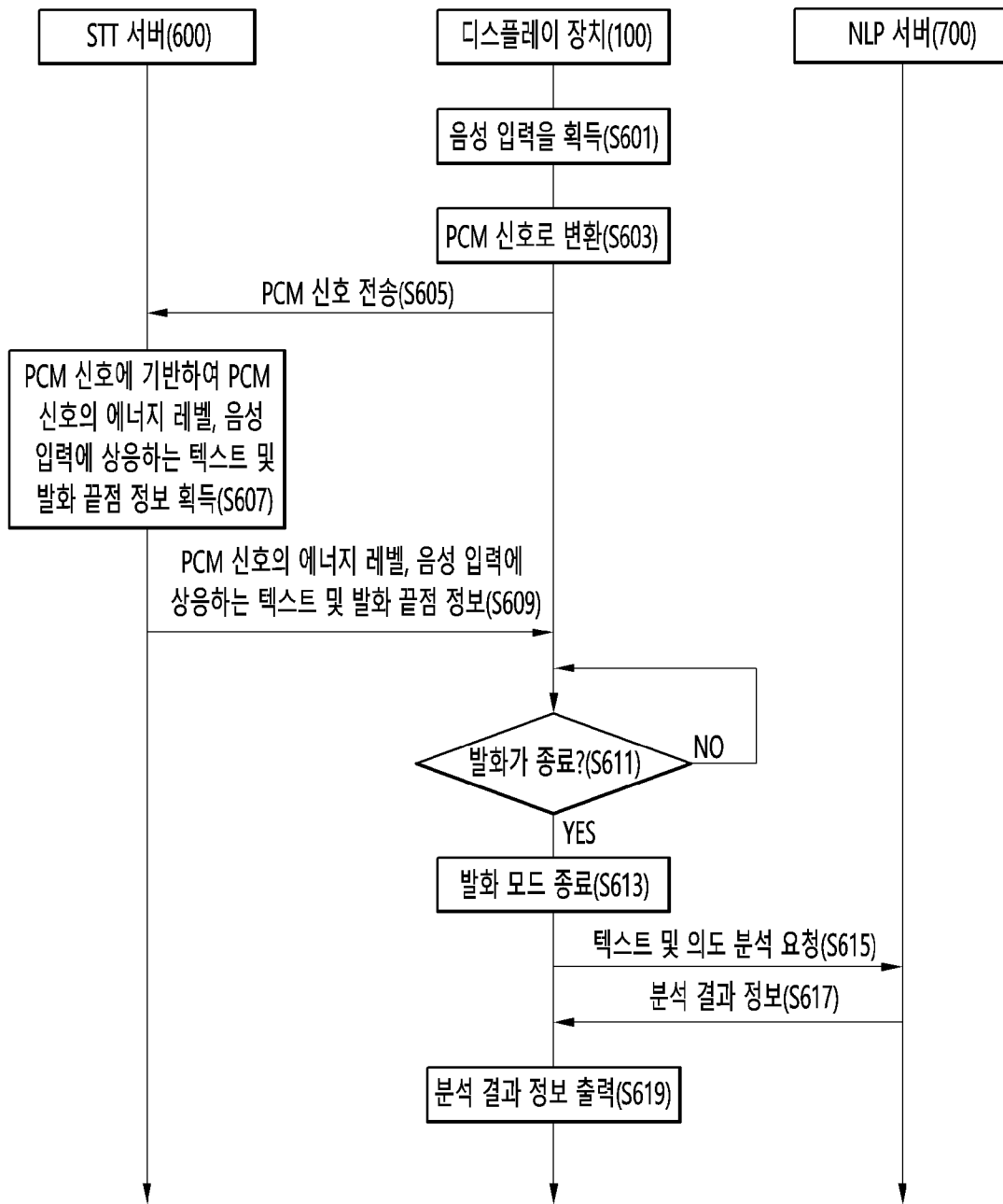
[도4]



[도5]

