(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2024 년 4 월 11 일 (11.04.2024)

WIPO | PCT

(10) 국제공개번호
WO 2024/076114 A1

(51) 국제특허분류:
G06F 3/16 (2006.01) *G10L 15/22* (2006.01)
G06F 3/0484 (2013.01) *G06F 3/048* (2006.01)
G06F 3/0481 (2013.01) *G10L 15/26* (2006.01)
G10L 15/02 (2006.01) *G10L 25/51* (2013.01)

(21) 국제출원번호: PCT/KR2023/015158

(22) 국제출원일: 2023 년 9 월 27 일 (27.09.2023)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(30) 우선권정보:
10-2022-0129140 2022 년 10 월 8 일 (08.10.2022) KR
10-2022-0144807 2022 년 11 월 2 일 (02.11.2022) KR

(71) 출원인: 삼성전자주식회사 (**SAMSUNG ELECTRONICS CO., LTD.**) [KR/KR]; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).

(72) 발명자: **김성한 (KIM, Sunghan)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). **김솔 (KIM, Sol)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). **김영빈 (KIM, Youngbin)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). **손지수 (SON, Jisu)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).

(74) 대리인: **특허법인 광앤장 (KWANG AND JANG INTELLECTUAL PROPERTY LAW FIRM)**; 06775 서울특별시 서초구 논현로17길 16, 4층, Seoul (KR).

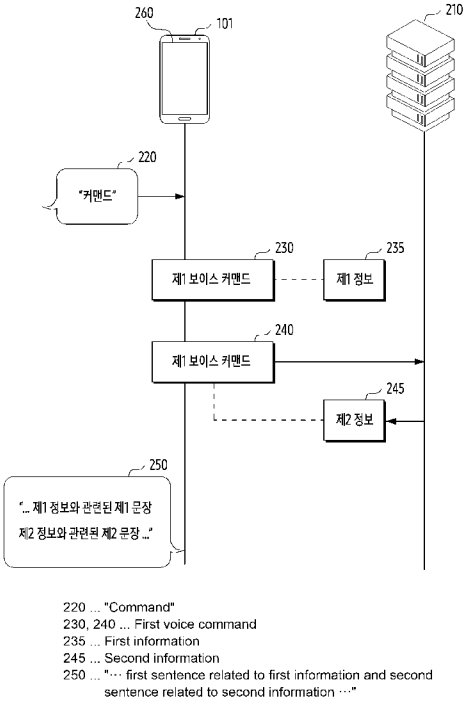
(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KW, KZ,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

WO 2024/076114 A1

(54) **Title:** ELECTRONIC DEVICE FOR CONTROLLING EXECUTION OF VOICE COMMAND, AND METHOD THEREOF

(54) **발명의 명칭:** 보이스 커맨드의 실행을 제어하기 위한 전자 장치 및 그 방법



220 ... "Command"
230, 240 ... First voice command
235 ... First information
245 ... Second information
250 ... "··· first sentence related to first information and second sentence related to second information ···"

(57) Abstract: The electronic device may comprise a communication circuit, a microphone, a speaker, a memory, and a processor. The processor may identify a first audio signal received through the microphone. The processor may perform at least one first function among multiple functions on the basis of the identification of the first audio signal including an utterance for sequentially executing the multiple functions. The processor may perform at least one second function among the multiple functions by using an external electronic device connected through the communication circuit on the basis of the identification of the first audio signal including the utterance. The at least one second function may be different from the at least one first function. The processor may sequentially output, through the speaker of the electronic device according to utterance-related orders, one or more second audio signals representing one or more results obtained by performing the at least one first function and the at least one second function.

(57) 요약서: 전자 장치(electronic device)는, 통신 회로, 마이크, 스피커, 메모리 및 프로세서를 포함할 수 있다. 프로세서는, 마이크를 통하여 수신된 제1 오디오 신호를 식별할 수 있다. 프로세서는, 복수의 기능들을 순차적으로 실행하기 위한 발언을 포함하는 제1 오디오 신호를 식별하는 것에 기반하여, 복수의 기능들 중 적어도 하나의 제1 기능을 수행할 수 있다. 프로세서는, 제1 오디오 신호가 발언을 포함함을 식별하는 것에 기반하여, 통신 회로를 통해 연결된 외부 전자 장치를 이용하여, 복수의 기능들 중에서, 적어도 하나의 제2 기능을 수행할 수 있다. 적어도 하나의 제2 기능은 적어도 하나의 제1 기능과 상이할 수 있다. 프로세서는, 전자 장치의 스피커를 통하여, 적어도 하나의 제1 기능, 및 적어도 하나의 제2 기능을 수행한 적어도 하나의 결과를 표현하는 적어도 하나의 제2 오디오 신호를, 발언과 관련된 순서를 따라 순차적으로 출력할 수 있다.

[다음 쪽 계속]

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, ME, 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))

명세서

발명의 명칭: 보이스 커맨드의 실행을 제어하기 위한 전자 장치 및 그 방법

기술분야

[1] 본 개시(disclosure)는 보이스 커맨드의 실행을 제어하기 위한 전자 장치 및 그 방법에 관한 것이다.

배경기술

[2] 전자 장치, 및 사용자 사이의 인터페이스는 키보드, 및/또는 마우스를 포함할 수 있다. 전자 장치를 보다 직관적으로 제어하기 위하여, 전자 장치, 및 사용자 사이의 인터페이스들의 타입들이 확대될 수 있다. 예를 들어, 전자 장치는, 전자 장치를 제어하기 위한 사용자의 발언(speech)을 식별하기 위하여, 마이크를 이용할 수 있다.

발명의 상세한 설명

과제 해결 수단

[3] 일 실시예에 따른, 전자 장치는, 통신 회로, 마이크, 스피커, 적어도 하나의 인스트럭션를 저장하기 위한 메모리 및 적어도 하나의 프로세서를 포함할 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 마이크를 통하여 수신된 제1 오디오 신호를 식별하도록, 구성될 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 복수의 기능들을 순차적으로 실행하기 위한 발언을 포함하는 제1 오디오 신호를 식별하는 것에 기반하여, 상기 복수의 기능들 중 적어도 하나의 제1 기능을 수행하도록, 구성될 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 제1 오디오 신호가 발언을 포함함을 식별하는 것에 기반하여, 상기 통신 회로를 통해 외부 전자 장치의 상기 복수의 기능들 중에서 적어도 하나의 제2 기능을 수행하도록, 구성될 수 있다. 상기 적어도 하나의 제2 기능은, 상기 복수의 기능들 중에서, 상기 적어도 하나의 제1 기능과 상이할 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 스피커를 통하여, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능을 수행한 적어도 하나의 결과를 표현하는(representing) 제2 오디오 신호들을, 상기 발언과 관련된 순서에 기반하여 순차적으로 출력하도록, 구성될 수 있다.

[4] 일 실시예에 따른, 전자 장치(electronic device)의 방법은, 상기 전자 장치의 마이크를 통하여 제1 오디오 신호를 수신하는 것에 기반하여, 상기 제1 오디오 신호 내 발언을 식별하는 동작을 포함할 수 있다. 상기 방법은, 상기 발언으로부터 복수의 보이스 커맨드들을 식별하는 것에 기반하여, 상기 전자 장치에 의해 실행되는 제1 보이스 엔진의 상태, 및 외부 전자 장치에 의해 실행되는 제2 보이스 엔

진의 상태를 획득하는 동작을 포함할 수 있다. 상기 외부 전자 장치는 상기 전자 장치의 통신 회로를 통해 연결될 수 있다. 상기 방법은, 상기 제1 보이스 엔진의 상태 및 상기 제2 보이스 엔진의 상태에 기반하여, 상기 복수의 보이스 커맨드들을, 상기 제1 보이스 엔진에 의해 수행되는 적어도 하나의 제1 보이스 커맨드, 및 상기 제2 보이스 엔진에 의해 수행되는 제2 보이스 커맨드로 구분하기 위한 정보를 획득하는 동작을 포함할 수 있다. 상기 방법은, 상기 적어도 하나의 제1 보이스 커맨드에 대응하는 제1 기능을, 상기 제1 보이스 엔진을 이용하여 실행하고, 상기 적어도 하나의 제2 보이스 커맨드에 대응하는 제2 기능을 수행하기 위하여 외부 전자 장치로 요청하는 동작을 포함할 수 있다. 상기 방법은, 상기 전자 장치의 스피커를 통하여, 상기 발언에 의해 지시되는 상기 복수의 보이스 커맨드들의 순서에 기반하여, 상기 제1 기능, 및 상기 제2 기능을 실행한 적어도 하나의 결과를 표현한(representing) 적어도 하나의 제2 오디오 신호를 출력하는 동작을 포함할 수 있다.

[5] 일 실시예에 따른, 전자 장치의 방법은, 상기 전자 장치의 마이크를 통해 수신된 제1 오디오 신호를 식별하는 동작을 포함할 수 있다. 상기 방법은, 상기 제1 오디오 신호가 복수의 기능들을 순차적으로 실행하기 위한 발언을 포함함을 식별하는 것에 기반하여, 상기 복수의 기능들로부터 적어도 하나의 제1 기능을 수행하는 동작을 포함할 수 있다. 상기 방법은, 상기 제1 오디오 신호가 상기 발언을 포함함을 식별하는 것에 기반하여, 상기 외부 전자 장치의 통신 회로를 통해 연결된 외부 전자 장치에서, 상기 복수의 기능들 중 적어도 하나의 제2 기능을 수행하는 동작을 포함할 수 있다. 상기 적어도 하나의 제2 기능은 상기 적어도 하나의 제1 기능과 다를 수 있다. 상기 방법은, 상기 전자 장치의 스피커를 통하여, 상기 발언과 관련된 발언에 기반하여 상기 적어도 하나의 제1 기능 및 상기 적어도 하나의 제2 기능을 수행한 적어도 하나의 결과를 표현하는 적어도 하나의 제2 오디오 신호를 순차적으로 출력하는 동작을 포함할 수 있다.

[6] 일 실시예(an embodiment)에 따른, 전자 장치(an electronic device)는, 통신 회로, 마이크, 스피커, 및 프로세서를 포함할 수 있다. 상기 프로세서는, 상기 마이크를 통하여 제1 오디오 신호를 식별하도록, 구성될 수 있다. 상기 프로세서는, 상기 제1 오디오 신호로부터 복수의 기능들을 순차적으로 실행하기 위한 발언을 식별하는 것에 기반하여, 상기 복수의 기능들 중 적어도 하나의 제1 기능을 실행하도록, 구성될 수 있다. 상기 프로세서는, 상기 발언을 식별하는 것에 기반하여, 상기 통신 회로를 통해 연결된 외부 전자 장치를 이용하여, 상기 복수의 기능들 중에서, 상기 적어도 하나의 제1 기능과 상이한 적어도 하나의 제2 기능을 실행하도록, 구성될 수 있다. 상기 프로세서는, 상기 스피커를 통하여, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능을 실행한 결과들을 표현하는(representing) 제2 오디오 신호들을, 상기 발언과 관련된 순서를 따라 순차적으로 출력하도록, 구성될 수 있다.

[7] 일 실시예에 따른, 전자 장치의 방법은, 상기 전자 장치의 마이크를 통하여 제1 오디오 신호를 식별하는 동작을 포함할 수 있다. 상기 방법은, 상기 제1 오디오 신호로부터 복수의 기능들을 순차적으로 실행하기 위한 발언을 식별하는 것에 기반하여, 상기 전자 장치의 프로세서에 기반하여 상기 복수의 기능들 중 적어도 하나의 제1 기능을 실행하는 동작을 포함할 수 있다. 상기 방법은, 상기 발언을 식별하는 것에 기반하여, 상기 전자 장치의 통신 회로를 통해 연결된 외부 전자 장치를 이용하여, 상기 복수의 기능들 중에서, 상기 적어도 하나의 제1 기능과 상이한 적어도 하나의 제2 기능을 실행하는 동작을 포함할 수 있다. 상기 방법은, 상기 전자 장치의 스피커를 통하여, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능을 실행한 결과들을 표현하는 제2 오디오 신호들을, 상기 발언과 관련된 순서를 따라 순차적으로 출력하는 동작을 포함할 수 있다.

[8] 일 실시예에 따른, 전자 장치(electronic device)는, 통신 회로, 마이크, 스피커, 및 프로세서를 포함할 수 있다. 상기 프로세서는, 상기 마이크를 통하여 제1 오디오 신호를 수신하는 것에 기반하여, 상기 제1 오디오 신호에 포함된 발언을 식별하도록, 구성될 수 있다. 상기 프로세서는, 상기 발언으로부터 복수의 보이스 커맨드들(voice commands)을 식별하는 것에 기반하여, 보이스 커맨드(a voice command)를 처리하기 위하여 상기 프로세서에 의해 실행되는 제1 보이스 엔진, 및 상기 통신 회로를 통해 연결된 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 상태들을 획득하도록, 구성될 수 있다. 상기 프로세서는, 상기 상태들에 기반하여, 상기 복수의 보이스 커맨드들을, 상기 제1 보이스 엔진에 대응하는 제1 보이스 커맨드, 및 상기 제2 보이스 엔진에 대응하는 제2 보이스 커맨드로 구분하기 위한 정보를 획득하도록, 구성될 수 있다. 상기 프로세서는, 상기 정보에 기반하여, 상기 제1 보이스 커맨드에 대응하는 제1 기능을, 상기 제1 보이스 엔진을 이용하여 실행하고, 상기 외부 전자 장치로 상기 제2 보이스 커맨드에 대응하는 제2 기능의 실행을 요청하도록, 구성될 수 있다. 상기 프로세서는, 상기 스피커를 통하여, 상기 발언에 의해 지시되는(indicated by) 상기 복수의 보이스 커맨드들의 순서에 기반하여, 상기 제1 기능, 및 상기 제2 기능을 실행한 결과들을 표현한(representing) 제2 오디오 신호를 출력하도록, 구성될 수 있다.

[9] 일 실시예에 따른, 전자 장치(electronic device)의 방법은, 상기 전자 장치의 마이크를 통하여 제1 오디오 신호를 수신하는 것에 기반하여, 상기 제1 오디오 신호에 포함된 발언을 식별하는 동작을 포함할 수 있다. 상기 방법은, 상기 발언으로부터 복수의 보이스 커맨드들을 식별하는 것에 기반하여, 보이스 커맨드를 처리하기 위하여 상기 전자 장치의 프로세서에 의해 실행되는 제1 보이스 엔진, 및 상기 전자 장치의 통신 회로를 통해 연결된 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 상태들을 획득하는 동작을 포함할 수 있다. 상기 방법은, 상기 상태들에 기반하여, 상기 복수의 보이스 커맨드들을, 상기 제1 보이스 엔진에 대응하는 제1 보이스 커맨드, 및 상기 제2 보이스 엔진에 대응하는 제2 보이스 커맨드로 구분하기 위한 정보를 획득하는 동작을 포함할 수 있다. 상기 방법은, 상기 정

보에 기반하여, 상기 제1 보이스 커맨드에 대응하는 제1 기능을, 상기 제1 보이스 엔진을 이용하여 실행하고, 상기 외부 전자 장치로 상기 제2 보이스 커맨드에 대응하는 제2 기능의 실행을 요청하는 동작을 포함할 수 있다. 상기 방법은, 상기 전자 장치의 스피커를 통하여, 상기 발언에 의해 지시되는 상기 복수의 보이스 커맨드들의 순서에 기반하여, 상기 제1 기능, 및 상기 제2 기능을 실행한 결과들을 표현한(representing) 제2 오디오 신호를 출력하는 동작을 포함할 수 있다.

도면의 간단한 설명

[10] 도 1은, 일 실시예에 따른, 네트워크 환경 내의 전자 장치의 블록도이다.

[11] 도 2는, 일 실시예에 따른, 전자 장치가 발언(speech)에 기반하여 하나 이상의 기능들을 실행하는 동작의 일 예를 도시한다.

[12] 도 3은, 일 실시예에 따른, 전자 장치, 및 외부 전자 장치의 블록도이다.

[13] 도 4는, 일 실시예에 따른, 전자 장치에 의해 실행되는 프로그램을 설명하기 위한 블록도이다.

[14] 도 5는, 일 실시예에 따른, 전자 장치 내에 포함된 보이스 커맨드 데이터베이스의 블록도이다.

[15] 도 6은, 일 실시예에 따른, 전자 장치에 의해 실행되는 프로그램을 설명하기 위한 블록도이다.

[16] 도 7은, 일 실시예에 따른, 전자 장치에 의해 실행되는 프로그램을 설명하기 위한 블록도이다.

[17] 도 8a 및 도 8b는, 일 실시예에 따른, 전자 장치가 UI(user interface)를 표시하는 예시적인 상태들을 도시한다.

[18] 도 9는, 일 실시예에 따른, 전자 장치가 발언에 기반하여 복수의 기능들을 실행하는 동작을 도시한다.

[19] 도 10은, 일 실시예에 따른, 전자 장치에 의해 수행되는 동작의 흐름도이다.

[20] 도 11은, 일 실시예에 따른, 전자 장치에 의해 수행되는 동작의 흐름도이다.

[21] 도 12는, 일 실시예에 따른, 전자 장치에 의해 수행되는 동작의 흐름도이다.

[22] 도 13은 일 실시 예에 따른 인공 지능(artificial intelligence)(AI) 시스템을 나타낸 블록도이다.

[23] 도 14는, 일 실시 예에 따른, 컨셉과 액션의 관계 정보를 데이터베이스에 저장하기 위한 스키마(schema)를 나타낸 도면이다.

[24] 도 15는, 일 실시 예에 따라, 지능형 앱을 통해 수신된 음성 입력을 처리하는 화면을 표시하는 사용자 단말을 도시하는 도면이다.

발명의 실시를 위한 형태

[25] 이하, 본 문서의 실시예들이 첨부된 도면을 참조하여 설명된다.

[26] 본 문서의 다양한 실시 예들 및 이에 사용된 용어들은 본 문서에 기재된 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 해당 실시 예의 다양한 변경, 균등물, 및/또는 대체물을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관

련하여, 유사한 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함할 수 있다. 본 문서에서, "A 또는 B", "A 및/또는 B 중 적어도 하나", "A, B 또는 C" 또는 "A, B 및/또는 C 중 적어도 하나" 등의 표현은 함께 나열된 항목들의 모든 가능한 조합을 포함할 수 있다. "제 1", "제 2", "첫째" 또는 "둘째" 등의 표현들은 해당 구성요소들을, 순서 또는 중요도에 상관없이 수식할 수 있고, 한 구성요소를 다른 구성요소와 구분하기 위해 사용될 뿐 해당 구성요소들을 한정하지 않는다. 어떤 (예: 제 1) 구성요소가 다른(예: 제 2) 구성요소에 "(기능적으로 또는 통신적으로) 연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 상기 어떤 구성요소가 상기 다른 구성요소에 직접적으로 연결되거나, 다른 구성요소(예: 제 3 구성요소)를 통하여 연결될 수 있다.

[27] 본 문서에서 사용된 용어 "모듈"은 하드웨어, 소프트웨어 또는 펌웨어로 구성된 유닛을 포함하며, 예를 들면, 로직, 논리 블록, 부품, 또는 회로 등의 용어와 상호 호환적으로 사용될 수 있다. 모듈은, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는 최소 단위 또는 그 일부가 될 수 있다. 예를 들면, 모듈은 ASIC(application-specific integrated circuit)으로 구성될 수 있다.

[28] 도 1은, 일 실시예들에 따른, 네트워크 환경(100) 내의 전자 장치(101)의 블록도이다.

[29] 도 1을 참조하면, 네트워크 환경(100)에서 전자 장치(101)는 제 1 네트워크(198)(예: 근거리 무선 통신 네트워크)를 통하여 외부 전자 장치(102)와 통신하거나, 또는 제 2 네트워크(199)(예: 원거리 무선 통신 네트워크)를 통하여 외부 전자 장치(104) 또는 서버(108) 중 적어도 하나 와 통신할 수 있다. 일실시예에 따르면, 전자 장치(101)는 서버(108)를 통하여 외부 전자 장치(104)와 통신할 수 있다. 일실시예에 따르면, 전자 장치(101)는 프로세서(120), 메모리(130), 입력 모듈(150), 음향 출력 모듈(155), 디스플레이 모듈(160), 오디오 모듈(170), 센서 모듈(176), 인터페이스(177), 연결 단자(178), 햅틱 모듈(179), 카메라 모듈(180), 전력 관리 모듈(188), 배터리(189), 통신 모듈(190), 가입자 식별 모듈(196), 또는 안테나 모듈(197)을 포함할 수 있다. 어떤 실시예에서는, 전자 장치(101)에는, 이 구성요소들 중 적어도 하나(예: 연결 단자(178))가 생략되거나, 하나 이상의 다른 구성요소가 추가될 수 있다. 어떤 실시예에서는, 이 구성요소들 중 일부들(예: 센서 모듈(176), 카메라 모듈(180), 또는 안테나 모듈(197))은 하나의 구성요소(예: 디스플레이 모듈(160))로 통합될 수 있다.

[30] 프로세서(120)는, 예를 들면, 소프트웨어(예: 프로그램(140))를 실행하여 프로세서(120)에 연결된 전자 장치(101)의 적어도 하나의 다른 구성요소(예: 하드웨어 또는 소프트웨어 구성요소)를 제어할 수 있고, 다양한 데이터 처리 또는 연산을 수행할 수 있다. 일실시예에 따르면, 데이터 처리 또는 연산의 적어도 일부로서, 프로세서(120)는 다른 구성요소(예: 센서 모듈(176) 또는 통신 모듈(190))로부터 수신된 명령 또는 데이터를 휘발성 메모리(132)에 저장하고, 휘발성 메모리(132)

에 저장된 명령 또는 데이터를 처리하고, 결과 데이터를 비휘발성 메모리(134)에 저장할 수 있다. 일실시예에 따르면, 프로세서(120)는 메인 프로세서(121)(예: 중앙 처리 장치 또는 어플리케이션 프로세서) 또는 이와는 독립적으로 또는 함께 운영 가능한 보조 프로세서(123)(예: 그래픽 처리 장치, 신경망 처리 장치(NPU: neural processing unit), 이미지 시그널 프로세서, 센서 허브 프로세서, 또는 커뮤니케이션 프로세서)를 포함할 수 있다. 예를 들어, 전자 장치(101)가 메인 프로세서(121) 및 보조 프로세서(123)를 포함하는 경우, 보조 프로세서(123)는 메인 프로세서(121)보다 저전력을 사용하거나, 지정된 기능에 특화되도록 설정될 수 있다. 보조 프로세서(123)는 메인 프로세서(121)와 별개로, 또는 그 일부로서 구현될 수 있다.

[31] 보조 프로세서(123)는, 예를 들면, 메인 프로세서(121)가 인액티브(예: 슬립) 상태에 있는 동안 메인 프로세서(121)를 대신하여, 또는 메인 프로세서(121)가 액티브(예: 어플리케이션 실행) 상태에 있는 동안 메인 프로세서(121)와 함께, 전자 장치(101)의 구성요소들 중 적어도 하나의 구성요소(예: 디스플레이 모듈(160), 센서 모듈(176), 또는 통신 모듈(190))와 관련된 기능 또는 상태들의 적어도 일부를 제어할 수 있다. 일실시예에 따르면, 보조 프로세서(123)(예: 이미지 시그널 프로세서 또는 커뮤니케이션 프로세서)는 기능적으로 관련 있는 다른 구성요소(예: 카메라 모듈(180) 또는 통신 모듈(190))의 일부로서 구현될 수 있다. 일실시예에 따르면, 보조 프로세서(123)(예: 신경망 처리 장치)는 인공지능 모델의 처리에 특화된 하드웨어 구조를 포함할 수 있다. 인공지능 모델은 기계 학습을 통해 생성될 수 있다. 이러한 학습은, 예를 들어, 인공지능 모델이 수행되는 전자 장치(101) 자체에서 수행될 수 있고, 별도의 서버(예: 서버(108))를 통해 수행될 수도 있다. 학습 알고리즘은, 예를 들어, 지도형 학습(supervised learning), 비지도형 학습(unsupervised learning), 준지도형 학습(semi-supervised learning) 또는 강화 학습(reinforcement learning)을 포함할 수 있으나, 전술한 예에 한정되지 않는다. 인공지능 모델은, 복수의 인공 신경망 레이어들을 포함할 수 있다. 인공 신경망은 심층 신경망(DNN: deep neural network), CNN(convolutional neural network), RNN(recurrent neural network), RBM(restricted boltzmann machine), DBN(deep belief network), BRDNN(bidirectional recurrent deep neural network), 심층 Q-네트워크(deep Q-networks) 또는 상기 중 둘 이상의 조합 중 하나일 수 있으나, 전술한 예에 한정되지 않는다. 인공지능 모델은 하드웨어 구조 이외에, 추가적으로 또는 대체적으로, 소프트웨어 구조를 포함할 수 있다.

[32] 메모리(130)는, 전자 장치(101)의 적어도 하나의 구성요소(예: 프로세서(120) 또는 센서 모듈(176))에 의해 사용되는 다양한 데이터를 저장할 수 있다. 데이터는, 예를 들어, 소프트웨어(예: 프로그램(140)) 및, 이와 관련된 명령에 대한 입력 데이터 또는 출력 데이터를 포함할 수 있다. 메모리(130)는, 휘발성 메모리(132) 또는 비휘발성 메모리(134)를 포함할 수 있다.

[33] 프로그램(140)은 메모리(130)에 소프트웨어로서 저장될 수 있으며, 예를 들면, 운영 체제(142), 미들 웨어(144) 또는 어플리케이션(146)을 포함할 수 있다.

[34] 입력 모듈(150)은, 전자 장치(101)의 구성요소(예: 프로세서(120))에 사용될 명령 또는 데이터를 전자 장치(101)의 외부(예: 사용자)로부터 수신할 수 있다. 입력 모듈(150)은, 예를 들면, 마이크, 마우스, 키보드, 키(예: 버튼), 또는 디지털 펜(예: 스타일러스 펜)을 포함할 수 있다.

[35] 음향 출력 모듈(155)은 음향 신호를 전자 장치(101)의 외부로 출력할 수 있다. 음향 출력 모듈(155)은, 예를 들면, 스피커 또는 리시버를 포함할 수 있다. 스피커는 멀티미디어 재생 또는 녹음 재생과 같이 일반적인 용도로 사용될 수 있다. 리시버는 착신 전화를 수신하기 위해 사용될 수 있다. 일실시예에 따르면, 리시버는 스피커와 별개로, 또는 그 일부로서 구현될 수 있다.

[36] 디스플레이 모듈(160)은 전자 장치(101)의 외부(예: 사용자)로 정보를 시각적으로 제공할 수 있다. 디스플레이 모듈(160)은, 예를 들면, 디스플레이, 홀로그램 장치, 또는 프로젝터 및 해당 장치를 제어하기 위한 제어 회로를 포함할 수 있다. 일실시예에 따르면, 디스플레이 모듈(160)은 터치를 감지하도록 설정된 터치 센서, 또는 상기 터치에 의해 발생되는 힘의 세기를 측정하도록 설정된 압력 센서를 포함할 수 있다.

[37] 오디오 모듈(170)은 소리를 전기 신호로 변환시키거나, 반대로 전기 신호를 소리로 변환시킬 수 있다. 일실시예에 따르면, 오디오 모듈(170)은, 입력 모듈(150)을 통해 소리를 획득하거나, 음향 출력 모듈(155), 또는 전자 장치(101)와 직접 또는 무선으로 연결된 외부 전자 장치(예: 외부 전자 장치(102))(예: 스피커 또는 헤드폰)를 통해 소리를 출력할 수 있다.

[38] 센서 모듈(176)은 전자 장치(101)의 작동 상태(예: 전력 또는 온도), 또는 외부의 환경 상태(예: 사용자 상태)를 감지하고, 감지된 상태에 대응하는 전기 신호 또는 데이터 값을 생성할 수 있다. 일실시예에 따르면, 센서 모듈(176)은, 예를 들면, 제스처 센서, 자이로 센서, 기압 센서, 마그네틱 센서, 가속도 센서, 그립 센서, 근접 센서, 컬러 센서, IR(infrared) 센서, 생체 센서, 온도 센서, 습도 센서, 또는 조도 센서를 포함할 수 있다.

[39] 인터페이스(177)는 전자 장치(101)가 외부 전자 장치(예: 외부 전자 장치(102))와 직접 또는 무선으로 연결되기 위해 사용될 수 있는 하나 이상의 지정된 프로토콜들을 지원할 수 있다. 일실시예에 따르면, 인터페이스(177)는, 예를 들면, HDMI(high definition multimedia interface), USB(universal serial bus) 인터페이스, SD카드 인터페이스, 또는 오디오 인터페이스를 포함할 수 있다.

[40] 연결 단자(178)는, 그를 통해서 전자 장치(101)가 외부 전자 장치(예: 외부 전자 장치(102))와 물리적으로 연결될 수 있는 커넥터를 포함할 수 있다. 일실시예에 따르면, 연결 단자(178)는, 예를 들면, HDMI 커넥터, USB 커넥터, SD 카드 커넥터, 또는 오디오 커넥터(예: 헤드폰 커넥터)를 포함할 수 있다.

[41] 햅틱 모듈(179)은 전기적 신호를 사용자가 촉각 또는 운동 감각을 통해서 인지할 수 있는 기계적인 자극(예: 진동 또는 움직임) 또는 전기적인 자극으로 변환할 수 있다. 일실시예에 따르면, 햅틱 모듈(179)은, 예를 들면, 모터, 압전 소자, 또는 전기 자극 장치를 포함할 수 있다.

[42] 카메라 모듈(180)은 정지 영상 및 동영상을 촬영할 수 있다. 일실시예에 따르면, 카메라 모듈(180)은 하나 이상의 렌즈들, 이미지 센서들, 이미지 시그널 프로세서들, 또는 플래시들을 포함할 수 있다.

[43] 전력 관리 모듈(188)은 전자 장치(101)에 공급되는 전력을 관리할 수 있다. 일실시예에 따르면, 전력 관리 모듈(188)은, 예를 들면, PMIC(power management integrated circuit)의 적어도 일부로서 구현될 수 있다.

[44] 배터리(189)는 전자 장치(101)의 적어도 하나의 구성요소에 전력을 공급할 수 있다. 일실시예에 따르면, 배터리(189)는, 예를 들면, 재충전 불가능한 1차 전지, 재충전 가능한 2차 전지 또는 연료 전지를 포함할 수 있다.

[45] 통신 모듈(190)은 전자 장치(101)와 외부 전자 장치(예: 외부 전자 장치(102), 외부 전자 장치(104), 또는 서버(108)) 간의 직접(예: 유선) 통신 채널 또는 무선 통신 채널의 수립, 및 수립된 통신 채널을 통한 통신 수행을 지원할 수 있다. 통신 모듈(190)은 프로세서(120)(예: 어플리케이션 프로세서)와 독립적으로 운영되고, 직접(예: 유선) 통신 또는 무선 통신을 지원하는 하나 이상의 커뮤니케이션 프로세서를 포함할 수 있다. 일실시예에 따르면, 통신 모듈(190)은 무선 통신 모듈(192)(예: 셀룰러 통신 모듈, 근거리 무선 통신 모듈, 또는 GNSS(global navigation satellite system) 통신 모듈) 또는 유선 통신 모듈(194)(예: LAN(local area network) 통신 모듈, 또는 전력선 통신 모듈)을 포함할 수 있다. 이들 통신 모듈 중 해당하는 통신 모듈은 제 1 네트워크(198)(예: 블루투스, WiFi(wireless fidelity) direct 또는 IrDA(infrared data association)와 같은 근거리 통신 네트워크) 또는 제 2 네트워크(199)(예: 레거시 셀룰러 네트워크, 5G 네트워크, 차세대 통신 네트워크, 인터넷, 또는 컴퓨터 네트워크(예: LAN 또는 WAN)와 같은 원거리 통신 네트워크)를 통하여 외부의 전자 장치(104)와 통신할 수 있다. 이런 여러 종류의 통신 모듈들은 하나의 구성요소(예: 단일 칩)로 통합되거나, 또는 서로 별도의 복수의 구성요소들(예: 복수 칩들)로 구현될 수 있다. 무선 통신 모듈(192)은 가입자 식별 모듈(196)에 저장된 가입자 정보(예: 국제 모바일 가입자 식별자(IMSI))를 이용하여 제 1 네트워크(198) 또는 제 2 네트워크(199)와 같은 통신 네트워크 내에서 전자 장치(101)를 확인 또는 인증할 수 있다.

[46] 무선 통신 모듈(192)은 4G 네트워크 이후의 5G 네트워크 및 차세대 통신 기술, 예를 들어, NR 접속 기술(new radio access technology)을 지원할 수 있다. NR 접속 기술은 고용량 데이터의 고속 전송(eMBB(enhanced mobile broadband)), 단말 전력 최소화와 다수 단말의 접속(mMTC(massive machine type communications)), 또는 고신뢰도와 저지연(URLLC(ultra-reliable and low-latency communications))을 지원할 수 있다. 무선 통신 모듈(192)은, 예를 들어, 높은 데이터 전송률 달

성을 위해, 고주파 대역(예: mmWave 대역)을 지원할 수 있다. 무선 통신 모듈(192)은 고주파 대역에서의 성능 확보를 위한 다양한 기술들, 예를 들어, 빔포밍(beamforming), 거대 배열 다중 입출력(massive MIMO(multiple-input and multiple-output)), 전차원 다중입출력(FD-MIMO: full dimensional MIMO), 어레이 안테나(array antenna), 아날로그 빔형성(analog beam-forming), 또는 대규모 안테나(large scale antenna)와 같은 기술들을 지원할 수 있다. 무선 통신 모듈(192)은 전자 장치(101), 외부 전자 장치(예: 외부 전자 장치(104)) 또는 네트워크 시스템(예: 제 2 네트워크(199))에 규정되는 다양한 요구사항을 지원할 수 있다. 일실시예에 따르면, 무선 통신 모듈(192)은 eMBB 실현을 위한 Peak data rate(예: 20Gbps 이상), mMTC 실현을 위한 손실 Coverage(예: 164dB 이하), 또는 URLLC 실현을 위한 U-plane latency(예: 다운링크(DL) 및 업링크(UL) 각각 0.5ms 이하, 또는 라운드 트립 1ms 이하)를 지원할 수 있다.

[47] 안테나 모듈(197)은 신호 또는 전력을 외부(예: 외부의 전자 장치)로 송신하거나 외부로부터 수신할 수 있다. 일실시예에 따르면, 안테나 모듈(197)은 서브스트레이트(예: PCB) 위에 형성된 도전체 또는 도전성 패턴으로 이루어진 방사체를 포함하는 안테나를 포함할 수 있다. 일실시예에 따르면, 안테나 모듈(197)은 복수의 안테나들(예: 어레이 안테나)을 포함할 수 있다. 이런 경우, 제 1 네트워크(198) 또는 제 2 네트워크(199)와 같은 통신 네트워크에서 사용되는 통신 방식에 적합한 적어도 하나의 안테나가, 예를 들면, 통신 모듈(190)에 의하여 상기 복수의 안테나들로부터 선택될 수 있다. 신호 또는 전력은 상기 선택된 적어도 하나의 안테나를 통하여 통신 모듈(190)과 외부의 전자 장치 간에 송신되거나 수신될 수 있다. 어떤 실시예에 따르면, 방사체 이외에 다른 부품(예: RFIC(radio frequency integrated circuit))이 추가로 안테나 모듈(197)의 일부로 형성될 수 있다.

[48] 다양한 실시예에 따르면, 안테나 모듈(197)은 mmWave 안테나 모듈을 형성할 수 있다. 일실시예에 따르면, mmWave 안테나 모듈은 인쇄 회로 기판, 상기 인쇄 회로 기판의 제 1 면(예: 아래 면)에 또는 그에 인접하여 배치되고 지정된 고주파 대역(예: mmWave 대역)을 지원할 수 있는 RFIC, 및 상기 인쇄 회로 기판의 제 2 면(예: 윗 면 또는 측 면)에 또는 그에 인접하여 배치되고 상기 지정된 고주파 대역의 신호를 송신 또는 수신할 수 있는 복수의 안테나들(예: 어레이 안테나)을 포함할 수 있다.

[49] 상기 구성요소들 중 적어도 일부는 주변 기기들간 통신 방식(예: 버스, GPIO(general purpose input and output), SPI(serial peripheral interface), 또는 MIPI(mobile industry processor interface))을 통해 서로 연결되고 신호(예: 명령 또는 데이터)를 상호간에 교환할 수 있다.

[50] 일실시예에 따르면, 명령 또는 데이터는 제 2 네트워크(199)에 연결된 서버(108)를 통해서 전자 장치(101)와 외부의 전자 장치(104)간에 송신 또는 수신될 수 있다. 외부의 외부 전자 장치(102, 또는 104) 각각은 전자 장치(101)와 동일한 또는 다른 종류의 장치일 수 있다. 일실시예에 따르면, 전자 장치(101)에서 실행

되는 동작들의 전부 또는 일부는 외부의 전자 장치들(102, 104, 또는 108) 중 하나 이상의 외부의 전자 장치들에서 실행될 수 있다. 예를 들면, 전자 장치(101)가 어떤 기능이나 서비스를 자동으로, 또는 사용자 또는 다른 장치로부터의 요청에 반응하여 수행해야 할 경우에, 전자 장치(101)는 기능 또는 서비스를 자체적으로 실행시키는 대신에 또는 추가적으로, 하나 이상의 외부의 전자 장치들에게 그 기능 또는 그 서비스의 적어도 일부를 수행하라고 요청할 수 있다. 상기 요청을 수신한 하나 이상의 외부의 전자 장치들은 요청된 기능 또는 서비스의 적어도 일부, 또는 상기 요청과 관련된 추가 기능 또는 서비스를 실행하고, 그 실행의 결과를 전자 장치(101)로 전달할 수 있다. 전자 장치(101)는 상기 결과를, 그대로 또는 추가적으로 처리하여, 상기 요청에 대한 응답의 적어도 일부로서 제공할 수 있다. 이를 위하여, 예를 들면, 클라우드 컴퓨팅, 분산 컴퓨팅, 모바일 에지 컴퓨팅(MEC: mobile edge computing), 또는 클라이언트-서버 컴퓨팅 기술이 이용될 수 있다. 전자 장치(101)는, 예를 들어, 분산 컴퓨팅 또는 모바일 에지 컴퓨팅을 이용하여 초저지연 서비스를 제공할 수 있다. 다른 실시예에 있어서, 외부의 전자 장치(104)는 IoT(internet of things) 기기를 포함할 수 있다. 서버(108)는 기계 학습 및/또는 신경망을 이용한 지능형 서버일 수 있다. 일실시예에 따르면, 외부의 전자 장치(104) 또는 서버(108)는 제 2 네트워크(199) 내에 포함될 수 있다. 전자 장치(101)는 5G 통신 기술 및 IoT 관련 기술을 기반으로 지능형 서비스(예: 스마트 홈, 스마트 시티, 스마트 카, 또는 헬스 케어)에 적용될 수 있다.

[51] 도 2는, 일 실시예에 따른, 전자 장치(101)가 발언(speech)(220)에 기반하여 하나 이상의 기능들(functions)을 실행하는 동작을 도시한다.

[52] 도 2의 전자 장치(101)는, 도 1의 전자 장치(101)를 포함할 수 있다. 전자 장치(101)는 사용자에 의해 소유되는(be owned by) 단말일 수 있다. 단말은, 예를 들어, 랩톱 및 데스크톱과 같은 개인용 컴퓨터(personal computer, PC), 스마트폰(smartphone), 스마트패드(smartpad), 태블릿 PC를 포함할 수 있다. 단말은, 스마트워치(smartwatch) 및/또는 HMD(head-mounted device)와 같은 스마트액세서리를 포함할 수 있다.

[53] 일 실시예에 따른, 전자 장치(101)는 소리를 이용하여 사용자와 상호작용(interact with)할 수 있다. 전자 장치(101)는 마이크(예, 도 1의 입력 모듈(150))를 이용하여, 전자 장치(101)를 포함하는 외부 공간의 소리에 종속된(dependent to) 전기 신호인, 오디오 신호를 획득할 수 있다. 전자 장치(101)는 상기 오디오 신호로부터, 사용자의 발언(220)을 식별할 수 있다. 발언(220)을 식별하는 것에 응답하여, 전자 장치(101)는 발언(220)과 관련된 하나 이상의 기능들을 실행할 수 있다. 전자 장치(101)는 상기 하나 이상의 기능들을 실행한 결과들을 포함하는 정보를, 상기 사용자에게 출력할 수 있다. 도 2를 참고하면, 전자 장치(101)는 발언(250)을 포함하는 오디오 신호를 생성할 수 있다. 전자 장치(101)는 상기 오디오 신호에 기반하여 스피커(예, 도 1의 음향 출력 모듈(155))를 제어하여, 상기 오디

오 신호에 종속된 소리를 재생할 수 있다. 마이크, 및/또는 스피커를 포함하는 전자 장치(101)의 구조가, 도 3을 참고하여 설명된다.

[54] 일 실시예에 따른, 전자 장치(101)는, 발언(220)으로부터 전자 장치(101)와 관련된 적어도 하나의 기능을 실행함을 나타내는 입력을 식별할 수 있다. 이하에서, 보이스 커맨드(voice command)는, 발언(220)과 같이 소리에 기반하여 전자 장치(101)로 입력된 커맨드(예, 자연어의 커맨드)를 포함할 수 있다. 상기 보이스 커맨드는, 전자 장치(101)에 의해 식별된 소리에 의하여, 전자 장치(101), 및/또는 전자 장치(101)에 연결된 외부 전자 장치(210)가 실행하는 기능, 및/또는 작업(task)의 단위일 수 있다. 보이스 커맨드를 실행하는 것은, 전자 장치(101)가 보이스 커맨드에 대응하는 적어도 하나의 기능을 실행하는 동작을 포함할 수 있다. 전자 장치(101)는 발언(220)에 포함된 자연어(natural language)에 기반하여, 발언(220)에 포함된 하나 이상의 보이스 커맨드들을 식별할 수 있다. 전자 장치(101)에 연결된 상기 외부 전자 장치(210)는, 발언(220)을 포함하는 오디오 신호를 처리하기 위해 전자 장치(101)에 연결된 서버(예, 도 1의 서버(108))를 포함할 수 있다.

[55] 도 2를 참고하면, 전자 장치(101)가 마이크를 통하여 식별된 오디오 신호로부터, 복수의 보이스 커맨드들(예, 제1 보이스 커맨드(230), 및/또는 제2 보이스 커맨드(240))를 식별한 예시적인 케이스가 도시된다. 전자 장치(101)는 상기 오디오 신호로부터 복수의 기능들을 순차적으로(sequentially) 실행하기 위한 발언(220)을 식별할 수 있다. 발언(220)은, 전자 장치(101)에 의한 발언(220)의 인식을 촉발하기(triggering) 위한 지정된 자연어(예, "헤이, 빅스비"와 같은 단어들)를 포함할 수 있다. 상기 지정된 자연어를 식별한 전자 장치(101)는, 발언(220)을 인식하여, 발언(220) 내에 포함된 하나 이상의 보이스 커맨드들을 식별할 수 있다. 예를 들어, 전자 장치(101)에 의해 인식되는 발언(220)은, 상기 복수의 기능들 각각을 표현하는 자연어(예, 단어, 구, 및/또는 문장)를 포함할 수 있다. 예를 들어, 전자 장치(101)에 의해 인식되는 발언(220)은, 상기 복수의 기능들의 그룹에 할당된(assigned) 명칭(name)을 포함할 수 있다. 상기 그룹에 할당된 명칭이, 퀵 커맨드, 및/또는 단축어로 지칭될 수 있다. 상기 명칭을 발언(220)으로부터 식별하는 것에 기반하여, 전자 장치(101)는 상기 명칭에 의해 구분되는 그룹에 포함된, 복수의 기능들을 실행할 수 있다. 복수의 기능들을 실행하기 위한 보이스 커맨드들의 그룹은, 전자 장치(101), 및 사용자 사이의 상호작용에 기반하여, 전자 장치(101)로 등록될 수 있다. 전자 장치(101)가 상기 그룹을 등록하기 위한 입력을 식별하기 위해 표시하는 UI(user interface)의 일 예가, 도 8a를 참고하여 설명된다.

[56] 일 실시예에 따른, 전자 장치(101)는 제1 보이스 커맨드(230), 및 제2 보이스 커맨드(240)의 순차적인 실행을 위한 발언(220)을 식별하는 것에 기반하여, 상기 제1 보이스 커맨드(230), 및 상기 제2 보이스 커맨드(240)에 대응하는 기능들을 실행할 수 있다. 상기 기능들은, 전자 장치(101), 및/또는 외부 전자 장치(210)에 의해 실행되는 어플리케이션에 기반하여, 실행될 수 있다. 보이스 커맨드를 실행하기 위하여, 전자 장치(101), 및/또는 외부 전자 장치(210)에 의해 실행되는 상기 어

플리케이션이, 보이스 엔진(voice engine), 및/또는 NLU(natural language unit)로 지칭될 수 있다. 보이스 엔진이 실행되는 전자 장치(예, 전자 장치(101), 및/또는 외부 전자 장치(210))는, 보이스 커맨드에 매칭된 동작을 수행하도록, 제어될 수 있다. 보이스 엔진의 실행에 의해 보이스 커맨드에 매칭된 상기 동작이 수행되는 기간(duration), 및/또는 속도는, 보이스 엔진이 실행되는 전자 장치의 성능, 및/또는 리소스의 상태에 종속될 수 있다.

[57] 보이스 엔진들이 네트워크를 통해 연결된 전자 장치(101), 및 외부 전자 장치(210) 전부에서 실행되는 도 2의 일 실시예에서, 제1 보이스 커맨드(230), 및 제2 보이스 커맨드(240)에 대응하는 기능들이, 상기 보이스 엔진들 중 적어도 하나에 의해 실행될 수 있다. 제1 보이스 커맨드(230), 및 제2 보이스 커맨드(240)의 순차적인 실행을 위한 발언(220)을 식별한 전자 장치(101)는, 전자 장치(101)에 의해 실행되는 제1 보이스 엔진, 및 외부 전자 장치(210)에 의해 실행되는 제2 보이스 엔진으로, 제1 보이스 커맨드(230)에 대응하는 제1 기능, 및 제2 보이스 커맨드(240)에 대응하는 제2 기능 각각을 매칭할 수 있다. 예를 들어, 전자 장치(101)는 제1 보이스 엔진, 및 제2 보이스 엔진과, 제1 보이스 커맨드(230), 및 제2 보이스 커맨드(240)를 매칭하여, 상기 제1 기능, 및 상기 제2 기능의 실행을 제어하거나, 또는 스케줄링할 수 있다. 예를 들어, 상기 제1 기능, 및 상기 제2 기능 전부가, 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 중 어느 하나에 의해 집중적으로 실행될 수 있다. 예를 들어, 상기 제1 기능, 및 상기 제2 기능 각각이 상기 제1 보이스 엔진, 및 상기 제2 보이스 엔진에 의해 분산하여 실행될 수 있다. 전자 장치(101)가 복수의 보이스 엔진들, 및 발언(220)에 포함된 복수의 보이스 커맨드들을 매칭하기 위해 실행하는 어플리케이션의 일 예가, 도 3 내지 도 7을 참고하여 설명된다.

[58] 도 2를 참고하면, 일 실시예에 따른, 전자 장치(101)가, 상이한 보이스 엔진들을 이용하여 제1 보이스 커맨드(230), 및 제2 보이스 커맨드(240) 각각의 기능들을 실행한 예시적인 상태가 도시된다. 전자 장치(101)는, 전자 장치(101)를 포함하는 상이한 전자 장치들에 의해 실행되는 보이스 엔진들의 상태들에 기반하여, 발언(220)에 포함된 적어도 하나의 보이스 커맨드의 실행에 이용될 보이스 엔진을 선택할 수 있다. 예를 들어, 전자 장치(101)에 의해 실행되는 제1 보이스 엔진의 상태, 또는 외부 전자 장치(210)에 의해 실행되는 제2 보이스 엔진의 상태 중 적어도 하나를 모니터링할 수 있다. 보이스 엔진의 상태는, 상기 보이스 엔진의 실행에 기반하여 보이스 커맨드에 대응하는 기능을 실행 가능한지 여부, 상기 기능의 실행에 요구되는 시간, 및/또는 상기 기능을 실행하기 이전의 대기 시간(latency time)과 관련될 수 있다. 모니터링된 상태들에 기반하여, 전자 장치(101)는 발언(220)에 포함된 제1 보이스 커맨드(230), 및 제2 보이스 커맨드(240) 각각의 실행에 이용될 보이스 엔진을, 상기 제1 보이스 엔진, 및 상기 제2 보이스 엔진 중에서 선택할 수 있다. 도 2의 예시적인 케이스 내에서, 전자 장치(101)는, 상기 제1 보이

스 엔진, 및 제1 보이스 커맨드(230)를 매칭하고, 상기 제2 보이스 엔진, 및 제2 보이스 커맨드(240)를 매칭할 수 있다.

[59] 상기 제1 보이스 엔진에 제1 보이스 커맨드(230)를 매칭하는 것에 기반하여, 전자 장치(101)는 제1 보이스 커맨드(230)에 대응하는 제1 기능을 실행할 수 있다. 전자 장치(101)는 상기 제1 기능을 실행한 결과를 포함하는 제1 정보(235)를, 상기 제1 보이스 엔진의 실행에 기반하여, 획득할 수 있다. 상기 제2 보이스 엔진에 제2 보이스 커맨드(240)를 매칭하는 것에 기반하여, 전자 장치(101)는 제2 보이스 커맨드(240)에 대응하는 제2 기능을 실행할 것을, 상기 제2 보이스 엔진을 실행하는 외부 전자 장치(210)로 요청할 수 있다. 전자 장치(101)는, 외부 전자 장치(210)로부터 상기 요청에 대한 응답으로(as a response to), 상기 제2 기능과 관련된 제2 정보(245)를 수신할 수 있다. 상기 제2 정보(245)는, 외부 전자 장치(210), 및/또는 제2 보이스 엔진에 기반하여, 상기 제2 기능을 실행한 결과를 포함할 수 있다. 상기 제2 정보(245)는, 상기 제2 기능의 실행에 기반하여 전자 장치(101)를 제어하기 위한 정보를 포함할 수 있다.