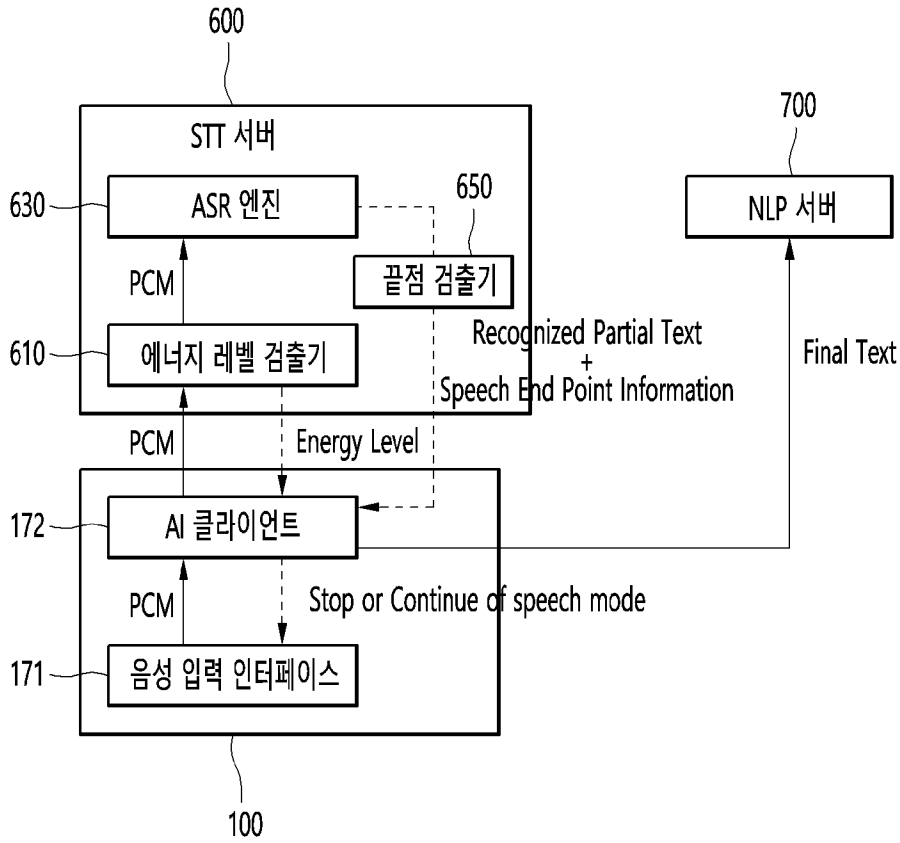


[도6]

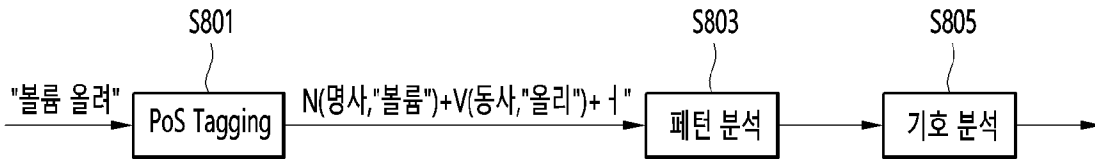


[도7]