[60] 일 실시예에 따른, 전자 장치(101)는 상기 제1 정보(235), 및/또는 제2 정보(245)를 획득하는 것에 기반하여, 제1 보이스 커맨드(230)에 매칭된 제1 기능, 및 제2 보이스 커맨드(240)에 매칭된 제2 기능을 실행한 결과들을 표현하는 발언(250)을 포함하는 오디오 신호를 생성할 수 있다. 전자 장치(101)는 하나 이상의 자연어 문장들에 기반하여, 상기 제1 기능을 실행한 결과를 포함하는 상기 제1 정보(235), 및 상기 제2 기능을 실행한 결과를 포함하는 상기 제2 정보(245)를 표현한, 발언(250)을 획득할 수 있다. 전자 장치(101)는 발언(250)을 포함하는 상기 오디오 신호를 생성할 수 있다. 전자 장치(101)는 생성된 오디오 신호에 기반하여 스피커를 제어하여, 상기 스피커를 통하여 상기 오디오 신호에 종속된 소리를 출력할 수 있다. 전자 장치(101)가 발언(250)을 출력하는 형태는, 도 2의 오디오 신호, 및/또는 소리에 제한되지 않는다. 전자 장치(101)가 디스플레이(260)와 같은 하드웨어를 이용하여, 발언(220)에 대한 적어도 하나의 보이스 커맨드를 실행한 결과를 시각화하는 일 실시예가, 도 8b를 참고하여 설명된다.

[61] 상술한 바와 같이, 일 실시예에 따른, 전자 장치(101)는 마이크로부터 출력된 오디오 신호로부터 복수의 기능들을 순차적으로 실행하기 위한 발언(220)을 식별하는 것에 기반하여, 상기 복수의 기능들을 실행할 보이스 엔진들을 스케줄링할 수 있다. 예를 들어, 전자 장치(101)는 상기 보이스 엔진들 중 적어도 하나의 상태에 기반하여, 상기 복수의 기능들을, 제1 보이스 엔진에 의해 실행될 적어도 하나의 제1 기능, 및 제2 보이스 엔진에 의해 실행될 적어도 하나의 제2 기능으로 구분할 수 있다. 보이스 엔진들을 스케줄링한 결과에 기반하여, 전자 장치(101)는 상기 복수의 기능들 중 적어도 하나의 제1 기능을 실행하고, 외부 전자 장치(210)를 이용하여, 상기 복수의 기능들 중에서, 상기 적어도 하나의 제1 기능과 상이한 적어도 하나의 제2 기능을 실행할 수 있다. 도 2의 제1 정보(235)는, 전자 장치(101)에 의해 실행되는 제1 보이스 엔진에 기반하여, 상기 적어도 하나의 제1

기능을 실행한 결과의 일 예일 수 있다. 도 2의 제2 정보(245)는, 외부 전자 장치(210)에 의해 실행되는 제2 보이스 엔진에 기반하여, 상기 적어도 하나의 제2 기능을 실행한 결과의 일 예일 수 있다. 전자 장치(101)는, 스피커를 통하여, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능을 실행한 결과들을 표현하는 제2 오디오 신호들을 출력할 수 있다. 전자 장치(101)가 상기 제2 오디오 신호들을 출력하는 순서는, 발언(220)과 관련된, 상기 복수의 기능들의 순서에 매칭될 수 있다. 발언(220)에 대응하는 복수의 기능들이 상이한 전자 장치들에 의해 실행되는 보이스 엔진들에 기반하는 분산 처리에 의해 실행되기 때문에, 전자 장치(101)는 상기 복수의 기능들을 실행한 결과들을 보다 신속하게 획득할 수 있다.

[62] 이하에서는, 도 3을 참고하여, 제1 보이스 커맨드(230), 및 제2 보이스 커맨드(240)를 포함하는 보이스 커맨드를 처리하기 위한 전자 장치(101), 및 외부 전자 장치(210)의 예시적인 구조가 도 3을 참고하여 설명된다.

[63] 도 3은, 일 실시예에 따른, 전자 장치(101), 및 외부 전자 장치(210)의 블록도이다. 도 3의 전자 장치(101)는 도 1 내지 도 2의 전자 장치(101)를 포함할 수 있다. 도 3의 외부 전자 장치(210)는 도 2의 외부 전자 장치(210)를 포함할 수 있다.

[64] 전자 장치(101)는, 프로세서(120), 메모리(130), 디스플레이(260), 스피커(310), 마이크(320), 또는 통신 회로(330) 중 적어도 하나를 포함할 수 있다. 프로세서(120), 메모리(130), 디스플레이(260), 스피커(310), 마이크(320), 및 통신 회로(330)는 통신 버스(a communication bus)(305)와 같은 전자 부품(electronic component)에 의해 서로 전기적으로 및/또는 작동적으로 연결될 수 있다(electronically and/or operably coupled with each other). 이하에서, 하드웨어들이 작동적으로 결합된 것은, 하드웨어들 중 제1 하드웨어에 의해 제2 하드웨어가 제어되도록, 하드웨어들 사이의 직접적인 연결, 또는 간접적인 연결이 유선으로, 또는 무선으로 수립된 것을 의미할 수 있다. 상이한 블록들에 기반하여 도시되었으나, 실시예가 이에 제한되는 것은 아니며, 도 3의 하드웨어들 중 일부분(예, 프로세서(120), 메모리(130), 및 통신 회로(330)의 적어도 일부분)이 SoC(system on a chip)와 같이 단일 집적 회로(single integrated circuit)에 포함될 수 있다. 전자 장치(101)에 포함된 하드웨어 컴포넌트의 타입 및/또는 개수는 도 3에 도시된 바에 제한되지 않는다. 예를 들어, 전자 장치(101)는 도 3에 도시된 하드웨어 중 일부만 포함할 수 있다.

[65] 전자 장치(101)의 프로세서(120)는, 하나 이상의 인스트럭션들에 기반하여 데이터를 처리하기 위한 하드웨어 컴포넌트를 포함할 수 있다. 데이터를 처리하기 위한 하드웨어 컴포넌트는, 예를 들어, ALU(arithmetic and logic unit), FPU(floating point unit), FPGA(field programmable gate array), CPU(central processing unit), 및/또는 AP(application processor)를 포함할 수 있다. 일 실시예에 따르면, 프로세서(120)는 복수의 프로세서들을 포함할 수 있다. 예를 들어, 프로세서(120)는 듀얼 코어(dual core), 쿼드 코어(quad core) 또는 헥사 코어(hexa core)

와 같은 멀티-코어 프로세서의 구조를 가질 수 있다. 도 3의 프로세서(120)는 도 1의 프로세서(120)를 포함할 수 있다.

[66] 전자 장치(101)의 메모리(130)는, 프로세서(120)에 입력 및/또는 출력되는 데이터 및/또는 인스트럭션을 저장하기 위한 하드웨어 컴포넌트를 포함할 수 있다. 메모리(130)는, 예를 들어, RAM(random-access memory)과 같은 휘발성 메모리(volatile memory) 및/또는 ROM(read-only memory)과 같은 비휘발성 메모리(non-volatile memory)를 포함할 수 있다. 휘발성 메모리는, 예를 들어, DRAM(dynamic RAM), SRAM(static RAM), Cache RAM, PSRAM (pseudo SRAM) 중 적어도 하나를 포함할 수 있다. 비휘발성 메모리는, 예를 들어, PROM(programmable ROM), EPROM (erasable PROM), EEPROM (electrically erasable PROM), 플래시 메모리, 하드디스크, 컴팩트 디스크, SSD(solid state drive), eMMC(embedded multi media card) 중 적어도 하나를 포함할 수 있다. 도 3의 메모리(130)는 도 1의 메모리(130)를 포함할 수 있다.

[67] 전자 장치(101)의 디스플레이(260)는 사용자에게 시각화된 정보(예를 들어, 도 8a 및 도 8b의 화면들 중 적어도 하나)를 출력할 수 있다. 예를 들어, 디스플레이(260)는, GPU(graphic processing unit), 및/또는 프로세서(120)와 같은 컨트롤러에 의해 제어되어, 사용자에게 시각화된 정보(visualized information)를 출력할 수 있다. 디스플레이(260)는 FPD(flat panel display) 및/또는 전자 종이(electronic paper)를 포함할 수 있다. 상기 FPD는 LCD(liquid crystal display), PDP(plasma display panel) 및/또는 하나 이상의 LED(light emitting diode)를 포함할 수 있다. 상기 LED는 OLED(organic LED)를 포함할 수 있다. 도 3의 디스플레이(260)는 도 1의 디스플레이 모듈(160)을 포함할 수 있다.

[68] 일 실시예에 따른, 전자 장치(101)의 디스플레이(260)는, 디스플레이(260) 상의 외부 객체(예, 사용자의 손가락)를 탐지하기 위한 센서(예, TSP(touch sensor panel))를 포함할 수 있다. 예를 들어, TSP에 기반하여, 전자 장치(101)는 디스플레이(260)와 접촉하거나, 또는 디스플레이(260) 상을 부유하는(floating) 외부 객체를 탐지할 수 있다. 상기 외부 객체를 탐지하는 것에 응답하여, 전자 장치(101)는 디스플레이(260) 내에 표시되고 있는 시각적 객체들 중에서 상기 외부 객체의 디스플레이(260) 상에서의 위치에 대응하는 특정 시각적 객체와 관련된 기능을 실행할 수 있다.

[69] 전자 장치(101)는, 정보를 시각화한 형태 외에 다른 형태로 출력하기 위한 출력 수단으로, 스피커(310)를 포함할 수 있다. 스피커(310)는 프로세서(120)로부터 수신된 오디오 신호(예, 도 2의 발언(250)을 포함하는 오디오 신호)에 의하여 진동되는 회로 요소를 포함할 수 있다. 전자 장치(101) 내에 포함된 스피커(310)의 개수는, 도 3의 일 예에 제한되지 않으며, 전자 장치(101)는 하나 이상의 스피커들을 포함할 수 있다. 전자 장치(101)는 정보를 시각적인 형태, 청각적인 형태 외에 다른 형태로 출력하기 위한 다른 출력 수단을 포함할 수 있다. 예를 들어, 전자 장치(101)는 진동에 기반하는 햅틱 피드백을 제공하기 위한 모터를 포함할 수 있다.

[70] 전자 장치(101)의 마이크(320)는 대기의 진동을 지시하는 전기 신호를 출력할 수 있다. 예를 들어, 전자 장치(101)는 마이크(320)로부터 출력된 전기 신호인, 오디오 신호로부터, 사용자의 발언(예, 도 2의 발언(220))을 식별할 수 있다. 오디오 신호에 포함된 사용자의 발언은, 프로세서(120)에 의해 실행되는 어플리케이션, 및/또는 프로세스인, 음성 인식 모델 및/또는 자연어 이해 모델에 기반하여, 전자 장치(101)에 의해 인식 가능한 포맷의 정보로 변환될 수 있다. 예를 들어, 전자 장치(101)는 사용자의 발언을 인식하여, 전자 장치(101)에 의해 제공 가능한 복수의 기능들 중에서 하나 이상의 기능들을 실행할 수 있다. 도 3의 스피커(310), 및/또는 마이크(320)는, 도 1의 음향 출력 모듈(155), 및/또는 오디오 모듈(170)을 포함할 수 있다.

[71] 전자 장치(101)의 통신 회로(330)는, 전자 장치(101), 및 외부 전자 장치(210) 사이의 전기 신호의 송신 및/또는 수신을 지원하기 위한 하드웨어를 포함할 수 있다. 통신 회로(330)를 통해 전자 장치(101)에 연결된 외부 전자 장치(210)의 개수는, 도 2, 및/또는 도 3의 일 실시예에 제한되지 않는다. 통신 회로(330)는, 예를 들어, 모뎀(MODEM), 안테나, O/E(optic/electronic) 변환기 중 적어도 하나를 포함할 수 있다. 통신 회로(330)는, 이더넷(ethernet), LAN(local area network), WAN(wide area network), WiFi(wireless fidelity), Bluetooth, BLE(bluetooth low energy), ZigBee, LTE(long term evolution), 5G NR(new radio)와 같은 다양한 타입의 프로토콜에 기반하여 전기 신호의 송신 및/또는 수신을 지원할 수 있다. 도 3의 통신 회로(330)는 도 1의 통신 모듈(190), 가입자 식별 모듈(196), 및/또는 안테나 모듈(197)을 포함할 수 있다.

[72] 메모리(130) 내에서, 프로세서(120)가 데이터에 수행할 연산, 및/또는 동작을 나타내는 하나 이상의 인스트럭션들(또는 명령어들)이 저장될 수 있다. 하나 이상의 인스트럭션들의 집합은, 펌웨어, 운영 체제(예, 도 1의 운영 체제(142)), 프로그램(예, 도 1의 프로그램(140)), 프로세스, 루틴, 서브-루틴 및/또는 어플리케이션(예, 도 1의 어플리케이션(146))으로 참조될 수 있다. 예를 들어, 전자 장치(101), 및/또는 프로세서(120)는, 운영체제, 펌웨어, 드라이버, 및/또는 어플리케이션 형태로 배포된 복수의 인스트럭션의 집합(set of a plurality of instructions)이 실행될 시에, 도 10 내지 도 12의 동작들 중 적어도 하나를 수행할 수 있다. 이하에서, 어플리케이션이 전자 장치(101)에 설치되었다는 것은, 어플리케이션의 형태로 제공된 하나 이상의 인스트럭션들이 전자 장치(101)의 메모리(130) 내에 저장된 것으로써, 상기 하나 이상의 어플리케이션들이 전자 장치(101)의 프로세서(120)에 의해 실행 가능한(executable) 포맷(예, 전자 장치(101)의 운영 체제에 의해 지정된 확장자를 가지는 파일)으로 저장된 것을 의미할 수 있다.

[73] 도 3의 일 실시예에서, 외부 전자 장치(210)는 음성 인식과 관련된 기능을 실행하기 위하여, 전자 장치(101)와 유선으로, 또는 무선으로 연결될 수 있다. 외부 전자 장치(210)는 서버를 포함할 수 있다. 서버는 하나 이상의 PC 및/또는 워크스테이션을 포함할 수 있다. 일 실시예에서, 상기 서버는, 전자 장치(101)의 마이크

(320)로부터 출력된 오디오 신호로부터 식별된 발언에 대응하는 하나 이상의 기능들을 실행하는 서비스로써, 예를 들어, 음성 인식 서비스를 포함할 수 있다. 외부 전자 장치(210)는, 프로세서(120), 메모리(130), 및 통신 회로(330) 중 적어도 하나를 포함할 수 있다. 외부 전자 장치(210) 내에서, 프로세서(120), 메모리(130), 및 통신 회로(330)는, 통신 버스(305)를 통해 전기적으로, 및/또는 작동적으로 결합될 수 있다. 외부 전자 장치(210) 내에 포함된, 프로세서(120), 메모리(130), 및 통신 회로(330)는, 전자 장치(101)의 프로세서(120), 메모리(130), 및 통신 회로(330)에 대응하는 하드웨어 부품, 및/또는 회로를 포함할 수 있다. 이하에서, 반복을 줄이기 위하여, 외부 전자 장치(210) 내에 포함된 프로세서(120), 메모리(130), 및 통신 회로(330)의 설명은, 전자 장치(101) 내 프로세서(120), 메모리(130), 및 통신 회로(330)와 중복되는 범위 내에서 생략될 수 있다.

[74] 도 3을 참고하면, 전자 장치(101)의 프로세서(120)에 의해 실행되는 프로그램으로, 입력 수신기(340), 퀵 커맨드 처리기(350), 발언 처리기(360), 및/또는 제1 보이스 엔진(370)이 도시된다. 예를 들어, 메모리(130) 내에 저장된 복수의 인스트럭션들의 집합, 및/또는 프로세서(120)에 의해 실행되는 프로세스들이, 입력 수신기(340), 퀵 커맨드 처리기(350), 발언 처리기(360), 및/또는 제1 보이스 엔진(370)로 구분될 수 있다. 입력 수신기(340), 퀵 커맨드 처리기(350), 발언 처리기(360), 및/또는 제1 보이스 엔진(370)은, 전자 장치(101) 내에 설치된 일 어플리케이션 내에 포함될 수 있다. 상기 어플리케이션은, 자연어 문장을 포함하는 발언에 기반하여 전자 장치(101)를 제어하기 위하여, 전자 장치(101)의 프로세서(120)에 의해 실행될 수 있다.

[75] 도 3을 참고하면, 외부 전자 장치(210)의 프로세서(120)에 의해 실행되는 프로그램으로, 제2 보이스 엔진(380)이 도시된다. 제2 보이스 엔진(380)은, 메모리(130) 내에 저장된 복수의 어플리케이션들의 집합(예, 어플리케이션)을 포함할 수 있다. 제2 보이스 엔진(380)은, 외부 전자 장치(210)에 연결된 다른 전자 장치(예, 전자 장치(101))로부터 요청된 보이스 커맨드를 실행하기 위하여, 외부 전자 장치(210)의 프로세서(120)에 의해 실행될 수 있다. 전자 장치(101)는, 외부 전자 장치(210)에 의해 실행되는 상기 제2 보이스 엔진(380)을 이용하여 상기 전자 장치(101)에 의해 식별된 적어도 하나의 보이스 커맨드를 실행하기 위하여, 통신 회로(330)를 이용하여 외부 전자 장치(210)와 통신 링크를 수립할 수 있다. 예를 들어, 전자 장치(101)는 마이크(320)로부터 출력된 오디오 신호로부터, 적어도 하나의 보이스 커맨드에 대응하는 자연어 문장의 발화를 가이드하기 위한 지정된 자연어를 식별하는 것에 기반하여, 상기 통신 링크를 수립할 수 있다.

[76] 일 실시예에 따른, 전자 장치(101)는 입력 수신기(340)의 실행에 기반하여, 마이크(320)로부터 출력된 오디오 신호로부터 발언을 식별할 수 있다. 입력 수신기(340)가 실행된 상태 내에서, 전자 장치(101)는 마이크(320)를 통하여 오디오 신호를 수신하는 것에 기반하여, 오디오 신호에 포함된 발언을 식별하기 위한 동작을 수행할 수 있다. 입력 수신기(340)에 의해 처리된 상기 오디오 신호는, 전자 장

치(101)의 프로세서(120)에 의해 외부 전자 장치(210)의 제2 보이스 엔진(380)으로 송신될 수 있다. 입력 수신기(340)는 마이크(320)로부터 출력된 오디오 신호뿐만 아니라, 디스플레이(260)를 통해 표시되는 소프트웨어 키보드, 및/또는 전자 장치(101)에 연결된 하드웨어 키보드를 통해 입력된 하나 이상의 문자들로부터, 사용자의 발언에 대응하는 텍스트를 식별할 수 있다.

[77] 전자 장치(101)는, 제2 보이스 엔진(380)을 실행하는 외부 전자 장치(210)로부터, 상기 오디오 신호에 포함된 발언을 나타내는 텍스트를 획득할 수 있다. 전자 장치(101)는 퀵 커맨드 처리기(350)의 실행에 기반하여, 외부 전자 장치(210)로부터 획득된 상기 텍스트가 하나 이상의 보이스 커맨드들의 그룹에 할당된 명칭(예, 퀵 커맨드)을 포함하는지 여부를 식별할 수 있다. 전자 장치(101)는 퀵 커맨드 처리기(350)의 실행에 기반하여, 사용자로부터 상기 명칭, 및 상기 명칭에 대응하는 하나 이상의 보이스 커맨드들의 그룹을 획득할 수 있다. 전자 장치(101)는 퀵 커맨드 처리기(350)의 실행에 기반하여, 사용자의 발언이 상기 그룹에 할당된 명칭을 포함하는지 여부, 또는 상기 그룹을 등록하기 위한 사용자 입력을 처리할 수 있다. 퀵 커맨드 처리기(350)의 예시적인 구조가, 도 6을 참고하여 설명된다.

[78] 마이크(320)를 통해 획득된 오디오 신호에 포함된 발언이 하나 이상의 보이스 커맨드들의 그룹에 할당된 명칭을 포함하는 경우, 전자 장치(101)는 발언 처리기(360)의 실행에 기반하여, 상기 그룹에 포함된 하나 이상의 보이스 커맨드들의 실행을 제어할 수 있다. 전자 장치(101)는 발언 처리기(360)의 실행에 기반하여, 상기 하나 이상의 보이스 커맨드들이 입력될 보이스 엔진을, 제1 보이스 엔진(370), 또는 제2 보이스 엔진(380) 중에서 선택할 수 있다. 예를 들어, 전자 장치(101)는 발언 처리기(360)의 실행에 기반하여, 상기 하나 이상의 보이스 커맨드들에 대응하는 하나 이상의 기능들의 실행을 스케줄링할 수 있다. 상기 발언으로부터 복수의 보이스 커맨드들을 식별하는 것에 기반하여, 전자 장치(101)는 보이스 커맨드를 처리하기 위해 실행되는 하나 이상의 보이스 엔진들(예, 제1 보이스 엔진(370), 및/또는 제2 보이스 엔진(380))의 상태들을 획득할 수 있다. 상기 상태들에 기반하여, 전자 장치(101)는 상기 하나 이상의 기능들의 실행을 스케줄링할 수 있다. 전자 장치(101)는 발언 처리기(360)의 실행에 기반하여, 상기 그룹 내 하나 이상의 보이스 커맨드들을, 제1 보이스 엔진(370)에 입력하거나, 또는 전자 장치(101)의 통신 회로(330)를 통해 외부 전자 장치(210)으로 송신할 수 있다.

[79] 전자 장치(101)는 발언 처리기(360)의 실행에 기반하여, 상기 하나 이상의 보이스 커맨드들에 대응하는 하나 이상의 기능들을 실행한 결과들을 포함하는 정보(예, 도 2의 제1 정보(235), 및/또는 제2 정보(245))를, 제1 보이스 엔진(370), 및/또는 제2 보이스 엔진(380)으로부터 획득할 수 있다. 전자 장치(101)는 발언 처리기(360)를 이용하여 획득된 상기 정보를, 스피커(310), 및/또는 디스플레이(260)를 통해 출력할 수 있다. 예를 들어, 전자 장치(101)는 상기 정보를 표현하는 오디오 신호를, 스피커(310)로 송신할 수 있다. 예를 들어, 전자 장치(101)는 디스플레이(260)를 제어하여, 상기 정보를 포함하는 시각적 객체, 및/또는 화면을 표시할 수

있다. 보이스 커맨드의 스케줄링, 및/또는 보이스 커맨드에 대응하는 기능을 실행한 결과들을 획득하기 위해 전자 장치(101)에 의해 실행되는 발언 처리기(360)의 예시적인 구조가, 도 4 내지 도 7을 참고하여 설명된다.

[80] 전자 장치(101)는 제1 보이스 엔진(370)의 실행에 기반하여, 발언 처리기(360)에 의해 제1 보이스 엔진(370)에 매칭된 적어도 하나의 보이스 커맨드를 식별할 수 있다. 전자 장치(101)는 상기 적어도 하나의 보이스 커맨드에 대응하는 적어도 하나의 기능을 실행할 수 있다. 상기 적어도 하나의 기능은, 전자 장치(101)에 포함된 하드웨어를 제어하기 위한 기능(예, 스피커(310)의 볼륨을 조절하는 기능), 메모리(130) 내에 설치된 어플리케이션에 의해 지원되는 기능(예, 알람, 음악 재생), 및/또는 통신 회로(330)를 통해 네트워크로부터 정보를 검색하는 기능(예, 날씨 검색) 중 적어도 하나를 포함할 수 있다. 전자 장치(101)는 상기 적어도 하나의 보이스 커맨드를 실행한 결과를, 발언 처리기(360)를 이용하여 저장하거나, 또는 사용자에게 출력할 수 있다.

[81] 외부 전자 장치(210)의 프로세서(120)는, 제2 보이스 엔진(380)의 실행에 기반하여, 전자 장치(101)로부터 요청된 적어도 하나의 보이스 커맨드에 대응하는 적어도 하나의 기능을 실행할 수 있다. 전자 장치(101)는 발언 처리기(360)에 기반하여, 적어도 하나의 보이스 커맨드를 실행할 것을, 외부 전자 장치(210)로 요청할 수 있다. 제2 보이스 엔진(380)이 실행된 상태 내에서, 외부 전자 장치(210)의 프로세서(120)는 상기 요청에 포함된 상기 적어도 하나의 보이스 커맨드에 대응하는 적어도 하나의 기능을 실행할 수 있다. 외부 전자 장치(210)의 프로세서(120)는, 통신 회로(330)를 이용하여 전자 장치(101)로, 상기 적어도 하나의 기능을 실행한 결과를 송신할 수 있다. 전자 장치(101)로 송신된 상기 결과는, 실행되는 발언 처리기(360)를 실행하는 전자 장치(101)의 프로세서(120)에 의해, 사용자에게 출력될 수 있다.

[82] 상술한 바와 같이, 일 실시예에 따른, 전자 장치(101)는 발언 처리기(360)와 같은 어플리케이션에 기반하여, 복수의 보이스 엔진들(예, 제1 보이스 엔진(370), 및 제2 보이스 엔진(380))에 의해 실행될 적어도 하나의 기능의 실행을 스케줄링하거나, 또는 제어할 수 있다. 복수의 보이스 엔진들의 스케줄링에 기반하여, 전자 장치(101)는 상기 복수의 보이스 엔진들을 효율적으로 이용하면서, 상기 복수의 보이스 엔진들의 실행을 위해 점유되는, 전자 장치(101), 및/또는 외부 전자 장치(210)의 리소스를 줄일 수 있다.

[83] 일 실시예에 따른, 전자 장치(101)는, 사용자와 상호작용하여 보이스 커맨드들의 그룹을 등록하는 상태 내에서, 상기 그룹에 포함된 상기 보이스 커맨드들의 속성(attributes)을 식별할 수 있다. 상기 속성은, 상기 그룹 내에 포함된 보이스 커맨드들에 대응하는 기능들을 실행할 보이스 엔진의 스케줄링을 위해 이용될 수 있다. 이하에서는, 도 4를 참고하여, 일 실시예에 따른, 전자 장치(101)가 복수의 보이스 커맨드들의 그룹을 저장하는 동작의 일 예가 설명된다.

[84] 도 4는, 일 실시예에 따른, 전자 장치(101)에 의해 실행되는 프로그램의 블록도이다.

[85] 도 4의 전자 장치(101)는 도 3의 전자 장치(101)의 일 예일 수 있다. 예를 들어, 도 3의 전자 장치(101), 메모리(130), 및 프로세서(120)에 의해 실행되는 입력 수신기(340), 퀵 커맨드 처리기(350), 발언 처리기(360), 및 제1 보이스 엔진(370)은, 도 4의 전자 장치(101), 메모리(130), 입력 수신기(340), 퀵 커맨드 처리기(350), 발언 처리기(360), 및 제1 보이스 엔진(370)을 포함할 수 있다.

[86] 도 4를 참고하면, 사용자의 발언(예, 오디오 신호 내에 포함된 발언)과 관련된 기능을 실행하기 위한 제1 보이스 엔진(370)의 인스트럭션들, 및/또는 서브-루틴들은, 발언 인식기(450), TTS(text-to-speech) 변환기(455), 자연어 처리기(460), 종속성(dependency) 관리기(470), 및/또는 이벤트 핸들러(475)로 구분될 수 있다. 자연어 처리기(460)의 인스트럭션들, 및/또는 서브-루틴들은, 도메인 분류기(462), 및/또는 목표(goal) 분류기(464)로 구분될 수 있다.

[87] 도 4를 참고하면, 사용자의 발언(예, 오디오 신호 내에 포함된 발언)으로부터 하나 이상의 보이스 커맨드들을 식별하기 위한 발언 처리기(360)의 인스트럭션들, 및/또는 서브-루틴들은, 발언 타입 분류기(410), 보이스 커맨드 인터프리터(420), 및/또는 보이스 커맨드 컨트롤러(440)로 구분될 수 있다. 발언 처리기(360)의 실행에 기반하여, 전자 장치(101)는 메모리(130) 내 보이스 커맨드 데이터베이스(430)를 관리할(manage) 수 있다. 보이스 커맨드 데이터베이스(430) 내에 저장된 정보의 일 예가 도 5를 참고하여 설명된다. 보이스 커맨드 인터프리터(420) 내 인스트럭션들, 및/또는 서브-루틴들은, 메타 데이터 분석기(422), 및/또는 보이스 커맨드 생성기(424)로 구분될 수 있다. 보이스 커맨드 컨트롤러(440) 내 인스트럭션들, 및/또는 서브-루틴들은, 보이스 커맨드 매니저(442), 보이스 엔진 스케줄러(446), 발언 스토리지(448)로 구분될 수 있다.

[88] 발언 처리기(360)의 실행에 기반하여, 전자 장치(101)는 발언에 포함된 사용자의 의도에 매칭되는 하나 이상의 보이스 커맨드들의 실행을 제어할 수 있다. 사용자로부터 복수의 보이스 커맨드들의 그룹을 등록하기 위한 입력을 식별한 상태 내에서, 전자 장치(101)는 발언 타입 분류기(410)의 실행에 기반하여, 상기 그룹 내에 포함된 상기 복수의 보이스 커맨드들 각각에 대한 속성들을 획득할 수 있다. 전자 장치(101)는 상기 속성들에 기반하여, 보이스 커맨드의 메타 데이터를 생성할 수 있다. 상기 메타 데이터는, 전자 장치(101)의 메모리(130) 내에 저장될 수 있다. 표 1은, 보이스 커맨드의 메타 데이터 내에 포함된 속성들의 일 예를 포함할 수 있다.

[89] [표1]

속성의 명칭	설명
Prompt	보이스 커맨드에 대응하는 기능을 실행하기 위하여, 퀵 커맨드와 상이한 파라미터가 요구되는지 여부
Time	보이스 커맨드에 대응하는 기능의 시간 종속성
Location	보이스 커맨드에 대응하는 기능의 위치 종속성
Device context	보이스 커맨드에 대응하는 기능의 단말 상태 종속성
onDevice supported	보이스 커맨드에 대응하는 기능이 단말에 의해 실행되는 보이스 엔진에 의해 실행 가능한지 여부
Event list	보이스 커맨드에 대응하여 실행되어야 하는 적어도 하나의 기능의 리스트

[90] 표 1의 명칭은, 속성에 할당된 예시적인 명칭일 수 있다. Prompt 명칭을 가지는 속성은, 부울(boolean) 자료형에 기반하여, 보이스 커맨드의 기능을 실행하기 위하여 추가적인 파라미터가 요구되는지 여부를 나타낼 수 있다. Time 명칭을 가지는 속성은, 상기 기능을 실행한 결과가, 시간에 따라 달라지는지 여부를 나타낼 수 있다. Location 명칭을 가지는 속성은, 상기 기능을 실행한 결과가, 전자 장치(101)의 위치에 따라 달라지는지 여부를 나타낼 수 있다. Device context 명칭을 가지는 속성은, 부울 자료형에 기반하여, 상기 기능을 실행한 결과가, 전자 장치(101)의 상태에 따라 달라지는지 여부를 나타낼 수 있다. onDevice supported 명칭을 가지는 속성은, 부울 자료형에 기반하여, 상기 기능이 제1 보이스 엔진(370)에 의해 실행 가능한지 여부를 나타낼 수 있다. 메타 데이터 내에 포함된 속성들(예, Prompt, Time, Location, Device context, onDevice supported 명칭들을 가지는 속성들)은, 종속성 관리기(470)에 의해 관리되는 정보(예, 종속성 맵(dependency map))를 포함할 수 있다.

[91] 표 1을 참고하면, Event list 명칭을 가지는 속성은, 보이스 커맨드에 대응하는 하나 이상의 기능들의 리스트를 포함할 수 있다. 예를 들어, Event list 명칭을 가지는 속성 내에, 상기 하나 이상의 기능들의 명칭들이 저장될 수 있다. 예를 들어, Event list 명칭을 가지는 속성 내에 'message' 텍스트가 저장되는 것은, 메타 데이터에 대응하는 보이스 커맨드가, 메시지를 표시하는 기능의 실행을 요구함을 나타낼 수 있다. 예를 들어, Event list 명칭을 가지는 속성 내에 'TTS_Response' 텍스트가 저장되는 것은, 메타 데이터에 대응하는 보이스 커맨드가, TTS 기능의 실행을 요구함을 나타낼 수 있다. 예를 들어, Event list 명칭을 가지는 속성 내에 'appLaunch' 텍스트가 저장되는 것은, 메타 데이터에 대응하는 보이스 커맨드가, 전자 장치(101)에 설치된 적어도 하나의 어플리케이션의 실행을 요구함을 나타낼 수 있다.

[92] 전자 장치(101)가 보이스 커맨드에 대한 메타 데이터를 획득하는 것은, 발언 타입 분류기(410)와 함께 실행되는 제1 보이스 엔진(370)에 기반하여 수행될 수 있다. 제1 보이스 엔진(370) 내 자연어 처리기(460)의 실행에 기반하여, 전자 장치(101)는 상기 보이스 커맨드에 포함된 자연어에 대한 전처리를 수행할 수 있다. 예를 들어, 전자 장치(101)는, 도메인 분류기(462)의 실행에 기반하여, 상기 보이스 커맨드가 제1 보이스 엔진(370)에 의해 처리될 수 있는지 여부를 식별할 수 있다. 보이스 커맨드가 제1 보이스 엔진(370)에 의해 처리될 수 있는지 여부가, 표 1의 onDevice supported 명칭을 가지는 속성 내에 저장될 수 있다. 전자 장치(101)는 목표 분류기(464)의 실행에 기반하여, 상기 보이스 커맨드에 의해 실행될 기능(또는 액션)의 목표를 식별할 수 있다. 종속성 관리기(470), 및/또는 이벤트 핸들러(475)의 실행에 기반하여, 전자 장치(101)는 전처리된 보이스 커맨드에 대한 후처리를 수행할 수 있다. 종속성 관리기(470)의 실행에 기반하여, 전자 장치(101)는 보이스 커맨드의 처리에 이용되는 리소스의 디바이스 의존성을 식별할 수 있다. 상기 디바이스 의존성은, 예를 들어, 상기 보이스 커맨드에 대응하는 기능이 제1 보이스 엔진(370), 및/또는 전자 장치(101)에 의해 처리될 수 있는지 여부를 포함할 수 있다. 이벤트 핸들러(475)의 실행에 기반하여, 전자 장치(101)는 보이스 커맨드에 대응하는 기능들의 리스트를 획득할 수 있다. 전자 장치(101)가 획득한 상기 리스트가, 메타 데이터 내에서 Event list 명칭을 가지는 속성 내에 저장될 수 있다.

[93] 발언 타입 분류기(410)의 실행에 기반하여, 전자 장치(101)는 퀵 커맨드에 의해 그룹핑된 하나 이상의 보이스 커맨드들에 대응하는 메타 데이터를 획득할 수 있다. 획득된 메타 데이터는, 전자 장치(101)의 메모리(130) 내에 저장될 수 있다. 전자 장치(101)가 퀵 커맨드를 등록하는 것은, 메모리(130) 내에 상기 메타 데이터를 저장하는 동작을 포함할 수 있다. 전자 장치(101)가 퀵 커맨드를 등록하는 것은, 상기 하나 이상의 보이스 커맨드들을, 보이스 커맨드 데이터베이스(430) 내에 저장하는 동작을 포함할 수 있다.

[94] 전자 장치(101)는 발언 타입 분류기(410)의 실행에 기반하여 획득된, 상기 메타 데이터를, 보이스 커맨드 인터프리터(420)에 기반하여 처리할 수 있다. 보이스 커맨드 인터프리터(420)의 실행에 기반하여, 전자 장치(101)는 상기 메타 데이터로부터 보이스 커맨드 데이터베이스(430) 내에 저장될 정보를 식별할 수 있다. 메타 데이터 분석기(422)의 실행에 기반하여, 전자 장치(101)는 발언 타입 분류기(410)에 기반하여 메모리(130) 내에 저장된 메타 데이터에 대한 파싱을 수행할 수 있다. 보이스 커맨드 생성기(424)의 실행에 기반하여, 전자 장치(101)는 메타 데이터 분석기(422)에 의해 파싱된 정보를, 보이스 커맨드 데이터베이스(430) 내에 저장할 수 있다. 보이스 커맨드 데이터베이스(430)의 스키마(schema)의 일 예가 도 5를 참고하여 설명된다.

[95] 상술한 바와 같이, 일 실시예에 따른, 전자 장치(101)는 퀵 커맨드를 등록함을 나타내는 입력에 기반하여, 상기 퀵 커맨드에 매칭되는 하나 이상의 보이스 커맨

드들에 대응하는 메타 데이터를 획득할 수 있다. 전자 장치(101)는 하나 이상의 보이스 커맨드들에 대응하는 정보를, 보이스 커맨드 데이터베이스(430) 내에 저장할 수 있다. 퀵 커맨드를 등록한 이후, 상기 퀵 커맨드를 포함하는 발언(예, 도 2의 발언(220))을 식별하는 것에 기반하여, 전자 장치(101)는 상기 저장된 메타 데이터를 이용하여, 상기 퀵 커맨드에 의해 실행될 하나 이상의 보이스 커맨드들, 및 상기 하나 이상의 보이스 커맨드들의 속성들을 식별할 수 있다. 발언 타입 분류기(410)에 의해 메모리(130) 내에 저장된 메타 데이터, 및 보이스 커맨드 데이터베이스(430) 내에 저장된 정보에 기반하여, 전자 장치(101)는 상기 하나 이상의 보이스 커맨드들, 및 상기 속성들을, 보다 신속하게 식별할 수 있다. 전자 장치(101)가 상기 퀵 커맨드를 포함하는 상기 발언을 식별하는 것에 기반하여 수행하는 동작이, 도 6을 참고하여 설명된다.

[96] 이하에서는, 도 5를 참고하여, 퀵 커맨드를 등록함을 나타내는 입력에 응답하여, 전자 장치(101)가 보이스 커맨드 데이터베이스(430) 내에 저장하는 정보의 일 예가 설명된다.

[97] 도 5는, 일 실시예에 따른, 전자 장치 내에 포함된 보이스 커맨드 데이터베이스(430)의 블록도이다.

[98] 도 5의 전자 장치는, 도 3 및 도 4의 전자 장치(101)의 일 예일 수 있다. 예를 들어, 도 4의 보이스 커맨드 데이터베이스(430)는 도 5의 보이스 커맨드 데이터베이스(430)를 포함할 수 있다. 도 5의 보이스 커맨드 데이터베이스(430)는, 도 3 내지 도 4의 발언 처리기(360)의 실행에 기반하여, 관리될 수 있다.

[99] 일 실시예에 따른, 전자 장치는 보이스 커맨드 데이터베이스(430) 내에 포함된 정보를, 보이스 커맨드에 대응하는 파라미터들을 저장하기 위한 제1 테이블(510), 및 상기 보이스 커맨드에 의해 실행될 하나 이상의 이벤트들과 관련된 파라미터들을 저장하기 위한 제2 테이블(520)로 구분하여 저장할 수 있다. 전자 장치는, 하나 이상의 보이스 커맨드들을 포함하는 퀵 커맨드를 등록하는 상태 내에서, 상기 하나 이상의 보이스 커맨드들에 기반하여, 보이스 커맨드 데이터베이스(430)를 업데이트할 수 있다.

[100] 보이스 커맨드 데이터베이스(430)의 제1 테이블(510) 내에서, 전자 장치는 보이스 커맨드에 대응하는 파라미터들을 저장할 수 있다. 표 2는, 보이스 커맨드 데이터베이스(430)의 제1 테이블(510) 내에 저장하는 상기 파라미터들의 일 예를 포함할 수 있다.

[101] [표2]

파라미터(또는 필드)의 명칭	설명
ID	보이스 커맨드 데이터베이스(430)의 제1 테이블(510) 내 보이스 커맨드의 식별자
Command	보이스 커맨드에 대응하는 자연어
Engine type	보이스 커맨드에 대응하는 기능을 실행할 수 있는 보이스 엔진의 타입
Static event list	보이스 커맨드에 대응하는 기능(또는 이벤트)의 가변성
Event list	보이스 커맨드에 대응하는 하나 이상의 기능들을 실행한 결과들을 표시하기 위해 수행될 이벤트들의 리스트

[102] 표 2를 참고하면, ID 명칭을 가지는 파라미터는, 보이스 커맨드 데이터베이스(430)의 제1 테이블(510) 내 레코드들에 고유하게 할당되는(uniquely assigned) 식별자(예, 키 값)을 포함할 수 있다. Command 명칭을 가지는 파라미터는, 보이스 커맨드에 대응하는 자연어를 포함할 수 있다. Engine type 명칭을 가지는 파라미터는, 보이스 커맨드에 대응하는 하나 이상의 기능들을 실행할 수 있는 보이스 엔진을 식별하기 위한 값을 포함할 수 있다. 예를 들어, Engine type 명칭을 가지는 파라미터는, 상기 하나 이상의 기능들이, 전자 장치에 의해 실행되는 제1 보이스 엔진(예, 도 3의 제1 보이스 엔진(370)), 및 외부 전자 장치(예, 도 2 내지 도 3의 외부 전자 장치(210))에 의해 실행되는 제2 보이스 엔진(예, 도 3의 제2 보이스 엔진(380)) 각각에 의해 실행될 수 있는지 여부를 나타내는 값을 포함할 수 있다. Static event list 명칭을 가지는 파라미터는, 보이스 커맨드에 대응하는 기능이 변경되는지 여부를 나타낼 수 있다. Event list 명칭을 가지는 파라미터는, 보이스 커맨드에 대응하는 하나 이상의 기능들을 실행한 결과들이 출력되는 방식을 나타낼 수 있다. 제1 테이블(510) 내에서, Event list 명칭을 가지는 파라미터는, 제2 테이블(520)의 레코드에 고유하게 할당된 식별자를 포함할 수 있다. 전자 장치는, Event list 명칭을 가지는 파라미터에 기반하여, 제1 테이블(510)의 특정 레코드에 매칭된 보이스 커맨드의 하나 이상의 기능들을 실행한 결과들이 출력되는 방식이 저장된, 제2 테이블(520) 내 레코드를 식별할 수 있다. 예를 들어, Event list 명칭을 가지는 상기 파라미터는, 제1 테이블(510), 및 제2 테이블(520)의 결합(예, join 연산에 기반하는 테이블들의 결합)을 위해 이용될 수 있다.

[103] 제2 테이블(520) 내에서, 전자 장치는 event message(이벤트 메시지) 명칭을 가지는 파라미터(또는 필드)에 기반하여, 보이스 커맨드의 하나 이상의 기능들을 실행한 결과들이 출력되는 방식을 나타낼 수 있다. 상기 파라미터는, JSON(javascript object notation)과 같이 구조화된 텍스트에 기반하여, 상기 방식을

나타낼 수 있다. 실시예가 이에 제한되는 것은 아니며, event message 명칭을 가지는 파라미터는, xml(extended marked-up language)에 기반하여, 상기 결과들이 출력되는 방식을 나타낼 수 있다. Event list 명칭을 가지는 파라미터 내, 상기 하나 이상의 기능들을 실행한 결과가, 디스플레이를 통해 표시될지, 또는 자연어에 기반하는 발언의 형태로 표시될지 여부가 저장될 수 있다.

[104] 도 4 및 도 5를 참고하여 상술한 바와 같이, 일 실시예에 따른, 전자 장치는 복수의 기능들의 그룹을 실행하기 위한 퀵 커맨드를 등록함을 나타내는 입력을 처리할 수 있다. 상기 퀵 커맨드는, 예를 들어, "굿모닝"과 같은 단순화된(simplified) 자연어에 기반하여, 사용자에 의해 배열된 기능들(예, 음악을 재생하는 기능, 알람을 해제하는 기능, 및 기상 예보를 출력하는 기능)의 순차적인 실행을 야기할 수 있다. 전자 장치는 상기 기능들을 실행하기 위해 이용되는 정보를, 전자 장치의 메모리 내에, 상기 기능들 각각에 대응하는 보이스 커맨드들의 단위로 저장할 수 있다. 전자 장치에 의해 저장되는 상기 정보는, 상기 보이스 커맨드들의 스케줄링에 이용되는 파라미터(예, 보이스 커맨드가 실행될 수 있는 전자 장치, 및/또는 보이스 엔진의 타입)를 포함할 수 있다.

[105] 퀵 커맨드, 및 상기 퀵 커맨드에 대응하는 하나 이상의 보이스 커맨드들을 등록한 이후, 전자 장치는 상기 퀵 커맨드를 포함하는 발언에 대응하여, 등록된 하나 이상의 보이스 커맨드들에 대응하는 하나 이상의 기능들을 실행할 수 있다. 전자 장치가, 도 4의 보이스 커맨드 컨트롤러(440)의 실행에 기반하여, 상기 하나 이상의 기능들에 대한 스케줄링, 및/또는 실행을 제어할 수 있다. 이하에서는 도 6 및 도 7을 참고하여, 일 실시예에 따른, 전자 장치가, 마이크로부터 출력된 오디오 신호로부터 퀵 커맨드를 식별하는 것에 기반하여 수행하는 동작의 일 예가 설명된다.

[106] 도 6은, 일 실시예에 따른, 전자 장치(101)에 의해 실행되는 프로그램을 설명하기 위한 블록도이다.

[107] 도 6의 전자 장치(101)는 도 3 및 도 4의 전자 장치(101)의 일 예일 수 있다. 예를 들어, 도 3의 전자 장치(101), 메모리(130), 및 프로세서(120)에 의해 실행되는 입력 수신기(340), 퀵 커맨드 처리기(350), 발언 처리기(360), 및 제1 보이스 엔진(370)은, 도 6의 전자 장치(101), 메모리(130), 입력 수신기(340), 퀵 커맨드 처리기(350), 발언 처리기(360), 및 제1 보이스 엔진(370)을 포함할 수 있다.

[108] 도 6을 참고하면, 사용자의 발언이 퀵 커맨드를 포함하는지 여부를 식별하기 위한 퀵 커맨드 처리기(350)의 인스트럭션들, 및/또는 서브-루틴들은, 비교기(610), 실행기(620), 퀵 커맨드 생성기(630)로 구분될 수 있다. 퀵 커맨드 처리기(350)의 실행에 기반하여, 전자 장치(101)는 메모리(130) 내에 퀵 커맨드에 대한 정보를 포함하는 퀵 커맨드 데이터베이스(640)를 관리할 수 있다. 사용자로부터 퀵 커맨드를 등록하는 상태 내에서, 전자 장치(101)는 상기 퀵 커맨드에 대응하는 자연어, 언어 타입(예, 영어), 상기 퀵 커맨드에 고유하게 할당된 식별자, 상기 퀵 커맨드가 등록된 시간을, 퀵 커맨드 데이터베이스(640) 내에 저장할 수 있다. 퀵 커맨

드 데이터베이스(640) 내에 저장되는 상기 정보는, 전자 장치(101)에 연결된 외부 전자 장치 내에 저장될 수 있다(예, 도 2 내지 도 3의 외부 전자 장치(210)의 메모리(130)).

[109] 일 실시예에 따른, 전자 장치(101)는 마이크(예, 도 3의 마이크(320))로부터 출력된 오디오 신호로부터, 사용자의 발언을 식별할 수 있다. 전자 장치(101)가 발언을 식별하는 것은, 제1 보이스 엔진(370), 및/또는 외부 전자 장치의 제2 보이스 엔진(예, 도 2의 제2 보이스 엔진(380)) 내 발언 인식기(450)의 실행에 기반하여, 상기 오디오 신호에 대응하는 텍스트를 획득하는 동작을 포함할 수 있다. 전자 장치(101)는 비교기(610)의 실행에 기반하여, 상기 텍스트, 및 퀵 커맨드 데이터베이스(640) 내에 저장된 퀵 커맨드를 비교하여, 퀵 커맨드가 상기 발언 내에 포함되어 있는지 여부를 식별할 수 있다.

[110] 퀵 커맨드가 상기 발언 내에 포함되어 있지 않은 경우, 전자 장치(101)는 발언 처리기(360)의 실행에 기반하여, 상기 발언 내에 포함된 하나 이상의 보이스 커맨드들을 식별할 수 있다. 전자 장치(101)는 발언 타입 분류기(410)에 의해 메모리(130)로부터, 상기 하나 이상의 보이스 커맨드들에 대응하는 메타 데이터를 식별할 수 있다. 메모리(130) 내에 상기 하나 이상의 보이스 커맨드들에 대응하는 메타 데이터가 존재하는 경우, 전자 장치(101)는 상기 메타 데이터에 기반하여, 상기 하나 이상의 보이스 커맨드들의 속성들을 식별할 수 있다. 식별된 속성들에 기반하여, 전자 장치(101)는 상기 하나 이상의 보이스 커맨드들의 실행에 이용될 보이스 엔진을, 제1 보이스 엔진(370), 또는 제2 보이스 엔진(예, 도 3의 제2 보이스 엔진(380)) 중에서 선택할 수 있다.

[111] 퀵 커맨드가 상기 발언 내에 포함되어 있지 않고, 상기 발언으로부터 복수의 보이스 커맨드들을 식별한 경우, 전자 장치(101)는 상기 복수의 보이스 커맨드들을, 제1 보이스 엔진(370), 또는 제2 보이스 엔진(예, 도 3의 제2 보이스 엔진(380)) 중 적어도 하나를 이용하여 처리할 수 있다. 예를 들어, 전자 장치(101)는 상기 제1 보이스 엔진(370), 및 상기 제2 보이스 엔진 중에서, 유휴 상태(idle state)의 보이스 엔진을 선택할 수 있다. 상기 예시 내에서, 전자 장치(101)는 선택된 보이스 엔진을 이용하여 상기 복수의 보이스 커맨드들에 대응하는 복수의 기능들을 순차적으로 실행할 수 있다. 유휴 상태의 보이스 엔진을 선택하기 위하여, 전자 장치(101)는 상기 제2 보이스 엔진이 실행되는 외부 전자 장치(예, 도 2 내지 도 3의 외부 전자 장치(210))와 통신 링크를 수립하거나, 또는 수립되어 있는 통신 링크를 유지할 수 있다.