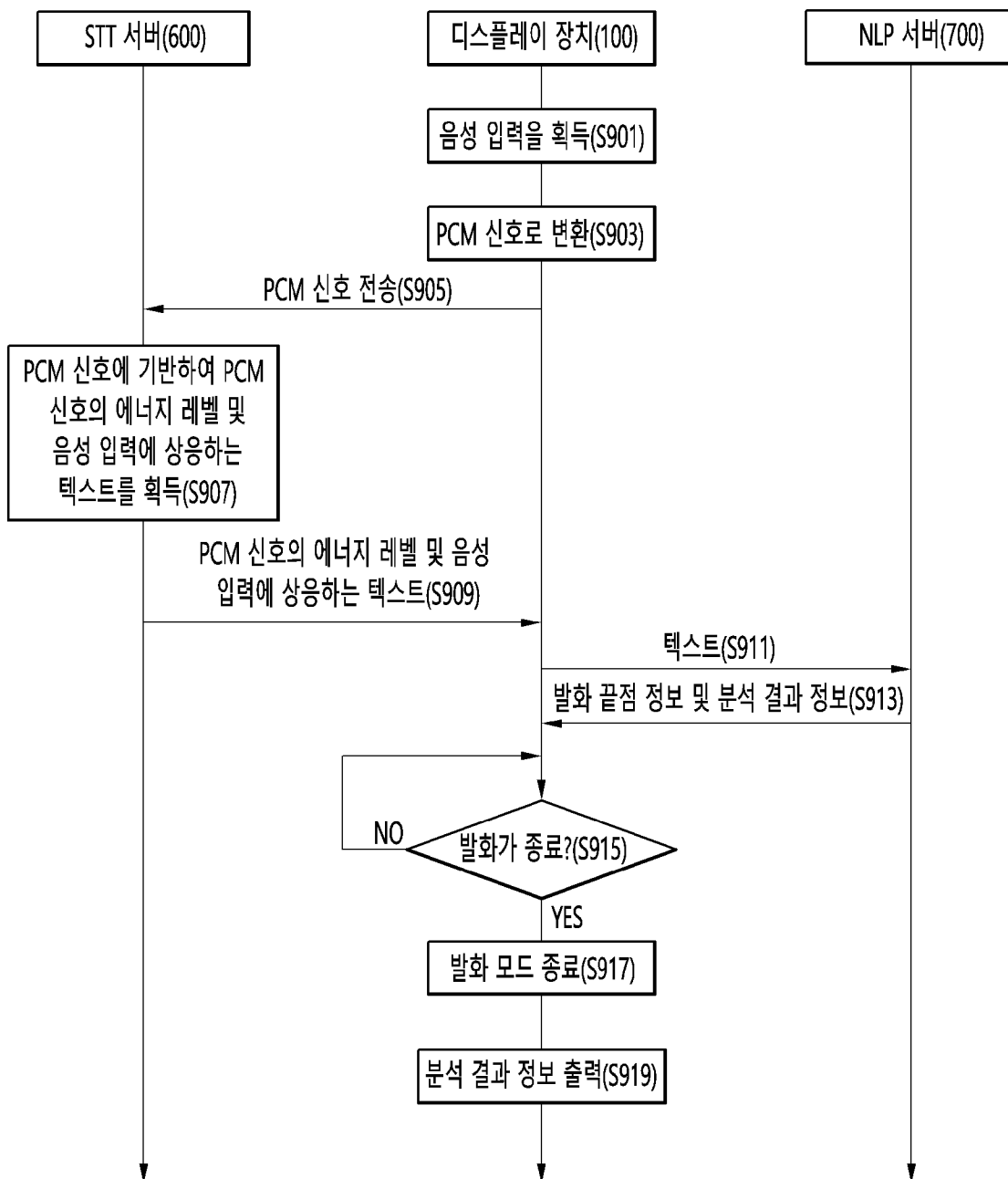
70



[도8]