[112] 상기 발언 내에 포함된 퀵 커맨드를 식별하는 것에 기반하여, 전자 장치(101)는 실행기(620)에 기반하여 상기 퀵 커맨드와 관련된 동작을 수행할 수 있다. 예를 들어, 전자 장치(101)는 실행기(620)에 기반하여, 발언 처리기(360)의 실행을 제어할 수 있다. 상기 발언 내에 포함된 퀵 커맨드를 식별한 상태 내에서, 전자 장치(101)는 발언 처리기(360)를 실행하여, 상기 퀵 커맨드와 관련된 메타 데이터, 및 보이스 커맨드 데이터베이스(430) 내에 저장된 파라미터들을 식별할 수 있다. 상

기 메타 데이터, 및 상기 파라미터들은, 상기 퀵 커맨드에 매칭되는 적어도 하나의 보이스 커맨드를 나타낼 수 있다. 상기 퀵 커맨드가 복수의 보이스 커맨드들의 실행을 위해 전자 장치(101) 내에 등록된 경우, 사용자가 상기 복수의 보이스 커맨드들 각각에 대응하는 문장들과 독립적으로, 상대적으로 짧은 길이의 상기 퀵 커맨드를 발언하여, 상기 복수의 보이스 커맨드들에 대응하는 기능들을 실행할 수 있다. 사용자가 자주 실행하는 보이스 커맨드들을 퀵 커맨드로 그룹핑한 경우, 전자 장치(101)는 상기 보이스 커맨드들에 대응하는 문장들을 인식하는 대신, 상대적으로 짧은 길이의 상기 퀵 커맨드를 인식하여, 상기 복수의 보이스 커맨드들을 실행할 수 있다.

[113] 퀵 커맨드와 관련된 메타 데이터, 및 보이스 커맨드 데이터베이스(430) 내 정보를 식별한 상태 내에서, 전자 장치(101)는 보이스 커맨드 컨트롤러(440)의 실행에 기반하여, 상기 퀵 커맨드에 대응하는 적어도 하나의 보이스 커맨드의 리스트를 획득할 수 있다. 전자 장치(101)는 상기 리스트에 기반하여, 상기 적어도 하나의 보이스 커맨드를 처리할 보이스 엔진을 스케줄링할 수 있다. 보이스 커맨드 컨트롤러(440)가 실행된 상태 내에서, 전자 장치(101)는 제1 보이스 엔진(370), 및 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 정보를 획득할 수 있다. 상기 제2 보이스 엔진의 정보는, 상기 제2 보이스 엔진을 실행하는 외부 전자 장치가 보이스 커맨드를 처리하는 것에 의해 발생되는 대기 시간(latency time), 및/또는 지연을 포함할 수 있다. 제2 보이스 엔진의 정보는, 상기 제2 보이스 엔진의 상태(예, 유휴 상태인지 여부)를 포함할 수 있다. 전자 장치(101)는 제1 보이스 엔진(370), 및 상기 제2 보이스 엔진의 상태들에 의해 지시되는, 상기 제1 보이스 엔진(370), 또는 상기 제2 보이스 엔진 각각이 보이스 커맨드에 대응하는 적어도 하나의 기능을 실행 가능한지 여부에 기반하여, 상기 퀵 커맨드에 대응하는 상기 적어도 하나의 보이스 커맨드를 분류할 수 있다. 예를 들어, 상기 리스트 내에서 복수의 보이스 커맨드들을 식별한 경우, 전자 장치(101)는 제1 보이스 엔진(370), 및 상기 제2 보이스 엔진의 상태들에 기반하여, 상기 복수의 보이스 커맨드들을, 제1 보이스 엔진(370)에 의해 실행될 제1 그룹, 및 상기 제2 보이스 엔진에 의해 실행될 제2 그룹으로 분류할 수 있다.

[114] 상술한 바와 같이, 일 실시예에 따른, 전자 장치(101)는 사용자의 발언으로부터 퀵 커맨드를 식별하는 것에 기반하여, 상기 퀵 커맨드에 대응하는 복수의 보이스 커맨드들을 식별할 수 있다. 전자 장치(101)는, 퀵 커맨드를 등록하는 동안 메모리(130) 내에 저장된 정보(예, 메타 데이터, 및/또는 보이스 커맨드 데이터베이스(430))에 기반하여, 상기 복수의 보이스 커맨드들의 스케줄링에 이용되는 상기 정보를 획득할 수 있다. 전자 장치(101)는 상기 정보에 기반하여 보이스 커맨드 컨트롤러(440)를 실행하여, 상기 복수의 보이스 커맨드들 각각을 처리할 보이스 엔진을, 상이한 보이스 엔진들(예, 제1 보이스 엔진(370), 및 상기 제2 보이스 엔진) 중에서 선택할 수 있다. 이하에서는, 도 7을 참고하여, 보이스 커맨드 컨트롤

러(440)가 실행된 상태 내에서, 전자 장치(101)가 상기 복수의 보이스 커맨드들 각각을 처리할 보이스 엔진을 선택하는 동작의 일 예가 설명된다.

[115] 도 7은, 일 실시예에 따른, 전자 장치(101)에 의해 실행되는 프로그램의 블록도이다.

[116] 도 7의 전자 장치(101)는 도 3 내지 도 6의 전자 장치(101)의 일 예일 수 있다. 예를 들어, 도 3 내지 도 6의 제1 보이스 엔진(370)은 도 7의 제1 보이스 엔진(370)을 포함할 수 있다. 도 4의 보이스 커맨드 컨트롤러(440), 및 보이스 커맨드 데이터베이스(430)는, 도 7의 보이스 커맨드 컨트롤러(440), 및 보이스 커맨드 데이터베이스(430)를 포함할 수 있다. 도 3의 외부 전자 장치(210)는, 도 7의 외부 전자 장치(210)를 포함할 수 있다.

[117] 도 7을 참고하면, 전자 장치(101)에 의해 실행되는 보이스 커맨드 컨트롤러(440)의 인스트럭션들, 및/또는 서브-루틴들은, 보이스 커맨드 매니저(442), 보이스 엔진 스케줄러(446), 및/또는 발언 스토리지(448)로 구분될 수 있다. 발언으로부터 전자 장치(101)에 등록된 퀵 커맨드를 식별하는 것에 기반하여, 전자 장치(101)는 보이스 커맨드 컨트롤러(440)를 실행할 수 있다. 전자 장치(101)는 보이스 커맨드 컨트롤러(440)의 실행을 위해 요구되는 정보를, 보이스 커맨드 데이터베이스(430)로부터 획득할 수 있다. 보이스 커맨드 데이터베이스(430)로부터, 전자 장치(101)는 상기 퀵 커맨드에 대응하는 적어도 하나의 보이스 커맨드와 관련된 정보를 획득할 수 있다.

[118] 보이스 커맨드 매니저(442)의 실행에 기반하여, 전자 장치(101)는 퀵 커맨드에 대응하는 적어도 하나의 보이스 커맨드와 관련된 정보를 획득할 수 있다. 전자 장치(101)는 상기 정보에 기반하여, 보이스 엔진에 기반하여 상기 적어도 하나의 보이스 커맨드를 처리할지 여부를 결정할 수 있다. 예를 들어, 전자 장치(101)의 메모리(예, 도 3의 메모리(130)) 내에, 보이스 커맨드에 대응하는 적어도 하나의 기능을 실행한 결과가 캐시된(cached) 경우, 전자 장치(101)는 보이스 엔진에 기반하는 보이스 커맨드의 처리를 바이패스하고, 발언 스토리지(448) 내에 상기 캐시된 결과를 저장할 수 있다. 발언 스토리지(448) 내에 상기 결과를 저장할지 여부는, 전자 장치(101)에 의해 식별된, 상기 결과가 유효한지 여부에 종속될 수 있다. 보이스 커맨드에 대응하는 적어도 하나의 기능을 실행한 결과가 메모리 내에 캐시되어 있지 않거나, 캐시된 결과가 유효하지 않은 경우, 전자 장치(101)는 보이스 엔진에 기반하여 상기 보이스 커맨드를 처리하기로 결정할 수 있다. 퀵 커맨드에 복수의 보이스 커맨드들이 대응하는 경우, 전자 장치(101)는 보이스 커맨드 매니저(442)의 실행에 기반하여, 상기 복수의 보이스 커맨드들 중에서 보이스 엔진에 의해 처리되어야 하는 보이스 커맨드를 식별할 수 있다.

[119] 보이스 엔진 스케줄러(446)의 실행에 기반하여, 전자 장치(101)는 보이스 엔진에 의해 처리되어야 하는 하나 이상의 보이스 커맨드들에 대한 스케줄링을 수행할 수 있다. 상기 스케줄링은, 전자 장치(101)에 의해 실행되는 제1 보이스 엔진(370), 및 외부 전자 장치(210)에 의해 실행되는 제2 보이스 엔진(380) 중에서, 상

기 하나 이상의 보이스 커맨드들의 처리에 이용될 보이스 엔진을 선택하는 동작을 포함할 수 있다. 복수의 보이스 커맨드들에 대한 스케줄링을 수행하는 상태 내에서, 전자 장치(101)는 복수의 보이스 커맨드들을, 제1 보이스 엔진(370)에 대응하는 제1 큐(710), 및 제2 보이스 엔진(380)에 대응하는 제2 큐(720) 중 어느 하나의 큐에 할당(allocate)할 수 있다. 제1 큐(710)는, 전자 장치(101)의 메모리(예, 도 3의 메모리(130)) 내에 형성된 영역으로, 제1 보이스 엔진(370)으로 순차적으로 입력될 하나 이상의 보이스 커맨드들에 대한 정보가 저장될 수 있다. 제2 큐(720)는, 상기 메모리 내에 형성된 다른 영역으로, 외부 전자 장치(210)로 순차적으로 송신될 하나 이상의 보이스 커맨드들에 대한 정보가 저장될 수 있다. 전자 장치(101)가 보이스 엔진 스케줄러(446)의 실행에 기반하여, FIFO(first in first out)의 데이터 구조를 가지는 제1 큐(710), 및 제2 큐(720)를 형성한 일 실시예가 도시되지만, 실시예가 이에 제한되는 것은 아니다.

[120] 보이스 엔진 스케줄러(446)가 실행된 상태 내에서, 전자 장치(101)는 제1 보이스 엔진(370), 및/또는 제2 보이스 엔진(380)의 상태들, 보이스 커맨드의 속성들 중 적어도 하나에 기반하여, 상기 보이스 커맨드를, 제1 큐(710), 또는 제2 큐(720) 중 어느 하나에 저장할 수 있다. 예를 들어, 상기 속성들 중에서, 표 1의 onDevice supported 명칭을 가지는 속성에 기반하여, 전자 장치(101)는 상기 보이스 커맨드가 특정 보이스 엔진(예, 제2 보이스 엔진(380))에 의하여 배타적으로 처리되어야 하는지 여부를 획득할 수 있다. 제2 보이스 엔진(380)에 의하여 배타적으로 처리되어야 하는 속성을 가진 보이스 커맨드를 식별하는 것에 응답하여, 전자 장치(101)는 상기 보이스 커맨드를 제1 큐(710), 또는 제2 큐(720) 중 제2 큐(720) 내에 저장할 수 있다. 예를 들어, 제1 보이스 엔진(370), 또는 제2 보이스 엔진(380) 각각이 유휴 상태인지 여부에 기반하여, 전자 장치(101)는 보이스 커맨드를 분류할 수 있다. 예를 들어, 제2 보이스 엔진(380)이 유휴 상태와 상이한 다른 상태인 경우, 전자 장치(101)는 보이스 엔진에 의해 처리되어야 하는 적어도 하나의 보이스 커맨드를, 제1 큐(710) 내에 저장할 수 있다.

[121] 보이스 엔진 스케줄러(446)의 실행에 기반하여, 전자 장치(101)는 비용 함수(cost function)를 이용하여, 퀵 커맨드에 포함된 복수의 보이스 커맨드들이 처리되는 총 기간(total duration)을 계산할 수 있다. 상기 총 기간을 최소화하도록, 전자 장치(101)는 복수의 보이스 커맨드들을, 제1 큐(710), 또는 제2 큐(720) 중 어느 하나의 큐로 분류할 수 있다. 상기 비용 함수는, 제1 보이스 엔진(370), 및 제2 보이스 엔진(380) 각각의 상태들, 및/또는 성능에 종속될 수 있다.

[122] 퀵 커맨드에 대응하는 복수의 보이스 커맨드들을, 보이스 엔진 스케줄러(446)의 실행에 기반하여, 상이한 보이스 엔진들로 분류한 결과를 이용하여, 전자 장치(101)는 상기 보이스 엔진들을 직접적으로, 또는 간접적으로 제어할 수 있다. 전자 장치(101)는 제1 큐(710)에 누적된 적어도 하나의 보이스 커맨드에 기반하여, 제1 보이스 엔진(370)을 실행할 수 있다. 전자 장치(101)는 제2 큐(720)에 누적된 적어도 하나의 보이스 커맨드를 포함하는 적어도 하나의 신호를, 제2 보이스

엔진(380)을 실행하는 외부 전자 장치(210)로 송신할 수 있다. 상기 적어도 하나의 신호에 기반하여, 전자 장치(101)는 외부 전자 장치(210)에 의해 실행되는 제2 보이스 엔진(380)을 간접적으로 제어할 수 있다.

[123] 전자 장치(101)는 발언 스토리지(448)의 실행에 기반하여, 제1 큐(710), 및 제2 큐(720) 각각에 저장된 보이스 커맨드들을 실행한 결과를, 전자 장치(101)의 메모리 내에 저장할 수 있다. 발언 스토리지(448)의 실행에 기반하여, 전자 장치(101)는 제1 보이스 엔진(370)을 이용하여 제1 큐(710) 내에 누적된 적어도 하나의 보이스 커맨드를 처리한 결과를, 메모리 내에 저장할 수 있다. 발언 스토리지(448)의 실행에 기반하여, 전자 장치(101)는 외부 전자 장치(210)로부터 수신된 신호에 포함되고, 제2 큐(720) 내에 누적된 적어도 하나의 보이스 커맨드를 처리한 결과를, 상기 메모리 내에 저장할 수 있다. 전자 장치(101)는 퀵 커맨드를 등록하는 동안 설정된, 복수의 보이스 커맨드들을 처리하는 순서에 기반하여, 상이한 보이스 엔진들에 의해 상기 보이스 커맨드들을 처리한 결과들을 저장할 수 있다. 예를 들어, 발언 스토리지(448)의 실행에 기반하여 전자 장치(101)의 메모리 내에 누적된 상기 결과들은, 상기 복수의 보이스 커맨드들을 처리하는 상기 순서에 기반하여, 정렬될 수 있다. 발언 스토리지(448)의 실행에 기반하여, 전자 장치(101)는 메모리 내에 복수의 보이스 커맨드들을 처리한 결과들이 저장되는 큐를 형성할 수 있다.

[124] 발언 스토리지(448)에 기반하여 전자 장치(101)의 메모리 내에 누적되고, 퀵 커맨드에 대응하는 복수의 보이스 커맨드들을 처리한 결과들로부터, 전자 장치(101)는 상기 결과들 각각에 대응하는 오디오 신호들을 획득할 수 있다. 복수의 보이스 커맨드들에 대한 순서에 기반하여, 전자 장치(101)는 스피커로 상기 오디오 신호들을 순차적으로 송신할 수 있다. 상기 오디오 신호들의 순차적인 송신에 기반하여, 전자 장치(101)는 상기 결과들을 표현하는 소리를, 상기 스피커를 통해 출력할 수 있다. 상기 오디오 신호들은, 상기 결과들을 표현하는 발언(예, 자연어 문장)을 포함할 수 있다.

[125] 상술한 바와 같이, 일 실시예에 따른, 전자 장치(101)는 하나의 퀵 커맨드에 대응하는 복수의 보이스 커맨드들을, 제1 보이스 엔진(370), 및 제2 보이스 엔진(380) 각각을 이용하여 분할하여 처리할 수 있다. 상기 복수의 보이스 커맨드들이 분할하여 처리되기 때문에, 제1 보이스 엔진(370), 및 제2 보이스 엔진(380) 각각의 대기 시간(idle time)이 최소화될 수 있다.

[126] 이하에서는, 도 8a 및 도 8b를 참고하여, 일 실시예에 따른, 전자 장치(101)가 퀵 커맨드를 등록하는 동작, 및 상기 퀵 커맨드를 포함하는 발언에 응답하여 상기 퀵 커맨드에 대응하는 복수의 보이스 커맨드들을 처리하는 동작의 일 예가 설명된다.

[127] 도 8a 및 도 8b는, 일 실시예에 따른, 전자 장치(101)가 UI(user interface)를 표시하는 예시적인 상태들(801, 802)을 도시한다.

[128] 도 8a 및 도 8b의 전자 장치(101)는 도 3 내지 도 7의 전자 장치(101)의 일 예일 수 있다. 예를 들어, 도 3의 전자 장치(101), 및 디스플레이(260)는, 도 8a 및 도 8b의 전자 장치(101), 및 디스플레이(260)를 포함할 수 있다. 도 8a 및 도 8b를 참고하면, 전자 장치(101)가 오디오 신호로부터 발언을 인식하기 위한 어플리케이션으로부터 제공된 UI를 표시하는 상이한 상태들(801, 802)이 도시된다. 상기 어플리케이션은, 도 3의 입력 수신기(340), 퀵 커맨드 처리기(350), 발언 처리기(360), 및/또는 제1 보이스 엔진(370)의 실행을 위한 하나 이상의 인스트럭션들을 포함할 수 있다.

[129] 도 8a를 참고하면, 상태(801) 내에서, 전자 장치(101)는 디스플레이(260) 내에 퀵 커맨드의 등록을 위한 화면을 표시할 수 있다. 상태(801) 내에서, 전자 장치(101)는 디스플레이(260)를 통하여, 퀵 커맨드의 등록을 위해 요구되는 정보를, 사용자로부터 수신할 수 있다. 전자 장치(101)는 디스플레이(260) 내에, 퀵 커맨드의 식별에 이용될 자연어를 획득하기 위한 텍스트 박스(810)를 표시할 수 있다. 텍스트 박스(810)와 함께, 전자 장치(101)는 발언에 기반하여 상기 자연어를 획득하기 위한 버튼(815)을 표시할 수 있다. 버튼(815)을 선택함을 나타내는 입력에 응답하여, 전자 장치(101)는 마이크로부터 이용하여 획득된 오디오 신호로부터 상기 자연어를 획득할 수 있다.

[130] 도 8a의 상태(801) 내에서, 전자 장치(101)는 디스플레이(260) 내에 퀵 커맨드에 의해 실행되는 하나 이상의 보이스 커맨드들을 추가하기 위한 버튼(860)을 표시할 수 있다. 버튼(860)을 선택함을 나타내는 입력에 응답하여, 전자 장치(101)는 텍스트, 및/또는 오디오 신호를 통해, 보이스 커맨드를 획득할 수 있다. 도 8a를 참고하면, 퀵 커맨드에 기반하여 4 개의 보이스 커맨드들을 획득한 예시적인 상태(801) 내에서, 전자 장치(101)는 디스플레이(260) 내에 상기 4 개의 보이스 커맨드들 각각에 대응하는 시각적 객체들(820, 830, 840, 850)을 표시할 수 있다. 시각적 객체들(820, 830, 840, 850)이 디스플레이(260) 내에서 배열된 순서는, 상기 퀵 커맨드에 기반하여 상기 제4 개의 보이스 커맨드들을 실행하는 순서를 나타낼 수 있다.

[131] 도 8a의 예시적인 상태(801) 내에서, 전자 장치(101)는 퀵 커맨드에 의해 실행될 보이스 커맨드들을, 시각적 객체들(820, 830, 840, 850)을 이용하여 시각화할 수 있다. 시각적 객체(820)는, 알람과 관련된 기능을 실행하기 위한 보이스 커맨드에 대응할 수 있다. 시각적 객체(820) 내에서, 전자 장치(101)는 상기 보이스 커맨드를 표현하는 자연어 문장(예, "알람 맞춰줘")을 표시할 수 있다. 시각적 객체(820) 내에서, 전자 장치(101)는 시각적 객체(820)에 대응하는 보이스 커맨드가 실행될 순서를 조절하기 위한 버튼(822)을 표시할 수 있다. 시각적 객체(820) 내에서, 전자 장치(101)는 시각적 객체(820)에 대응하는 보이스 커맨드를, 상기 퀵 커맨드에 의해 실행될 보이스 커맨드들로부터 제외하기 위한 버튼(824)을 표시할 수 있다. 유사하게, 시각적 객체(830) 내에서, 전자 장치(101)는 시각적 객체(830)에 대응하는 보이스 커맨드를 표현하는 자연어 문장(예, "마지막으로 온 문자 읽어줘")

을 표시할 수 있다. 시각적 객체(840) 내에서, 전자 장치(101)는 시각적 객체(840)에 대응하는 보이스 커맨드를 표현하는 자연어 문장(예, "10분 뒤에 리마인더 등록해줘")을 표시할 수 있다. 시각적 객체(850) 내에서, 전자 장치(101)는 시각적 객체(850)에 대응하는 보이스 커맨드를 표현하는 자연어 문장(예, "날씨 알려줘")을 표시할 수 있다. 시각적 객체들(820, 830, 840, 850), 및 텍스트 박스(810)를 참고하면, 퀵 커맨드를 위해 지정된 발언(예, "커맨드")은, 시각적 객체들(820, 830, 840, 850)의 실행을 위해 요구되는 상기 자연어 문장들과 다를 수 있다. 예를 들어, 상기 지정된 발언의 길이는, 상기 자연어 문장들 보다 짧을 수 있다.

[132] 도 8a를 참고하면, 전자 장치(101)는 디스플레이(260) 내에 표시된 정보에 기반하여 퀵 커맨드를 등록하기 위한 버튼(870)을 표시할 수 있다. 버튼(870)과 함께, 전자 장치(101)는 퀵 커맨드의 등록을 바이패스하기 위한 버튼(875)을 표시할 수 있다. 버튼(870)을 선택함을 나타내는 입력에 응답하여, 전자 장치(101)는 텍스트 박스(810)에 입력된 텍스트에 의해 실행되는 퀵 커맨드를 등록할 수 있다. 퀵 커맨드를 등록하기 위하여, 전자 장치(101)는 도 4 및 도 5를 참고하여 상술된 동작을 수행할 수 있다. 전자 장치(101)는 도 3 및 도 4의 발언 처리기(360)의 실행에 기반하여, 시각적 객체들(820, 830, 840, 850)에 대응하는 보이스 커맨드들에 대한 정보를 저장할 수 있다. 상기 정보는, 표 1을 참고하여 상술된, 상기 보이스 커맨드들에 대한 메타 데이터, 및/또는 표 2를 참고하여 상술된, 상기 보이스 커맨드들에 대한 속성들을 포함할 수 있다. 예를 들어, 전자 장치(101)는 표 3과 같이, 상기 보이스 커맨드들에 대한 속성들을 획득할 수 있다.

[133] [표3]

보이스 커맨드	Prompt	Time	Location	Device Context	OnDevice Supported	event list
알람 맞춰줘	O	X	X	X	O	Message, Renderer, TtsResponse
마지막으로 온 문자 읽어줘	X	X	X	O	O	Message, Renderer, TtsResponse
10분뒤에 리마인더 등록해줘	X	O	X	X	O	Message, Renderer, TtsResponse
날씨 알려줘	NA	NA	NA	NA	X	Message, Renderer, TtsResponse

[134] 표 3의 속성들은, 시각적 객체들(820, 830, 840, 850)에 대응하는 보이스 커맨드들에 대한 메타 데이터 내에 저장될 수 있다. 표 3을 참고하면, 시각적 객체(820)에 대응하는 제1 보이스 커맨드는, 알람과 관련된 자연어 문장(예, "알람 맞춰줘")을 포함할 수 있다. 상기 자연어 문장이 알람을 조절하기 위한 파라미터(예, 시간)를 포함하지 않기 때문에, 제1 보이스 커맨드를 처리하기 위하여, 상기 파라미터의 추가적인 입력이 요구될 수 있다. 전자 장치(101)는 상기 제1 보이스 커맨드에 대응하는 메타 데이터 내에서, prompt 명칭을 가지는 속성 내에, 상기 제1 보이스 커맨드의 처리를 위한 추가적인 입력이 요구됨을 나타내는 지정된 값을 할당할(assign) 수 있다.

[135] 표 3을 참고하면, 시각적 객체(830)에 대응하는 제2 보이스 커맨드는, SMS(short message service)와 관련된 자연어 문장(예, "마지막으로 온 문자 읽어줘")을 포함할 수 있다. 제2 보이스 커맨드는, 제2 보이스 커맨드를 처리하는 시점의 전자 장치(101)의 상태(예, 적어도 하나의 문자 메시지가 전자 장치(101) 내에 저장된 상태)에 종속될 수 있다. 전자 장치(101)는 제2 보이스 커맨드에 대응하는 메타 데이터 내에서, device context 명칭을 가지는 속성 내에, 상기 제2 보이스 커맨드의 처리를 위하여 전자 장치(101)의 상태를 식별하는 것이 요구됨을 나타내는 지정된 값을 할당할 수 있다.

[136] 표 3을 참고하면, 시각적 객체(840)에 대응하는 제3 보이스 커맨드는, 알람과 관련된 자연어 문장(예, "10분 뒤에 리마인더 등록해줘")을 포함할 수 있다. 제3 보이스 커맨드에 의해 실행되는 기능은, 제3 보이스 커맨드를 처리하는 시점에 종

속될 수 있다. 전자 장치(101)는 제3 보이스 커맨드에 대응하는 메타 데이터 내에서, time 명칭을 가지는 속성에 지정된 값을 할당하여, 상기 제3 보이스 커맨드의 처리를 위하여 전자 장치(101)의 현재 시점이 요구됨을 나타낼 수 있다.

[137] 표 3을 참고하면, 시각적 객체(850)에 대응하는 제4 보이스 커맨드는, 네트워크를 통한 정보의 획득을 위한 자연어 문장(예, "날씨 알려줘")을 포함할 수 있다. 제4 보이스 커맨드에 의해 실행되는 기능은, 전자 장치(101)와 상이한 외부 전자 장치(예, 날씨 정보를 제공하기 위한 서드-파티 서버)와 통신하는 것을 요구할 수 있다. 전자 장치(101)는 제4 보이스 커맨드에 대응하는 메타 데이터 내에서, "onDevice supported" 명칭을 가지는 속성에 지정된 값을 할당하여, 상기 제4 보이스 커맨드의 처리를 위하여 외부 전자 장치가 요구됨을 나타낼 수 있다.

[138] 표 3을 참고하여 상술한 바와 같이, "커맨드" 발언에 의해 식별되는 퀵 커맨드를 등록하는 동안, 전자 장치(101)는 사용자에 의해 상기 퀵 커맨드에 매칭된 4 개의 보이스 커맨드들의 속성들을 획득할 수 있다. 도 8a의 상태(801) 이후, 마이크로부터 출력된 오디오 신호로부터 "커맨드" 발언을 식별하는 것에 기반하여, 전자 장치(101)는 상기 4 개의 보이스 커맨드들을 실행하기 위하여 등록된, 상기 퀵 커맨드를 식별할 수 있다. 도 8b를 참고하면, 도 8a의 상태(801) 이후, 전자 장치(101)가 상기 퀵 커맨드를 실행하기 위한 상기 발언("커맨드")을 식별한 상태(802)가 도시된다.

[139] 도 8b의 상태(802) 내에서, 전자 장치(101)는 표 3을 참고하여 상술된, 퀵 커맨드에 대응하는 4 개의 보이스 커맨드들에 대한 메타 데이터를 획득할 수 있다. 상기 메타 데이터에 기반하여, 전자 장치(101)는 상기 보이스 커맨드들 각각을, 상이한 보이스 엔진들 중 어느 하나와 매칭할 수 있다. 상기 보이스 엔진들은, 전자 장치(101)에 의해 실행되는 제1 보이스 엔진(예, 도 3의 제1 보이스 엔진(370)), 및 전자 장치(101)와 상이한 외부 전자 장치에 의해 실행되는 제2 보이스 엔진(예, 도 3의 제2 보이스 엔진(380))을 포함할 수 있다. 전자 장치(101)는 상기 보이스 엔진들에 기반하여 상기 보이스 커맨드들이 처리되는 총 기간을 최소화하도록, 상기 보이스 엔진들로 상기 보이스 커맨드들을 할당할 수 있다. 전자 장치(101)가 상기 보이스 엔진들로 상기 보이스 커맨드들을 할당한 이후, 상기 보이스 엔진들에 기반하여 상기 보이스 커맨드들을 직접적으로, 또는 간접적으로 처리하는 것은, 도 6 내지 도 7을 참고하여 상술된 동작에 기반하여 수행될 수 있다.

[140] 일 실시예에 따른, 전자 장치(101)는 상기 보이스 커맨드들을 처리한 결과들을 표현한 오디오 신호들을, 상기 보이스 커맨드들의 순서에 따라 순차적으로 출력할 수 있다. 예를 들어, 전자 장치(101)는 도 8a의 상태(801) 내 시각적 객체들(820, 830, 840, 850)의 순서에 기반하여, 상기 오디오 신호들을 순차적으로 출력할 수 있다.

[141] 일 실시예에 따른, 전자 장치(101)는 디스플레이(260) 내에 보이스 커맨드들을 처리한 결과들을 시각화할 수 있다. 전자 장치(101)는 팝업-윈도우의 형태를 가지는 시각적 객체(880)를 통해, 상기 결과들 각각에 대응하는 시각적 객체들(882,

884, 886, 888)을 병렬로(in parallel) 표시할 수 있다. 전자 장치(101)가 시각적 객체들(882, 884, 886, 888)을 표시하는 순서 및/또는 레이아웃은, 도 8b의 일 실시예에 제한되지 않는다. 도 8b를 참고하면, 시각적 객체(882) 내에서, 전자 장치(101)는, 제1 보이스 커맨드를 실행한 결과를 포함하는 정보를 표시할 수 있다. 시각적 객체(884) 내에서, 전자 장치(101)는 제2 보이스 커맨드를 실행한 결과를 포함하는 정보를 표시할 수 있다. 시각적 객체(886) 내에서, 전자 장치(101)는 제3 보이스 커맨드를 실행한 결과를 포함하는 정보를 표시할 수 있다. 시각적 객체(888) 내에서, 전자 장치(101)는 제4 보이스 커맨드를 실행한 결과를 포함하는 정보를 표시할 수 있다. 전자 장치(101)는 상이한 보이스 엔진들을 이용하여 상기 보이스 커맨드들을 병렬로 실행한 결과들을, 시각적 객체들(882, 884, 886, 888)을 이용하여 개별적으로 표시할 수 있다. 예를 들어, 제2 보이스 커맨드의 처리가 완료되기 이전에, 전자 장치(101)는 시각적 객체(886)를 이용하여 제3 보이스 커맨드를 처리한 결과를 표시할 수 있다.

[142] 상술한 바와 같이, 일 실시예에 따른, 전자 장치(101)는 퀵 커맨드를 포함하는 발언에 응답하여, 상기 퀵 커맨드에 매칭된 복수의 보이스 커맨드들을, 상이한 보이스 엔진들을 이용하여 신속하게 실행할 수 있다. 전자 장치(101)는 스피커를 이용하여 상기 복수의 보이스 커맨드들을 처리한 결과들을 표현한 오디오 신호들을 출력할 수 있다. 전자 장치(101)는 디스플레이(260) 내에, 상기 결과들을 병렬로 표시할 수 있다. 실시예가 이에 제한되는 것은 아니며, 예를 들어, 전자 장치(101)는 디스플레이(260) 내에, 상기 결과들을 순차적으로 표시할 수 있다. 전자 장치(101)가 디스플레이(260)를 통하여 상기 결과들을 병렬로 표시하기 때문에, 전자 장치(101)는 상기 복수의 보이스 커맨드들 중 특정 보이스 커맨드의 처리가 실패하는 것과 독립적으로, 다른 보이스 커맨드를 처리한 결과를 사용자에게 표시할 수 있다.

[143] 도 9는, 일 실시예에 따른, 전자 장치(101)가 발언(910)에 기반하여 복수의 기능들을 실행하는 동작을 도시한다.

[144] 도 9의 전자 장치(101)는 도 3 내지 도 7의 전자 장치(101)의 일 예일 수 있다. 예를 들어, 도 3의 전자 장치(101), 및 디스플레이(260)는 도 9의 전자 장치(101), 및 디스플레이(260)를 포함할 수 있다. 도 9의 외부 전자 장치(210)는 도 2, 및/또는 도 7의 외부 전자 장치(210)를 포함할 수 있다.

[145] 일 실시예에 따른, 전자 장치(101)는 복수의 보이스 커맨드들을 포함하는 발언(910)을 식별할 수 있다. 상기 발언(910)은, 전자 장치(101)에 포함된 마이크를 통해 출력된 오디오 신호로부터 식별될 수 있다. 도 9를 참고하면, 전자 장치(101)는 발언(910) 내에 포함된 자연어 문장 내 제1 부분(912)(예, "7 시에 알람 맞춰주고")으로부터, 알람을 예약하기 위한 기능에 대응하는 제1 보이스 커맨드(920)를 식별할 수 있다. 전자 장치(101)는 발언(910) 내에 포함된 자연어 문장 내 제2 부분(914)(예, "뉴스 들려줘")로부터, 네트워크로부터 수집된 정보를 출력하기 위한 기능에 대응하는 제2 보이스 커맨드(930)를 식별할 수 있다. 전자 장치(101)는 도

4의 발언 처리기(360)의 실행에 기반하여, 발언(910)으로부터 상기 보이스 커맨드들(920, 930)을 식별할 수 있다.

[146] 발언(910)으로부터 복수의 보이스 커맨드들(예, 제1 보이스 커맨드(920), 및 제2 보이스 커맨드(930))을 식별하는 것에 기반하여, 전자 장치(101)는 복수의 보이스 커맨드들을 상이한 전자 장치들에 의해 실행되는 보이스 엔진들 중 적어도 하나로 입력할 수 있다. 상기 보이스 엔진들은, 전자 장치(101)에 의해 실행되는 제1 보이스 엔진(예, 도 3의 제1 보이스 엔진(370)), 및 전자 장치(101)에 연결된 외부 전자 장치(210)에 의해 실행되는 제2 보이스 엔진(예, 도 3의 제2 보이스 엔진(380))을 포함할 수 있다.

[147] 도 9를 참고하면, 전자 장치(101)는 제1 보이스 엔진을 이용하여 제1 보이스 커맨드(920)를 처리하여, 제1 보이스 커맨드(920)를 처리한 결과를 포함하는 제1 정보(925)를 획득할 수 있다. 전자 장치(101)는 제2 보이스 엔진을 이용하여 제2 보이스 커맨드(930)를 처리하기 위하여, 외부 전자 장치(210)로, 제2 보이스 커맨드(930)를 포함하는 신호를 송신할 수 있다. 전자 장치(101)는 외부 전자 장치(210)로부터 상기 신호에 대한 응답으로, 제2 정보(935)를 포함하는 다른 신호를 수신할 수 있다. 상기 다른 신호에 포함된 상기 제2 정보(935)는, 제2 보이스 엔진을 실행한 외부 전자 장치(210)에 의해 제2 보이스 커맨드(930)를 처리한 결과를 포함할 수 있다.

[148] 일 실시예에 따른, 전자 장치(101)는 발언(910)으로부터 식별된 복수의 보이스 커맨드들(예, 제1 보이스 커맨드(920), 및 제2 보이스 커맨드(930))을 처리한 결과들을 포함하는 정보(예, 제1 정보(925), 및 제2 정보(935))를 포함하는 발언(940)을, 사용자에게 출력할 수 있다. 전자 장치(101)는 스피커로 송신되는 오디오 신호를 이용하여, 상기 발언(940)을 출력할 수 있다. 도 9를 참고하면, 발언(940)은, 제1 보이스 커맨드(920)를 처리한 결과인 제1 정보(925)에 기반하는 자연어 문장(942)(예, "7 시에 알람이 울릴 것입니다"), 및 제2 보이스 커맨드(930)를 처리한 결과인 제2 정보(935)에 기반하는 자연어 문장(944)(예, "뉴스입니다. 오늘 S 회사는, ...")을 포함할 수 있다.

[149] 상술한 바와 같이, 일 실시예에 따른, 전자 장치(101)는, 퀵 커맨드를 포함하는 발언(예, 도 2의 발언(220)), 및/또는 복수의 보이스 커맨드들을 표현한 발언(예, 도 9의 발언(910))에 기반하여, 복수의 보이스 커맨드들을 처리하기 위한 입력을 식별할 수 있다. 상기 입력에 기반하여, 전자 장치(101)는, 전자 장치(101)에 의해 실행되는 제1 보이스 엔진, 및 외부 전자 장치(210)에 의해 실행되는 제2 보이스 엔진의 상태들에 기반하여, 상기 복수의 보이스 커맨드들을 병렬로 처리할 수 있다. 퀵 커맨드에 응답하여 복수의 보이스 커맨드들을 병렬로 처리한 경우, 전자 장치(101)는 복수의 보이스 커맨드들을 병렬로 처리한 결과들을, 상기 퀵 커맨드에 의해 정렬된 복수의 보이스 커맨드들의 순서를 따라 순차적으로 표시할 수 있다.

[150] 도 10은, 일 실시예에 따른, 전자 장치에 의해 수행되는 동작의 흐름도를 도시한다.

[151] 도 10의 전자 장치는 도 1 내지 도 9의 전자 장치(101)를 포함할 수 있다. 도 10의 동작들 중 적어도 하나는, 도 3의 전자 장치(101), 및/또는 도 3의 전자 장치(101) 내 프로세서(120)에 의해 실행될 수 있다. 이하 실시예에서 각 동작들은 순차적으로 수행될 수도 있으나, 반드시 순차적으로 수행되는 것은 아니다. 예를 들어, 각 동작들의 순서가 변경될 수도 있으며, 적어도 두 동작들이 병렬적으로 수행될 수도 있다.

[152] 도 10을 참고하면, 동작(1010) 내에서, 일 실시예에 따른, 전자 장치는, 마이크를 통하여 제1 오디오 신호를 식별할 수 있다. 제1 오디오 신호는, 마이크를 포함하는 외부 공간의 진동을 나타낼 수 있다. 동작(1010)의 마이크는, 도 3의 마이크(320), 헤드폰과 같이, 전자 장치와 유선으로 연결된 마이크, 또는 무선 이어폰과 같이, 도 3의 통신 회로(330)를 통해 전자 장치와 무선으로 연결된 마이크 중 적어도 하나를 포함할 수 있다.

[153] 도 10을 참고하면, 동작(1020) 내에서, 일 실시예에 따른, 전자 장치는, 제1 오디오 신호로부터 복수의 기능들을 순차적으로 실행하기 위한 발언을 식별하였는지 여부를 판단할 수 있다. 상기 발언은, 도 2의 발언(220)과 같이, 퀵 커맨드를 실행하기 위하여 전자 장치에 등록된 자연어를 포함할 수 있다. 상기 발언은, 도 9의 발언(910)과 같이, 복수의 기능들을 표현한 자연어 문장들을 포함할 수 있다.

[154] 제1 오디오 신호로부터 복수의 기능들을 순차적으로 실행하기 위한 발언을 식별한 상태 내에서(1020-예), 동작(1030)에 기반하여, 일 실시예에 따른, 전자 장치는, 프로세서에 기반하여 복수의 기능들 중 적어도 하나의 제1 기능을 실행할 수 있다. 상기 상태 내에서, 동작(1040)에 기반하여, 전자 장치는, 통신 회로(예, 도 3의 전자 장치(101) 내 통신 회로(330))를 통해 연결된 외부 전자 장치(예, 도 3의 외부 전자 장치(210))를 이용하여, 복수의 기능들 중에서, 적어도 하나의 제1 기능과 상이한 적어도 하나의 제2 기능을 실행할 수 있다. 동작들(1030, 1040)을 수행하기 위하여, 전자 장치는 제1 오디오 신호로부터 식별된 상기 발언에 매칭된 상기 복수의 기능들을, 전자 장치의 프로세서에 의해 실행될 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능으로 구분할 수 있다.

[155] 전자 장치가 상기 복수의 기능들을, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능으로 구분하는 것은, 도 6 및 도 7을 참고하여 상술된 동작에 기반하여, 수행될 수 있다. 예를 들어, 전자 장치는 프로세서에 의해 실행되는 제1 보이스 엔진, 및 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 상태들에 기반하여, 상기 복수의 기능들을, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능으로 구분할 수 있다. 전자 장치는 동작(1030)에 기반하여 상기 복수의 기능들 중에서 구분된 상기 적어도 하나의 제1 기능을, 전자 장치의 프로세서에 의해 실행되는 제1 보이스 엔진을 이용하여 처리할 수 있다. 전자 장치는 동작(1040)에 기반하여, 상기 적어도 하나의 제2 기능을, 외부 전자 장치로 송신

하고, 외부 전자 장치로부터 상기 적어도 하나의 제2 기능을 실행한 결과를 획득할 수 있다.

[156] 도 10을 참고하면, 동작(1050) 내에서, 일 실시예에 따른, 전자 장치는, 스피커를 통하여, 적어도 하나의 제1 기능, 및 적어도 하나의 제2 기능을 실행한 결과들을 표현하는 제2 오디오 신호들을 출력할 수 있다. 전자 장치는, 상기 제1 오디오 신호를 통해 식별된 상기 발언에 대한 응답으로, 상기 제2 오디오 신호들을 출력할 수 있다. 상기 제2 오디오 신호들 각각은, 동작(1020)의 복수의 기능들 각각을 처리한 결과들을 표현한 자연어 문장을 포함할 수 있다. 상기 제2 오디오 신호들에 의해 전자 장치로부터 출력되는 자연어 문장들의 순서는, 상기 발언에 의해 지시되는 복수의 기능들의 순서에 대응할 수 있다.

[157] 제1 오디오 신호로부터 단일의 기능(예, 제3 기능)을 실행하기 위한 발언을 식별한 경우(1020-아니오), 동작(1060)에 기반하여, 전자 장치는, 전자 장치의 프로세서, 또는 외부 전자 장치 중 어느 하나를 이용하여, 제1 오디오 신호에 대응하는 제3 기능을 실행할 수 있다. 전자 장치는 프로세서에 의해 실행되는 상기 제1 보이스 엔진의 상태, 및/또는 상기 외부 전자 장치에 의해 실행되는 상기 제2 보이스 엔진의 상태를 식별할 수 있다. 예를 들어, 전자 장치는 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 중에서, 유휴 상태인 보이스 엔진을 이용하여 상기 제3 기능을 실행할 수 있다. 동작(1060) 내에서, 동작(1050)과 유사하게, 전자 장치는 상기 제3 기능을 실행한 결과를 표현하는 제3 오디오 신호를 출력할 수 있다.

[158] 동작들(1050, 1060)을 참고하여 상술한 바와 같이, 전자 장치는 제1 오디오 신호에 포함된 발언에 기반하여, 제2 오디오 신호들, 및/또는 제3 오디오 신호를 출력할 수 있다. 실시예가 이에 제한되는 것은 아니며, 전자 장치는 도 8b를 참고하여 상술한 바와 같이, 전자 장치 내 디스플레이를 이용하여, 동작(1020)에 의해 식별된 발언에 대응하는 하나 이상의 기능들을 실행한 결과들을 시각화할 수 있다.

[159] 도 11은, 일 실시예에 따른, 전자 장치에 의해 수행되는 동작의 흐름도를 도시한다.

[160] 도 11의 전자 장치는 도 1 내지 도 9의 전자 장치(101)를 포함할 수 있다. 도 11의 동작들 중 적어도 하나는, 도 3의 전자 장치(101), 및/또는 도 3의 전자 장치(101) 내 프로세서(120)에 의해 실행될 수 있다. 도 11의 동작들 중 적어도 하나는 도 10의 동작들 중 적어도 하나와 관련될 수 있다. 이하 실시예에서 각 동작들은 순차적으로 수행될 수도 있으나, 반드시 순차적으로 수행되는 것은 아니다. 예를 들어, 각 동작들의 순서가 변경될 수도 있으며, 적어도 두 동작들이 병렬적으로 수행될 수도 있다.

[161] 도 11을 참고하면, 동작(1110) 내에서, 일 실시예에 따른, 전자 장치는, 오디오 신호로부터 발언을 식별할 수 있다. 전자 장치는, 전자 장치의 프로세서에 의해 실행되는 제1 보이스 엔진을 이용하여 발언을 식별하거나, 또는 전자 장치에 연결된 외부 전자 장치에 의해 실행되는 제2 보이스 엔진을 이용하여 발언을 식별

할 수 있다. 전자 장치가 발언을 식별하는 것은, 상기 발언에 대응하는 텍스트 데이터를 획득하는 동작을 포함할 수 있다.

[162] 도 11을 참고하면, 동작(1120) 내에서, 일 실시예에 따른, 전자 장치는, 발언에 대응하는 메타 데이터를 식별하였는지 여부를 식별하였는지 여부를 판단할 수 있다. 상기 메타 데이터는, 전자 장치에 의해 실행되는 발언 처리기(예, 도 3의 발언 처리기(360))를 이용하여 식별될 수 있다.

[163] 발언에 대응하는 메타 데이터를 식별한 상태 내에서(1120-예), 동작(1130)에 기반하여, 전자 장치는, 동작(1120)에 의해 식별된 메타 데이터에 기반하여, 발언을 처리할 보이스 엔진을 선택할 수 있다. 전자 장치는 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 중에서, 발언과 관련된 하나 이상의 기능들을 실행하기 위한 보이스 엔진을 선택할 수 있다. 발언으로부터 복수의 보이스 커맨드들을 식별한 경우, 전자 장치는 상기 복수의 보이스 커맨드들 각각에 대하여 동작(1130)을 독립적으로 수행할 수 있다. 예를 들어, 동작(1130)에 기반하여, 상기 복수의 보이스 커맨드들 각각이 처리될 보이스 엔진이, 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 중에서 독립적으로 선택될 수 있다. 전자 장치는 메타 데이터 내 속성들에 기반하여, 발언이 처리될 보이스 엔진을, 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 중에서 선택할 수 있다. 예를 들어, 메타 데이터를 통하여 상기 발언이 특정 보이스 엔진에 의해 선택적으로 실행됨을 나타내는 속성을 식별한 경우, 전자 장치는 상기 속성에 기반하여, 상기 발언을 처리할 보이스 엔진을 선택할 수 있다. 발언에 대응하는 메타 데이터를 식별하지 못한 경우(1120-아니오), 동작(1130)에 기반하는 보이스 엔진의 선택이 바이패스될 수 있다.

[164] 도 11을 참고하면, 동작(1140) 내에서, 일 실시예에 따른, 전자 장치는, 발언으로부터, 발언에 대응하는 기능을 실행하기 위해 요구되는 파라미터를 식별하였는지 여부를 판단할 수 있다. 발언으로부터, 상기 파라미터를 식별하지 못한 경우(1140-아니오), 동작(1160)에 기반하여, 전자 장치는 상기 기능을 실행하기 위해 요구되는 파라미터를, 사용자에게 요청할 수 있다. 발언으로부터 상기 파라미터를 식별하였거나, 또는 상기 기능이 어느 파라미터도 요구하지 않는 경우(1140-예), 동작(1150)에 기반하여, 전자 장치는 보이스 엔진을 이용하여 상기 발언을 처리할 수 있다. 전자 장치는, 스피커, 및/또는 디스플레이를 이용하여 상기 발언을 처리한 결과(예, 상기 발언에 포함된 하나 이상의 보이스 커맨드들을 실행한 결과들)를 출력할 수 있다.

[165] 도 12는, 일 실시예에 따른, 전자 장치에 의해 수행되는 동작의 흐름도를 도시한다.

[166] 도 12의 전자 장치는 도 1 내지 도 9의 전자 장치(101)를 포함할 수 있다. 도 12의 동작들 중 적어도 하나는, 도 3의 전자 장치(101), 및/또는 도 3의 전자 장치(101) 내 프로세서(120)에 의해 실행될 수 있다. 도 12의 동작들 중 적어도 하나는, 도 10 내지 도 11의 동작들 중 적어도 하나와 관련될 수 있다. 이하 실시예에서 각 동작들은 순차적으로 수행될 수도 있으나, 반드시 순차적으로 수행되는 것은 아니다.

예를 들어, 각 동작들의 순서가 변경될 수도 있으며, 적어도 두 동작들이 병렬적으로 수행될 수도 있다.

[167] 도 12를 참고하면, 동작(1210) 내에서, 일 실시예에 따른, 전자 장치는, 마이크를 통하여 수신된 제1 오디오 신호에 포함된 발언을 식별할 수 있다. 전자 장치는 도 11의 동작(1110)과 유사하게, 도 12의 동작(1210)을 수행할 수 있다.

[168] 도 12를 참고하면, 동작(1220) 내에서, 일 실시예에 따른, 전자 장치는, 발언으로부터 복수의 보이스 커맨드들을 식별하는 것에 기반하여, 프로세서에 의해 실행되는 제1 보이스 엔진, 및 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 상태들을 획득할 수 있다. 발언 내에서, 상기 복수의 보이스 커맨드들의 그룹의 실행을 위한 퀵 커맨드를 식별하는 것에 기반하여, 전자 장치는 동작(1220)을 수행할 수 있다. 전자 장치가 동작(1220)에 기반하여 획득하는 상기 상태들은, 상기 제1 보이스 엔진, 및/또는 상기 제2 보이스 엔진이 유휴 상태인지 여부를 나타낼 수 있다.

[169] 도 12를 참고하면, 동작(1230) 내에서, 일 실시예에 따른, 전자 장치는, 상태들에 기반하여, 복수의 보이스 커맨드들을, 제1 보이스 엔진에 대응하는 제1 보이스 커맨드, 및 제2 보이스 엔진에 대응하는 제2 보이스 커맨드로 구분할 수 있다. 전자 장치는, 상태들에 의해 나타나고, 제1 보이스 엔진 또는 제2 보이스 엔진 각각의 기능을 실행하는 가능성(possibility)에 기반하여, 복수의 기능들 각각을 적어도 하나의 제1 기능 및 적어도 하나의 제2 기능으로 구분할 수 있다. 전자 장치는, 상태들에 의해 나타나는, 제1 보이스 엔진 또는 제2 보이스 엔진 각각의 기능의 타당성(feasibility)에 기반하여, 복수의 기능들 각각을 적어도 하나의 제1 기능 및 적어도 하나의 제2 기능으로 구분할 수 있다. 동작(1230)에 기반하여, 전자 장치는 동작(1220)의 복수의 보이스 커맨드들에 대한 스케줄링을 수행할 수 있다. 전자 장치는 제1 보이스 커맨드, 및 제2 보이스 커맨드를, 전자 장치의 메모리 내에서 상이한 영역들(예, 도 7의 제1 큐(710), 및 제2 큐(720)) 내에 저장할 수 있다.