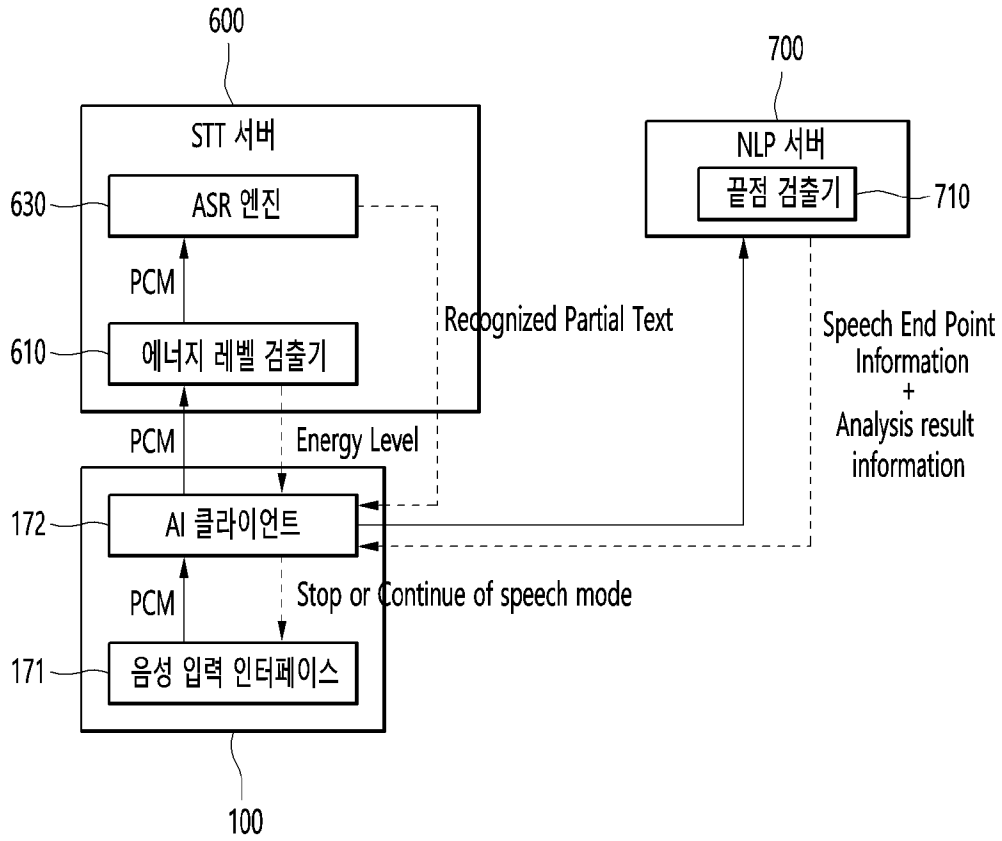


[도9]



[도 10]

1000



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2022/009357

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
G10L 15/22(2006.01)i; G10L 15/30(2013.01)i; G10L 15/26(2006.01)i; G10L 25/87(2013.01)i; G10L 15/18(2006.01)i; G10L 19/10(2006.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) G10L 15/22(2006.01); G06F 3/16(2006.01); G10L 15/00(2006.01); G10L 25/51(2013.01); G10L 25/78(2013.01); G10L 25/87(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: 음성(voice), 텍스트(text), 에너지 레벨(energy level), 발화(utterance), 끝점(end point)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2021-0031265 A (SAMSUNG ELECTRONICS CO., LTD.) 19 March 2021 (2021-03-19) See paragraphs [0035]-[0166]; claims 1-9; and figures 2-9.	1-15
Y	KR 10-2019-0049260 A (HYUNDAI MOTOR COMPANY et al.) 09 May 2019 (2019-05-09) See paragraphs [0041]-[0078]; claims 1-10; and figures 2-7.	1-15
Y	KR 10-2018-0084394 A (SAMSUNG ELECTRONICS CO., LTD.) 25 July 2018 (2018-07-25) See paragraphs [0059] and [0093]-[0231]; claims 1-13; and figures 5a-13.	3-6,8,11-15
Y	KR 10-2013-0122359 A (SAMSUNG ELECTRONICS CO., LTD.) 07 November 2013 (2013-11-07) See paragraphs [0046]-[0127]; claims 1-8; and figures 2-8.	7
A	US 2012-0265530 A1 (HETHERINGTON, Phil et al.) 18 October 2012 (2012-10-18) See paragraphs [0027]-[0053]; claims 1-11; and figures 3-10.	1-15
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 <b>20 March 2023</b>		Date of mailing of the international search report <b>20 March 2023</b>
Name and mailing address of the ISA/KR <b>Korean Intellectual Property Office Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208</b> Facsimile No. +82-42-481-8578		Authorized officer  Telephone No.

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/KR2022/009357**

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
KR 10-2021-0031265	A	19 March 2021	CN	112489691	A	12 March 2021	
			EP	3792919	A1	17 March 2021	
			US	2021-0074290	A1	11 March 2021	
			WO	2021-049795	A1	18 March 2021	
KR 10-2019-0049260	A	09 May 2019	US	10621985	B2	14 April 2020	
			US	2019-0130907	A1	02 May 2019	
KR 10-2018-0084394	A	25 July 2018	CN	110199350	A	03 September 2019	
			EP	3570275	A1	20 November 2019	
			EP	3570275	B1	20 April 2022	
			US	11211048	B2	28 December 2021	
			US	2019-0378493	A1	12 December 2019	
KR 10-2013-0122359	A	07 November 2013	WO	2018-135743	A1	26 July 2018	
			US	2013-00290001	A1	31 October 2013	
US 2012-0265530	A1	18 October 2012	US	2017-00223301	A1	03 August 2017	
			CN	101031958	A	05 September 2007	
US 2012-0265530	A1	18 October 2012	CN	101031958	B	16 May 2012	
			EP	1771840	A1	11 April 2007	
			JP	2008-508564	A	21 March 2008	
			JP	2011-107715	A	02 June 2011	
			JP	5331784	B2	30 October 2013	
			KR	10-2007-0088469	A	29 August 2007	
			US	2006-0287859	A1	21 December 2006	
			US	2007-0288238	A1	13 December 2007	
			US	2008-0228478	A1	18 September 2008	
			US	2012-0303366	A1	29 November 2012	
			US	8165880	B2	24 April 2012	
			US	8170875	B2	01 May 2012	
			US	8311819	B2	13 November 2012	
			US	8457961	B2	04 June 2013	
			US	8554564	B2	08 October 2013	
WO	2006-133537	A1	21 December 2006				