[170] 도 12를 참고하면, 동작(1240) 내에서, 일 실시예에 따른, 전자 장치는, 제1 보이스 커맨드에 대응하는 제1 기능을, 제1 보이스 엔진을 이용하여 실행할 수 있다. 제1 보이스 엔진이 전자 장치의 프로세서에 의해 실행되기 때문에, 전자 장치는 동작(1240)에 기반하여, 복수의 보이스 커맨드들 중에서, 제1 보이스 커맨드를 선택적으로 처리할 수 있다.

[171] 도 12를 참고하면, 동작(1250) 내에서, 일 실시예에 따른, 전자 장치는, 외부 전자 장치로 제2 보이스 커맨드에 대응하는 제2 기능의 실행을 요청할 수 있다. 전자 장치는, 제2 보이스 엔진을 실행하는 상기 외부 전자 장치로 상기 제2 보이스 커맨드와 관련된 정보를 송신하여, 상기 제2 기능의 실행을 요청할 수 있다. 전자 장치는 외부 전자 장치로부터, 상기 요청에 대한 응답으로, 상기 제2 기능을 실행한 결과를 포함하는 신호, 및/또는 정보를 수신할 수 있다. 전자 장치가 동작들(1240, 1250)을 실행하는 순서는, 도 12의 일 실시예에 제한되지 않는다. 예를 들어, 전자 장치는 동작들(1240, 1250)을 병렬적으로 수행할 수 있다.

[172] 도 12를 참고하면, 동작(1260) 내에서, 일 실시예에 따른, 전자 장치는, 발언에 의해 지시되는 복수의 보이스 커맨드들의 순서에 기반하여, 제1 기능, 및 제2 기능을 실행한 결과들을 표현한 제2 오디오 신호를 출력할 수 있다. 전자 장치가 제2 오디오 신호를 통해 출력하는 소리는, 상기 제1 기능, 및 상기 제2 기능을 실행한 결과들을 표현하는 자연어 문장들을 포함할 수 있다. 상기 자연어 문장들이 출력되는 순서는 큐 커맨드를 등록하는 상태 내에서 설정된, 상기 복수의 보이스 커맨드들이 처리되는 순서에 종속될 수 있다.

[173] 도 13은 일 실시 예에 따른 인공 지능(artificial intelligence, AI) 시스템을 나타낸 블록도이다.

[174] 도 13을 참조하면, 일 실시 예의 인공 지능 시스템(10)은 사용자 단말(1300), 지능형 서버(1400), 및 서비스 서버(1500)를 포함할 수 있다.

[175] 일 실시 예의 사용자 단말(1300)(예: 도 1의 전자 장치(101))은, 인터넷에 연결 가능한 단말 장치(또는, 전자 장치)일 수 있으며, 예를 들어, 휴대폰, 스마트폰, PDA(personal digital assistant), 노트북 컴퓨터, TV, 백색 가전, 웨어러블 장치, HMD, 또는 스마트 스피커일 수 있다.

[176] 일 실시 예에 따르면, 사용자 단말(1300)은 통신 인터페이스(1310), 마이크(1320), 스피커(1330), 디스플레이(1340), 메모리(1350), 및 프로세서(1360)를 포함할 수 있다. 상기 열거된 구성요소들은 서로 작동적으로 또는 전기적으로 연결될 수 있다.

[177] 일 실시 예에 따르면, 통신 인터페이스(1310)는 외부 장치와 연결되어 데이터를 송수신하도록 구성될 수 있다. 일 실시 예에 따르면, 마이크(1320)는 소리(예: 사용자 발화)를 수신하여, 전기적 신호로 변환할 수 있다. 일 실시 예에 따르면, 스피커(1330)는 전기적 신호를 소리(예: 음성)으로 출력할 수 있다. 일 실시 예에 따르면, 디스플레이(1340)는 이미지 또는 비디오를 표시하도록 구성될 수 있다. 일 실시 예에 따르면, 디스플레이(1340)는 실행되는 앱(app)(또는, 어플리케이션 프로그램(application program))의 그래픽 사용자 인터페이스(graphic user interface)(GUI)를 표시할 수 있다.

[178] 일 실시예의 디스플레이(1340)는 이미지 또는 비디오를 표시하도록 구성될 수 있다. 일 실시예의 디스플레이(1340)는 또한 실행되는 앱(app)(또는, 어플리케이션 프로그램(application program))의 그래픽 사용자 인터페이스(graphic user interface)(GUI)를 표시할 수 있다. 일 실시예의 디스플레이(1340)는 터치 센서를 통해 터치 입력을 수신할 수 있다. 예를 들어, 디스플레이(1340)는 디스플레이(1340) 내에 표시되는 화상 키보드 영역의 터치 센서를 통해 텍스트 입력을 수신할 수 있다.

[179] 일 실시 예에 따르면, 메모리(1350)는 클라이언트 모듈(1351), SDK(software development kit)(1353), 및 복수의 앱들(1355)을 저장할 수 있다. 상기 클라이언트 모듈(1351), 및 SDK(1353)는 범용적인 기능을 수행하기 위한 프레임워크(framework)(또는, 솔루션 프로그램(solution program))를 구성할 수 있다. 또한, 클

라이언트 모듈(1351) 또는 SDK(1353)는 사용자 입력(예: 음성 입력, 텍스트 입력, 터치 입력)을 처리하기 위한 프레임워크를 구성할 수 있다.

[180] 일 실시 예에 따르면, 상기 복수의 앱들(1355)은 지정된 기능을 수행하기 위한 프로그램일 수 있다. 일 실시 예에 따르면, 복수의 앱들(1355)은 제1 앱(1355_1), 제2 앱(1355_3) 을 포함할 수 있다. 일 실시 예에 따르면, 복수의 앱들(1355) 각각은 지정된 기능을 수행하기 위한 복수의 동작들을 포함할 수 있다. 예를 들어, 복수의 앱들(1355)은, 알람 앱, 메시지 앱, 및 스케줄 앱 중 적어도 하나를 포함할 수 있다. 일 실시 예에 따르면, 복수의 앱들(1355)은 프로세서(1360)에 의해 실행되어 상기 복수의 동작들 중 적어도 일부를 순차적으로 실행할 수 있다.

[181] 일 실시 예에 따르면, 프로세서(1360)는 사용자 단말(1300)의 전반적인 동작을 제어할 수 있다. 예를 들어, 프로세서(1360)는 통신 인터페이스(1310), 마이크(1320), 스피커(1330), 디스플레이(1340), 및 메모리(1350)와 전기적으로 연결되어 연결되어 지정된 동작을 수행할 수 있다.

[182] 일 실시 예에 따르면, 프로세서(1360)는 또한 상기 메모리(1350)에 저장된 프로그램을 실행시켜 지정된 기능을 수행할 수 있다. 예를 들어, 프로세서(1360)는 클라이언트 모듈(1351) 또는 SDK(1353) 중 적어도 하나를 실행하여, 사용자 입력을 처리하기 위한 이하의 동작을 수행할 수 있다. 프로세서(1360)는, 예를 들어, SDK(1353)를 통해 복수의 앱들(1355)의 동작을 제어할 수 있다. 클라이언트 모듈(1351) 또는 SDK(1353)의 동작으로 설명된 이하의 동작은 프로세서(1360)의 실행에 의한 동작일 수 있다.

[183] 일 실시 예에 따르면, 클라이언트 모듈(1351)은 사용자 입력을 수신할 수 있다. 예를 들어, 클라이언트 모듈(1351)은 마이크(1320)를 통해 감지된 사용자 발화에 대응되는 음성 신호를 생성할 수 있다. 또는, 클라이언트 모듈(1351)은 디스플레이(1340)를 통해 감지된 터치 입력을 수신할 수 있다. 또는, 클라이언트 모듈(1351)은 키보드 또는 화상 키보드를 통해 감지된 텍스트 입력을 수신할 수 있다. 이 외에도, 클라이언트 모듈(1351)은 사용자 단말(1300)에 포함된 입력 모듈 또는 사용자 단말(1300)에 연결된 입력 모듈을 통해 감지되는 다양한 형태의 사용자 입력을 수신할 수 있다. 상기 클라이언트 모듈(1351)은 수신된 사용자 입력을 지능형 서버(1400)로 송신할 수 있다. 일 실시 예에 따르면, 클라이언트 모듈(1351)은 수신된 사용자 입력과 함께, 사용자 단말(1300)의 상태 정보를 지능형 서버(1400)로 송신할 수 있다. 상기 상태 정보는, 예를 들어, 앱의 실행 상태 정보일 수 있다.

[184] 일 실시 예에 따르면, 클라이언트 모듈(1351)은 수신된 사용자 입력에 대응되는 결과를 수신할 수 있다. 예를 들어, 클라이언트 모듈(1351)은 지능형 서버(1400)로부터 사용자 입력에 대응되는 결과를 수신할 수 있다. 클라이언트 모듈(1351)은 상기 수신된 결과를 디스플레이(1340)에 표시할 수 있다. 또한, 클라이언트 모듈(1351)은 상기 수신된 결과를 스피커(1330)를 통해 오디오로 출력할 수 있다.

[185] 일 실시 예에 따르면, 클라이언트 모듈(1351)은 수신된 사용자 입력에 대응되는 플랜을 수신할 수 있다. 클라이언트 모듈(1351)은 플랜에 따라 앱의 복수의 동작들을 실행한 결과를 디스플레이(1340)에 표시할 수 있다. 예를 들어, 클라이언트 모듈(1351)은 복수의 동작들의 실행 결과를 순차적으로 디스플레이에 표시할 수 있고, 스피커(1330)를 통해 오디오를 출력할 수 있다. 다른 예를 들어, 사용자 단말(1300)은 복수의 동작들을 실행한 일부 결과(예: 마지막 동작의 결과)만을 디스플레이에 표시할 수 있으며, 스피커(1330)를 통해 오디오를 출력할 수 있다.

[186] 일 실시 예에 따르면, 클라이언트 모듈(1351)은 지능형 서버(1400)로부터 사용자 입력에 대응되는 결과를 산출하기 위해 필요한 정보를 획득하기 위한 요청을 수신할 수 있다. 상기 결과를 산출하기 위해 필요한 정보는, 예를 들어, 사용자 단말(1300)의 상태 정보일 수 있다. 일 실시 예에 따르면, 클라이언트 모듈(1351)은 상기 요청에 대응하여 상기 필요한 정보를 지능형 서버(1400)로 송신할 수 있다.

[187] 일 실시 예에 따르면, 클라이언트 모듈(1351)은 플랜에 따라 복수의 동작들을 실행한 결과 정보를 지능형 서버(1400)로 송신할 수 있다. 지능형 서버(1400)는 상기 결과 정보를 통해 수신된 사용자 입력이 올바르게 처리된 것을 확인할 수 있다.

[188] 일 실시 예에 따르면, 클라이언트 모듈(1351)은 음성 인식 모듈을 포함할 수 있다. 일 실시 예에 따르면, 클라이언트 모듈(1351)은 상기 음성 인식 모듈을 통해 제한된 기능을 수행하는 음성 입력을 인식할 수 있다. 예를 들어, 클라이언트 모듈(1351)은 지정된 입력(예: 웨이크 업!)을 통해 유기적인 동작을 수행하기 위한 음성 입력을 처리하기 위한 지능형 앱을 수행할 수 있다.

[189] 일 실시 예에 따르면, 지능형 서버(1400)는 통신 망을 통해 사용자 단말(1300)로부터 사용자 음성 입력과 관련된 정보를 수신할 수 있다. 일 실시 예에 따르면, 지능형 서버(1400)는 수신된 음성 입력과 관련된 데이터를 텍스트 데이터(text data)로 변경할 수 있다. 일 실시 예에 따르면, 지능형 서버(1400)는 상기 텍스트 데이터에 기초하여 사용자 음성 입력과 대응되는 태스크(task)를 수행하기 위한 플랜(plan)을 생성할 수 있다.

[190] 일 실시 예에 따르면, 플랜은 인공 지능(artificial intelligent)(AI) 시스템에 의해 생성될 수 있다. 인공 지능 시스템은 룰 베이스 시스템(rule-based system) 일 수도 있고, 신경망 베이스 시스템(neural network-based system)(예: 피드포워드 신경망(feedforward neural network)(FNN)), 순환 신경망(recurrent neural network)(RNN)) 일 수도 있다. 또는, 전술한 것의 조합 또는 이와 상이한 인공지능 시스템일 수도 있다. 일 실시 예에 따르면, 플랜은 미리 정의된 플랜의 집합에서 선택될 수 있거나, 사용자 요청에 응답하여 실시간으로 생성될 수 있다. 예를 들어, 인공지능 시스템은 미리 정의 된 복수의 플랜들 중 적어도 하나의 플랜을 선택할 수 있다.

[191] 일 실시 예에 따르면, 지능형 서버(1400)는 생성된 플랜에 따라 산출된 결과를 사용자 단말(1300)로 송신하거나, 생성된 플랜을 사용자 단말(1300)로 송신할 수 있다. 일 실시 예에 따르면, 사용자 단말(1300)은 플랜에 따라 산출된 결과를 디스

플레이에 표시할 수 있다. 일 실시 예에 따르면, 사용자 단말(1300)은 플랜에 따른 동작을 실행한 결과를 디스플레이에 표시할 수 있다.

[192] 일 실시 예의 지능형 서버(1400)는 프론트 엔드(front end)(1410), 자연어 플랫폼(natural language platform)(1420), 캡슐 데이터베이스(capsule DB)(1430), 실행 엔진(execution engine)(1440), 엔드 유저 인터페이스(end user interface)(1450), 매니지먼트 플랫폼(management platform)(1460), 빅 데이터 플랫폼(big data platform)(1470), 및 분석 플랫폼(analytic platform)(1480)을 포함할 수 있다.

[193] 일 실시 예에 따르면, 프론트 엔드(1410)는 사용자 단말(1300)로부터 수신된 사용자 입력을 수신할 수 있다. 프론트 엔드(1410)는 상기 사용자 입력에 대응되는 응답을 송신할 수 있다.

[194] 일 실시 예에 따르면, 자연어 플랫폼(1420)은 자동 음성 인식 모듈(automatic speech recognition module)(ASR module)(1421), 자연어 이해 모듈(natural language understanding module)(NLU module)(1423), 플래너 모듈(planner module)(1425), 자연어 생성 모듈(natural language generator module)(NLG module)(1427), 및 텍스트 음성 변환 모듈(text to speech module)(TTS module)(1429)를 포함할 수 있다.

[195] 일 실시 예에 따르면, 자동 음성 인식 모듈(1421)은 사용자 단말(1300)로부터 수신된 음성 입력을 텍스트 데이터로 변환할 수 있다. 일 실시 예에 따르면, 자연어 이해 모듈(1423)은 음성 입력의 텍스트 데이터를 이용하여 사용자의 의도를 파악할 수 있다. 예를 들어, 자연어 이해 모듈(1423)은 텍스트 데이터 형태의 사용자 입력에 대하여 문법적 분석(syntactic analyze) 또는 의미적 분석(semantic analyze)을 수행하여 사용자의 의도를 파악할 수 있다. 일 실시 예에 따르면, 자연어 이해 모듈(1423)은 형태소 또는 구의 언어적 특징(예: 문법적 요소)을 이용하여 사용자 입력으로부터 추출된 단어의 의미를 파악하고, 상기 파악된 단어의 의미를 의도에 매칭시켜 사용자의 의도를 결정할 수 있다. 자연어 이해 모듈(1423)은 사용자 발화에 대응되는 의도 정보(intent information)를 획득할 수 있다. 의도 정보는 텍스트 데이터를 해석하여 판단되는 사용자의 의도를 나타내는 정보일 수 있다. 의도 정보는 사용자가 디바이스를 이용하여 실행하고자 하는 동작 또는 기능을 나타내는 정보를 포함할 수 있다.

[196] 일 실시 예에 따르면, 플래너 모듈(1425)은 자연어 이해 모듈(1423)에서 결정된 의도 및 파라미터를 이용하여 플랜을 생성할 수 있다. 일 실시 예에 따르면, 플래너 모듈(1425)은 상기 결정된 의도에 기초하여 태스크를 수행하기 위해 필요한 복수의 도메인들을 결정할 수 있다. 플래너 모듈(1425)은 상기 의도에 기초하여 결정된 복수의 도메인들 각각에 포함된 복수의 동작들을 결정할 수 있다. 일 실시 예에 따르면, 플래너 모듈(1425)은 상기 결정된 복수의 동작들을 실행하는데 필요한 파라미터나, 상기 복수의 동작들의 실행에 의해 출력되는 결과 값을 결정할 수 있다. 상기 파라미터 및 상기 결과 값은 지정된 형식(또는, 클래스)과 관련된 컨셉으로 정의될 수 있다. 이에 따라, 플랜은 사용자의 의도에 의해 결정된 복수의 동작들, 및 복수의 컨셉들을 포함할 수 있다. 상기 플래너 모듈(1425)은 상기

복수의 동작들, 및 상기 복수의 컨셉들 사이의 관계를 단계적(또는, 계층적)으로 결정할 수 있다. 예를 들어, 플래너 모듈(1425)은 복수의 컨셉들에 기초하여 사용자의 의도에 기초하여 결정된 복수의 동작들의 실행 순서를 결정할 수 있다. 다시 말해, 플래너 모듈(1425)은 복수의 동작들의 실행에 필요한 파라미터, 및 복수의 동작들의 실행에 의해 출력되는 결과에 기초하여, 복수의 동작들의 실행 순서를 결정할 수 있다. 이에 따라, 플래너 모듈(1425)는 복수의 동작들, 및 복수의 컨셉들 사이의 연관 정보(예: 온톨로지(ontology))가 포함된 플랜를 생성할 수 있다. 상기 플래너 모듈(1425)은 컨셉과 동작의 관계들의 집합이 저장된 캡슐 데이터베이스(1430)에 저장된 정보를 이용하여 플랜을 생성할 수 있다.

[197] 일 실시 예에 따르면, 자연어 생성 모듈(1427)은 지정된 정보를 텍스트 형태로 변경할 수 있다. 상기 텍스트 형태로 변경된 정보는 자연어 발화의 형태일 수 있다. 일 실시 예의 텍스트 음성 변환 모듈(1429)은 텍스트 형태의 정보를 음성 형태의 정보로 변경할 수 있다.

[198] 일 실시 예에 따르면, 캡슐 데이터베이스(1430)는 복수의 도메인들에 대응되는 복수의 컨셉들과 동작들의 관계에 대한 정보를 저장할 수 있다. 예를 들어, 캡슐 데이터베이스(1430)는 플랜의 복수의 동작 오브젝트(action object 또는, 동작 정보)들 및 컨셉 오브젝트(concept object 또는 컨셉 정보)들을 포함하는 복수의 캡슐들을 저장할 수 있다. 일 실시 예에 따르면, 캡슐 데이터베이스(1430)는 CAN(concept action network)의 형태로 상기 복수의 캡슐들을 저장할 수 있다. 일 실시 예에 따르면, 복수의 캡슐들은 캡슐 데이터베이스(1430)에 포함된 기능 저장소(function registry)에 저장될 수 있다.

[199] 일 실시 예에 따르면, 캡슐 데이터베이스(1430)는 음성 입력에 대응되는 플랜을 결정할 때 필요한 전략 정보가 저장된 전략 레지스트리(strategy registry)를 포함할 수 있다. 상기 전략 정보는 사용자 입력에 대응되는 복수의 플랜들이 있는 경우, 하나의 플랜을 결정하기 위한 기준 정보를 포함할 수 있다. 일 실시 예에 따르면, 캡슐 데이터베이스(1430)는 지정된 상황에서 사용자에게 후속 동작을 제안하기 위한 후속 동작의 정보가 저장된 후속 동작 레지스트리(follow up registry)를 포함할 수 있다. 상기 후속 동작은, 예를 들어, 후속 발화를 포함할 수 있다. 일 실시 예에 따르면, 캡슐 데이터베이스(1430)는 사용자 단말(1300)을 통해 출력되는 정보의 레이아웃(layout) 정보를 저장하는 레이아웃 레지스트리(layout registry)를 포함할 수 있다. 일 실시 예에 따르면, 캡슐 데이터베이스(1430)는 캡슐 정보에 포함된 어휘(vocabulary) 정보가 저장된 어휘 레지스트리(vocabulary registry)를 포함할 수 있다. 일 실시 예에 따르면, 캡슐 데이터베이스(1430)는 사용자와의 대화(dialog)(또는, 인터렉션(interaction)) 정보가 저장된 대화 레지스트리(dialog registry)를 포함할 수 있다.

[200] 일 실시 예에 따르면, 캡슐 데이터베이스(1430)는 개발자 툴(developer tool)을 통해 저장된 오브젝트를 업데이트(update)할 수 있다. 상기 개발자 툴은, 예를 들어, 동작 오브젝트 또는 컨셉 오브젝트를 업데이트하기 위한 기능 에디터

(function editor)를 포함할 수 있다. 상기 개발자 툴은 어휘를 업데이트하기 위한 어휘 에디터(vocabulary editor)를 포함할 수 있다. 상기 개발자 툴은 플랜을 결정하는 전략을 생성 및 등록 하는 전략 에디터(strategy editor)를 포함할 수 있다. 상기 개발자 툴은 사용자와의 대화를 생성하는 대화 에디터(dialog editor)를 포함할 수 있다. 상기 개발자 툴은 후속 목표를 활성화하고, 힌트를 제공하는 후속 발화를 편집할 수 있는 후속 동작 에디터(follow up editor)를 포함할 수 있다. 상기 후속 목표는 현재 설정된 목표, 사용자의 선호도 또는 환경 조건에 기초하여 결정될 수 있다.

[201] 일 실시 예에 따르면, 캡슐 데이터베이스(1430)는 사용자 단말(1300) 내에도 구현될 수 있다. 다시 말해, 사용자 단말(1300)은 음성 입력에 대응되는 동작을 결정하기 위한 정보를 저장하는 캡슐 데이터베이스(1430)를 포함할 수 있다.

[202] 일 실시 예에 따르면, 실행 엔진(1440)은 상기 생성된 플랜을 이용하여 결과를 산출할 수 있다. 일 실시 예에 따르면, 엔드 유저 인터페이스(1450)는 산출된 결과를 사용자 단말(1300)로 송신할 수 있다. 이에 따라, 사용자 단말(1300)은 상기 결과를 수신하고, 상기 수신된 결과를 사용자에게 제공할 수 있다. 일 실시 예에 따르면, 매니지먼트 플랫폼(1460)은 지능형 서버(1400)에서 이용되는 정보를 관리할 수 있다. 일 실시 예에 따르면, 빅 데이터 플랫폼(1470)은 사용자의 데이터를 수집할 수 있다. 일 실시 예에 따르면, 분석 플랫폼(1480)을 지능형 서버(1400)의 QoS(quality of service)를 관리할 수 있다. 예를 들어, 분석 플랫폼(1480)은 지능형 서버(1400)의 구성 요소 및 처리 속도(또는, 효율성)를 관리할 수 있다.

[203] 일 실시 예에 따르면, 서비스 서버(1500)는 사용자 단말(1300)에 지정된 서비스(예: 음식 주문 또는 호텔 예약)를 제공할 수 있다. 일 실시 예에 따르면, 서비스 서버(1500)는 제3 자에 의해 운영되는 서버일 수 있다. 예를 들어, 서비스 서버(1500)는 서로 다른 제3 자에 의해 운영되는 제1 서비스 서버(1501), 제2 서비스 서버(1503), 및 제3 서비스 서버(1505)를 포함할 수 있다. 일 실시 예에 따르면, 서비스 서버(1500)는 수신된 음성 입력에 대응되는 플랜을 생성하기 위한 정보를 지능형 서버(1400)에 제공할 수 있다. 상기 제공된 정보는, 예를 들어, 캡슐 데이터베이스(1430)에 저장될 수 있다. 또한, 서비스 서버(1500)는 플랜에 따른 결과 정보를 지능형 서버(1400)에 제공할 수 있다.

[204] 상기에 기술된 인공 지능 시스템(13)에서, 상기 사용자 단말(1300)은, 사용자 입력에 응답하여 사용자에게 다양한 인텔리전트 서비스를 제공할 수 있다. 상기 사용자 입력은, 예를 들어, 물리적 버튼을 통한 입력, 터치 입력 또는 음성 입력을 포함할 수 있다.

[205] 일 실시 예에 따르면, 상기 사용자 단말(1300)은 내부에 저장된 지능형 앱(또는, 음성 인식 앱)을 통해 음성 인식 서비스를 제공할 수 있다. 이 경우, 예를 들어, 사용자 단말(1300)은 상기 마이크를 통해 수신된 사용자 발화(utterance) 또는 음성 입력(voice input)를 인식하고, 인식된 음성 입력에 대응되는 서비스를 사용자에게 제공할 수 있다.

[206] 일 실시 예에 따르면, 사용자 단말(1300)은 수신된 음성 입력에 기초하여, 단독으로 또는 상기 지능형 서버 및/또는 서비스 서버와 함께 지정된 동작을 수행할 수 있다. 예를 들어, 사용자 단말(1300)은 수신된 음성 입력에 대응되는 앱을 실행시키고, 실행된 앱을 통해 지정된 동작을 수행할 수 있다.