<b>A. 발명이 속하는 기술분류(국제특허분류(IPC))</b> <b>G10L 15/22(2006.01)i; G10L 15/30(2013.01)i; G10L 15/26(2006.01)i; G10L 25/87(2013.01)i; G10L 15/18(2006.01)i; G10L 19/10(2006.01)i</b>		
<b>B. 조사된 분야</b> 조사된 최소문헌(국제특허분류를 기재) G10L 15/22(2006.01); G06F 3/16(2006.01); G10L 15/00(2006.01); G10L 25/51(2013.01); G10L 25/78(2013.01); G10L 25/87(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: 음성(voice), 텍스트(text), 에너지 레벨(energy level), 발화 (utterance), 끝점(end point)		
<b>C. 관련 문헌</b>		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2021-0031265 A (삼성전자주식회사) 2021.03.19 단락 [0035]-[0166]; 청구항 1-9; 및 도면 2-9	1-15
Y	KR 10-2019-0049260 A (현대자동차주식회사 등) 2019.05.09 단락 [0041]-[0078]; 청구항 1-10; 및 도면 2-7	1-15
Y	KR 10-2018-0084394 A (삼성전자주식회사) 2018.07.25 단락 [0059], [0093]-[0231]; 청구항 1-13; 및 도면 5a-13	3-6,8,11-15
Y	KR 10-2013-0122359 A (삼성전자주식회사) 2013.11.07 단락 [0046]-[0127]; 청구항 1-8; 및 도면 2-8	7
A	US 2012-0265530 A1 (PHIL HETHERINGTON 등) 2012.10.18 단락 [0027]-[0053]; 청구항 1-11; 및 도면 3-10	1-15
<input type="checkbox"/> 추가 문헌이 C(계속)에 기재되어 있습니다. <input checked="" type="checkbox"/> 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		
국제조사의 실제 완료일 <b>2023년03월20일 (20.03.2023)</b>		국제조사보고서 발송일 <b>2023년03월20일 (20.03.2023)</b>
ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578		심사관  양정록 전화번호 +82-42-481-5709

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2021-0031265 A	2021/03/19	CN 112489691 A	2021/03/12
		EP 3792919 A1	2021/03/17
		US 2021-0074290 A1	2021/03/11
		WO 2021-049795 A1	2021/03/18
KR 10-2019-0049260 A	2019/05/09	US 10621985 B2	2020/04/14
		US 2019-0130907 A1	2019/05/02
KR 10-2018-0084394 A	2018/07/25	CN 110199350 A	2019/09/03
		EP 3570275 A1	2019/11/20
		EP 3570275 B1	2022/04/20
		US 11211048 B2	2021/12/28
		US 2019-0378493 A1	2019/12/12
KR 10-2013-0122359 A	2013/11/07	US 2013-00290001 A1	2013/10/31
		US 2017-00223301 A1	2017/08/03
US 2012-0265530 A1	2012/10/18	CN 101031958 A	2007/09/05
		CN 101031958 B	2012/05/16
		EP 1771840 A1	2007/04/11
		JP 2008-508564 A	2008/03/21
		JP 2011-107715 A	2011/06/02
		JP 5331784 B2	2013/10/30
		KR 10-2007-0088469 A	2007/08/29
		US 2006-0287859 A1	2006/12/21
		US 2007-0288238 A1	2007/12/13
		US 2008-0228478 A1	2008/09/18
		US 2012-0303366 A1	2012/11/29
		US 8165880 B2	2012/04/24
		US 8170875 B2	2012/05/01
		US 8311819 B2	2012/11/13
US 8457961 B2	2013/06/04		
US 8554564 B2	2013/10/08		
WO 2006-133537 A1	2006/12/21		