[207] 일 실시 예에 따르면, 사용자 단말(1300)이 지능형 서버(1400) 및/또는 서비스 서버와 함께 서비스를 제공하는 경우에는, 상기 사용자 단말은, 상기 마이크(1320)를 이용하여 사용자 발화를 감지하고, 상기 감지된 사용자 발화에 대응되는 신호(또는, 음성 데이터)를 생성할 수 있다. 상기 사용자 단말은, 상기 음성 데이터를 통신 인터페이스(1310)를 이용하여 지능형 서버(1400)로 송신할 수 있다.

[208] 일 실시 예에 따르면, 지능형 서버(1400)는 사용자 단말(1300)로부터 수신된 음성 입력에 대한 응답으로써, 음성 입력에 대응되는 태스크(task)를 수행하기 위한 플랜, 또는 상기 플랜에 따라 동작을 수행한 결과를 생성할 수 있다. 상기 플랜은, 예를 들어, 사용자의 음성 입력에 대응되는 태스크(task)를 수행하기 위한 복수의 동작들, 및 상기 복수의 동작들과 관련된 복수의 컨셉들을 포함할 수 있다. 상기 컨셉은 상기 복수의 동작들의 실행에 입력되는 파라미터나, 복수의 동작들의 실행에 의해 출력되는 결과 값을 정의한 것일 수 있다. 상기 플랜은 복수의 동작들, 및 복수의 컨셉들 사이의 연관 정보를 포함할 수 있다.

[209] 일 실시 예의 사용자 단말(1300)은, 통신 인터페이스(1310)를 이용하여 상기 응답을 수신할 수 있다. 사용자 단말(1300)은 상기 스피커(1330)를 이용하여 사용자 단말(1300) 내부에서 생성된 음성 신호를 외부로 출력하거나, 디스플레이(1340)를 이용하여 사용자 단말(1300) 내부에서 생성된 이미지를 외부로 출력할 수 있다.

[210] 도 14는 다양한 실시 예에 따른, 컨셉과 동작의 관계 정보를 데이터베이스하기 위한 스키마(schemar)를 도시하는 도면이다.

[211] 상기 지능형 서버(예: 도 13의 지능형 서버(1400))의 캡슐 데이터베이스(예: 도 13의 캡슐 데이터베이스(1430))는 CAN(concept action network)(1600) 형태로 복수의 캡슐들을 저장할 수 있다. 상기 캡슐 데이터베이스는 사용자의 음성 입력에 대응되는 태스크를 처리하기 위한 동작, 및 상기 동작을 위해 필요한 파라미터를 CAN(concept action network) 형태로 저장될 수 있다. 상기 CAN은 동작(action)과 상기 동작을 수행하는데 필요한 파라미터(parameter)를 정의한 컨셉(concept) 사이의 유기적인 관계를 나타낸 것일 수 있다.

[212] 상기 캡슐 데이터베이스는 복수의 도메인(예: 어플리케이션) 각각에 대응되는 복수의 캡슐(예: Capsule A(1601), Capsule B(1604))들을 저장할 수 있다. 일 실시 예에 따르면, 하나의 캡슐(예: Capsule A(1601))은 하나의 도메인(예: 어플리케이션)에 대응될 수 있다. 또한, 하나의 캡슐은 캡슐과 관련된 도메인의 기능을 수행하기 위한 적어도 하나의 서비스 제공자(예: CP 1(1602), CP 2 (1603), CP 3(1606), 또는 CP 4(1605))에 대응될 수 있다. 일 실시 예에 따르면, 하나의 캡슐은 지정된

기능을 수행하기 위한 적어도 하나 이상의 동작(1610) 및 적어도 하나 이상의 컨셉(1620)을 포함할 수 있다.

[213] 일 실시 예에 따르면, 자연어 플랫폼(예: 도 13의 자연어 플랫폼(1420))은 캡슐 데이터베이스에 저장된 캡슐을 이용하여 수신된 음성 입력에 대응하는 태스크를 수행하기 위한 플랜을 생성할 수 있다. 예를 들어, 자연어 플랫폼의 플래너 모듈(예: 도 13의 플래너 모듈(1425))은 캡슐 데이터베이스에 저장된 캡슐을 이용하여 플랜을 생성할 수 있다. 예를 들어, Capsule A (1601) 의 동작들(1711, 1713)과 컨셉들(1712, 1714) 및 Capsule B(1604)의 동작(1741)과 컨셉들(1742)을 이용하여 플랜(1607)을 생성할 수 있다.

[214] 도 15는 다양한 실시 예에 따른 사용자 단말이 지능형 앱을 통해 수신된 음성 입력을 처리하는 화면을 나타낸 도면이다.

[215] 사용자 단말(1300)은 지능형 서버(예: 도 13의 지능형 서버(1400))를 통해 사용자 입력을 처리하기 위해 지능형 앱을 실행할 수 있다.

[216] 일 실시 예에 따르면, 1510 화면에서, 사용자 단말(1300)은 지정된 음성 입력(예: 웨이크 업!)를 인식하거나 하드웨어 키(예: 전용 하드웨어 키)를 통한 입력을 수신하면, 음성 입력을 처리하기 위한 지능형 앱을 실행할 수 있다. 사용자 단말(1300)은, 예를 들어, 스케줄 앱을 실행한 상태에서 지능형 앱을 실행할 수 있다. 일 실시 예에 따르면, 사용자 단말(1300)은 지능형 앱에 대응되는 오브젝트(예: 아이콘)(1511)를 디스플레이(예: 도 13의 디스플레이(1340))에 표시할 수 있다. 일 실시 예에 따르면, 사용자 단말(1300)은 사용자 발화에 의한 음성 입력을 수신할 수 있다. 예를 들어, 사용자 단말(1300)은 "이번주 일정 알려줘!"라는 음성 입력을 수신할 수 있다. 일 실시 예에 따르면, 사용자 단말(1300)은 수신된 음성 입력의 텍스트 데이터가 표시된 지능형 앱의 UI(user interface)(1513)(예: 입력창)를 디스플레이에 표시할 수 있다.

[217] 일 실시 예에 따르면, 1520 화면에서, 사용자 단말(1300)은 수신된 음성 입력에 대응되는 결과를 디스플레이에 표시할 수 있다. 예를 들어, 사용자 단말(1300)은 수신된 사용자 입력에 대응되는 플랜을 수신하고, 플랜에 따라 '이번주 일정'을 디스플레이에 표시할 수 있다.

[218] 전자 장치가 퀵 커맨드와 같이, 복수의 기능들을 실행하기 위한 지정된 발언을 식별한 일 실시예에서, 상기 복수의 기능들을 보다 신속하게 실행하기 위한 방안이 요구될 수 있다. 상술한 바와 같은, 일 실시예에 따른, 전자 장치(an electronic device)(예, 도 3의 전자 장치(101))는, 통신 회로(예, 도 3의 통신 회로(330)), 마이크(예, 도 3의 마이크(320)), 스피커(예, 도 3의 스피커(310)), 및 프로세서(예, 도 3의 프로세서(120))를 포함할 수 있다. 상기 프로세서는, 상기 마이크를 통하여 제1 오디오 신호를 식별하도록, 구성될 수 있다. 상기 프로세서는, 상기 제1 오디오 신호로부터 복수의 기능들을 순차적으로 실행하기 위한 발언(예, 도 2의 발언(220))을 식별하는 것에 기반하여, 상기 복수의 기능들 중 적어도 하나의 제1 기능을 실행하도록, 구성될 수 있다. 상기 프로세서는, 상기 발언을 식별하는 것에

기반하여, 상기 통신 회로를 통해 연결된 외부 전자 장치(예, 도 3의 외부 전자 장치(210))를 이용하여, 상기 복수의 기능들 중에서, 상기 적어도 하나의 제1 기능과 상이한 적어도 하나의 제2 기능을 실행하도록, 구성될 수 있다. 상기 프로세서는, 상기 스피커를 통하여, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능을 실행한 결과들을 표현하는(representing) 제2 오디오 신호들을, 상기 발언과 관련된 순서를 따라 순차적으로 출력하도록, 구성될 수 있다. 일 실시예에 따른, 전자 장치는, 복수의 기능들을 실행하기 위한 발언에 기반하여, 전자 장치, 및 외부 전자 장치를 이용하여 상기 복수의 기능들을 보다 신속하게 실행할 수 있다.

[219] 예를 들어, 상기 프로세서는, 상기 프로세서에 의해 실행되는 제1 보이스 엔진(예, 도 3의 제1 보이스 엔진(370)), 및 상기 외부 전자 장치에 의해 실행되는 제2 보이스 엔진(예, 도 3의 제2 보이스 엔진(380))의 상태들을 획득하도록, 구성될 수 있다. 상기 프로세서는, 상기 상태들에 기반하여, 상기 복수의 기능들을, 상기 제1 보이스 엔진에 의해 실행될 상기 적어도 하나의 제1 기능으로, 상기 제2 보이스 엔진에 의해 실행될 상기 적어도 하나의 제2 기능으로 구분하도록, 구성될 수 있다.

[220] 예를 들어, 상기 프로세서는, 상기 상태들에 의해 지시되는, 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 각각이 기능(a function)을 실행 가능한지 여부에 기반하여, 상기 복수의 기능들 각각을, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능으로 구분하도록, 구성될 수 있다.

[221] 예를 들어, 상기 프로세서는, 상기 적어도 하나의 제1 기능에 대응하는 적어도 하나의 제1 보이스 커맨드를 상기 제1 보이스 엔진에 대응하는 제1 큐(예, 도 7의 제1 큐(710))로 할당하도록, 구성될 수 있다. 상기 프로세서는, 상기 적어도 하나의 제2 기능에 대응하는 적어도 하나의 제2 보이스 커맨드를, 또는 상기 제2 보이스 엔진에 대응하는 제2 큐(예, 도 7의 제2 큐(720))로 할당하도록, 구성될 수 있다.

[222] 예를 들어, 상기 프로세서는, 상기 보이스 커맨드들 중에서, 상기 제1 큐에 누적된(accumulated) 상기 적어도 하나의 제1 보이스 커맨드에 기반하여 상기 제1 보이스 엔진을 실행하도록, 구성될 수 있다.

[223] 예를 들어, 상기 프로세서는, 상기 보이스 커맨드들 중에서, 상기 제2 큐에 누적된 상기 적어도 하나의 제2 보이스 커맨드를 포함하는 적어도 하나의 신호를, 상기 외부 전자 장치로 송신하도록, 구성될 수 있다.

[224] 예를 들어, 상기 전자 장치는, 메모리(예, 도 3의 메모리(130))를 더 포함할 수 있다. 상기 프로세서는, 메모리 내에, 상기 결과들을 포함하고, 상기 결과들이 상기 순서를 따라 정렬된, 정보를 저장하도록, 구성될 수 있다.

[225] 예를 들어, 상기 정보는, 상기 순서를 따라 상기 제2 오디오 신호들에 포함된 발언들이 정렬된 큐(queue)를 포함할 수 있다.

[226] 예를 들어, 상기 전자 장치는, 디스플레이(예, 도 3의 디스플레이(260))를 더 포함할 수 있다. 상기 프로세서는, 상기 디스플레이 내에, 상기 결과들을 포함하는 상이한 시각적 객체들(예, 도 8의 시각적 객체들(882, 884, 886, 888))을 병렬로(in parallel) 표시하도록, 구성될 수 있다.

[227] 상술한 바와 같은, 일 실시예에 따른, 전자 장치(electronic device)의 방법은, 상기 전자 장치의 마이크를 통하여 제1 오디오 신호를 수신하는 것에 기반하여, 상기 제1 오디오 신호에 포함된 발언을 식별하는 동작(예, 도 12의 동작(1210))을 포함할 수 있다. 상기 방법은, 상기 발언으로부터 복수의 보이스 커맨드들을 식별하는 것에 기반하여, 보이스 커맨드를 처리하기 위하여 상기 전자 장치의 프로세서에 의해 실행되는 제1 보이스 엔진, 및 상기 전자 장치의 통신 회로를 통해 연결된 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 상태들을 획득하는 동작(예, 도 12의 동작(1220))을 포함할 수 있다. 상기 방법은, 상기 상태들에 기반하여, 상기 복수의 보이스 커맨드들을, 상기 제1 보이스 엔진에 대응하는 제1 보이스 커맨드, 및 상기 제2 보이스 엔진에 대응하는 제2 보이스 커맨드로 구분하기 위한 정보를 획득하는 동작(예, 도 12의 동작(1230))을 포함할 수 있다. 상기 방법은, 상기 정보에 기반하여, 상기 제1 보이스 커맨드에 대응하는 제1 기능을, 상기 제1 보이스 엔진을 이용하여 실행하고, 상기 외부 전자 장치로 상기 제2 보이스 커맨드에 대응하는 제2 기능의 실행을 요청하는 동작(예, 도 12의 동작들(1240, 1250))을 포함할 수 있다. 상기 방법은, 상기 전자 장치의 스피커를 통하여, 상기 발언에 의해 지시되는 상기 복수의 보이스 커맨드들의 순서에 기반하여, 상기 제1 기능, 및 상기 제2 기능을 실행한 결과들을 표현한(representing) 제2 오디오 신호를 출력하는 동작(예, 도 12의 동작(1260))을 포함할 수 있다.

[228] 예를 들어, 상기 발언을 식별하는 동작은, 상기 제1 오디오 신호를 수신하는 것에 응답하여, 상기 외부 전자 장치로 상기 제1 오디오 신호에 대응하는 텍스트 데이터를 요청하는 동작을 포함할 수 있다. 상기 발언을 식별하는 동작은, 상기 외부 전자 장치로부터 수신된 상기 텍스트 데이터에 기반하여 상기 발언을 식별하는 동작을 포함할 수 있다.

[229] 예를 들어, 상기 상태들을 획득하는 동작은, 상기 전자 장치의 메모리 내에서, 상기 순서에 기반하는 상기 복수의 보이스 커맨드들의 실행을 촉발(triggering)하기 위한 지정된 발언을 식별하는 동작을 포함할 수 있다. 상기 상태들을 획득하는 동작은, 상기 제1 오디오 신호에 포함된 상기 발언이 상기 지정된 발언에 대응하는지 여부에 기반하여, 상기 발언에 매칭된 상기 복수의 보이스 커맨드들을 식별하는 동작을 포함할 수 있다.

[230] 예를 들어, 상기 요청하는 동작은, 상기 외부 전자 장치로 상기 제2 기능의 실행을 요청한 이후, 상기 외부 전자 장치로부터 상기 제2 기능을 실행한 결과를 포함하는 정보를 수신하는 동작을 포함할 수 있다. 상기 요청하는 동작은, 상기 메모리 내에, 상기 순서에 기반하여 상기 외부 전자 장치로부터 수신된 상기 정보를 저장하는 동작을 포함할 수 있다.

[231] 예를 들어, 상기 정보를 획득하는 동작은, 상기 상태들에 기반하여, 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 각각이 상기 복수의 보이스 커맨드들을 실행 가능한지 여부에 기반하여, 상기 정보를 획득하는 동작을 포함할 수 있다.

[232] 예를 들어, 상기 출력하는 동작은, 상기 전자 장치의 디스플레이 내에, 상기 제1 기능, 및 상기 제2 기능을 실행한 결과들 각각에 대응하는 시각적 객체들을 표시하는 동작을 포함할 수 있다.

[233] 상술한 바와 같은, 일 실시예에 따른, 전자 장치의 방법은, 상기 전자 장치의 마이크를 통하여 제1 오디오 신호를 식별하는 동작(예, 도 10의 동작(1010))을 포함할 수 있다. 상기 방법은, 상기 제1 오디오 신호로부터 복수의 기능들을 순차적으로 실행하기 위한 발언을 식별하는 것에 기반하여, 상기 전자 장치의 프로세서에 기반하여 상기 복수의 기능들 중 적어도 하나의 제1 기능을 실행하는 동작(예, 도 10의 동작(1030))을 포함할 수 있다. 상기 방법은, 상기 발언을 식별하는 것에 기반하여, 상기 전자 장치의 통신 회로를 통해 연결된 외부 전자 장치를 이용하여, 상기 복수의 기능들 중에서, 상기 적어도 하나의 제1 기능과 상이한 적어도 하나의 제2 기능을 실행하는 동작(예, 도 10의 동작(1040))을 포함할 수 있다. 상기 방법은, 상기 전자 장치의 스피커를 통하여, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능을 실행한 결과들을 표현하는 제2 오디오 신호들을, 상기 발언과 관련된 순서를 따라 순차적으로 출력하는 동작(예, 도 10의 동작(1050))을 포함할 수 있다.

[234] 예를 들어, 상기 방법은, 상기 프로세서에 의해 실행되는 제1 보이스 엔진, 및 상기 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 상태들을 획득하는 동작을 포함할 수 있다. 상기 방법은, 상기 상태들에 기반하여, 상기 복수의 기능들 각각을, 상기 제1 보이스 엔진에 의해 실행될 상기 적어도 하나의 제1 기능으로, 상기 제2 보이스 엔진에 의해 실행될 상기 적어도 하나의 제2 기능으로 구분하는 동작을 포함할 수 있다.

[235] 예를 들어, 상기 구분하는 동작은, 상기 상태들에 의해 지시되는, 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 각각이 기능을 실행 가능한지 여부에 기반하여, 상기 복수의 기능들 각각을, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능으로 구분하는 동작을 포함할 수 있다.

[236] 예를 들어, 상기 구분하는 동작은, 상기 적어도 하나의 제1 기능에 대응하는 적어도 하나의 제1 보이스 커맨드를 상기 제1 보이스 엔진에 대응하는 제1 큐로 할당하는 동작을 포함할 수 있다. 상기 구분하는 동작은, 상기 적어도 하나의 제2 기능에 대응하는 적어도 하나의 제2 보이스 커맨드를, 또는 상기 제2 보이스 엔진에 대응하는 제2 큐로 할당하는 동작을 포함할 수 있다.

[237] 예를 들어, 상기 출력하는 동작은, 상기 결과들을 표현하고, 상기 순서를 따라 정렬된, 발언들의 큐에 기반하여, 상기 제2 오디오 신호들을 출력하는 동작을 포함할 수 있다.

[238] 상술한 바와 같은, 일 실시예에 따른, 전자 장치(electronic device)(예, 도 3의 전자 장치(101))는, 통신 회로(예, 도 3의통신 회로(330)), 마이크(예, 도 3의 마이크(320)), 스피커(예, 도 3의 스피커(310)), 및 프로세서(예, 도 3의 프로세서(120))를 포함할 수 있다. 상기 프로세서는, 상기 마이크를 통하여 제1 오디오 신호를 수신하는 것에 기반하여, 상기 제1 오디오 신호에 포함된 발언(예, 도 2의 발언(220))을 식별하도록, 구성될 수 있다. 상기 프로세서는, 상기 발언으로부터 복수의 보이스 커맨드들(voice commands)(예, 도 2의 보이스 커맨드들(230, 240))을 식별하는 것에 기반하여, 보이스 커맨드(a voice command)를 처리하기 위하여 상기 프로세서에 의해 실행되는 제1 보이스 엔진(예, 도 3의 제1 보이스 엔진(370)), 및 상기 통신 회로를 통해 연결된 외부 전자 장치(예, 도 3의 외부 전자 장치(210))에 의해 실행되는 제2 보이스 엔진(예, 도 3의 제2 보이스 엔진(380))의 상태들을 획득하도록, 구성될 수 있다. 상기 프로세서는, 상기 상태들에 기반하여, 상기 복수의 보이스 커맨드들을, 상기 제1 보이스 엔진에 대응하는 제1 보이스 커맨드, 및 상기 제2 보이스 엔진에 대응하는 제2 보이스 커맨드로 구분하기 위한 정보를 획득하도록, 구성될 수 있다. 상기 프로세서는, 상기 정보에 기반하여, 상기 제1 보이스 커맨드에 대응하는 제1 기능을, 상기 제1 보이스 엔진을 이용하여 실행하고, 상기 외부 전자 장치로 상기 제2 보이스 커맨드에 대응하는 제2 기능의 실행을 요청하도록, 구성될 수 있다. 상기 프로세서는, 상기 스피커를 통하여, 상기 발언에 의해 지시되는(indicated by) 상기 복수의 보이스 커맨드들의 순서에 기반하여, 상기 제1 기능, 및 상기 제2 기능을 실행한 결과들을 표현한(representing) 제2 오디오 신호를 출력하도록, 구성될 수 있다.

[239] 예를 들어, 상기 전자 장치는, 메모리(예, 도 3의 메모리(130))를 더 포함할 수 있다. 상기 프로세서는, 상기 메모리 내에서, 상기 순서에 기반하는 상기 복수의 보이스들의 실행을 촉발(triggering)하기 위한 지정된 발언을 식별하도록, 구성될 수 있다. 상기 프로세서는, 상기 제1 오디오 신호에 포함된 상기 발언이 상기 지정된 발언에 대응하는지 여부에 기반하여, 상기 발언에 매칭된 상기 복수의 보이스 커맨드들을 식별하도록, 구성될 수 있다.

[240] 예를 들어, 상기 프로세서는, 상기 외부 전자 장치로 상기 제2 기능의 실행을 요청한 이후, 상기 외부 전자 장치로부터 상기 제2 기능을 실행한 결과를 포함하는 정보를 수신하도록, 구성될 수 있다. 상기 프로세서는, 상기 메모리 내에, 상기 순서에 기반하여 상기 외부 전자 장치로부터 수신된 상기 정보를 저장하도록, 구성될 수 있다.

[241] 예를 들어, 상기 지정된 발언은, 상기 복수의 보이스 커맨드들과 관련된 하나 이상의 단어들과 상이한 다른 단어를 포함할 수 있다.

[242] 예를 들어, 상기 프로세서는, 상기 상태들에 기반하여, 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 각각이 상기 복수의 보이스 커맨드들을 실행 가능한지 여부에 기반하여, 상기 정보를 획득하도록, 구성될 수 있다.

[243] 예를 들어, 상기 프로세서는, 상기 제1 오디오 신호를 수신하는 것에 응답하여, 상기 외부 전자 장치로 상기 제1 오디오 신호에 대응하는 텍스트 데이터를 요청하도록, 구성될 수 있다. 상기 프로세서는, 상기 외부 전자 장치로부터 수신된 상기 텍스트 데이터에 기반하여 상기 발언을 식별하도록, 구성될 수 있다.

[244] 예를 들어, 상기 전자 장치는, 디스플레이를 더 포함할 수 있다. 상기 프로세서는, 상기 디스플레이 내에, 상기 스피커를 통하여 출력되는 상기 제2 오디오 신호와 함께, 상기 제1 기능, 및 상기 제2 기능을 실행한 결과들 각각에 대응하는 시각적 객체들을 표시하도록, 구성될 수 있다.

[245] 예를 들어, 상기 프로세서는, 상기 상태들에 대응하는, 상기 제1 보이스 엔진, 및 상기 제2 보이스 엔진의 대기 시간들(latency times)에 기반하여, 상기 복수의 보이스 커맨드들 각각을, 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 중 어느 하나로 매칭하여, 상기 정보를 획득하도록, 구성될 수 있다.

[246] 예를 들어, 상기 프로세서는, 상기 상태들에 기반하여, 상기 복수의 보이스 커맨드들을, 상기 제1 보이스 엔진, 또는 상기 제2 보이스 엔진 중 적어도 하나를 이용하여 실행하기 위한 스케줄을 나타내는 상기 정보를 획득하도록, 구성될 수 있다.

[247] 상술한 바와 같은, 일 실시예에 따른, 전자 장치는, 통신 회로, 마이크, 스피커, 적어도 하나의 인스트럭션를 저장하기 위한 메모리 및 적어도 하나의 프로세서를 포함할 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 마이크를 통하여 수신된 제1 오디오 신호를 식별하도록, 구성될 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 복수의 기능들을 순차적으로 실행하기 위한 발언을 포함하는 제1 오디오 신호를 식별하는 것에 기반하여, 상기 복수의 기능들 중 적어도 하나의 제1 기능을 수행하도록, 구성될 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 제1 오디오 신호가 발언을 포함함을 식별하는 것에 기반하여, 상기 통신 회로를 통해 외부 전자 장치의 상기 복수의 기능들 중에서 적어도 하나의 제2 기능을 수행하도록, 구성될 수 있다. 상기 적어도 하나의 제2 기능은, 상기 복수의 기능들 중에서, 상기 적어도 하나의 제1 기능과 상이할 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 스피커를 통하여, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능을 수행한 적어도 하나의 결과를 표현하는(representing) 제2 오디오 신호들을, 상기 발언과 관련된 순서에 기반하여 순차적으로 출력하도록, 구성될 수 있다.

[248] 예를 들어, 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 전자 장치에서 실행되는 제1 보이스 엔진의 상태, 및 상기 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 상태를 획득하도록, 구성될 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 제1 보이스 엔진의 상태 및 상기 제2 보이스 엔진의 상태에 기반하여,

상기 복수의 기능들을, 상기 제1 보이스 엔진에 의해 수행될 상기 적어도 하나의 제1 기능 및 상기 제2 보이스 엔진에 의해 실행될 상기 적어도 하나의 제2 기능으로 구분하도록, 구성될 수 있다.

[249] 예를 들어, 상기 제1 보이스 엔진의 상기 상태는 상기 제1 보이스 엔진이 상기 복수의 기능들 중 어느 하나를 수행할 수 있는 지 여부를 나타내고, 상기 제2 보이스 엔진의 상기 상태는 상기 제2 보이스 엔진이 상기 복수의 기능들 중 하나를 수행할 수 있는지 여부를 나타낼 수 있다.

[250] 예를 들어, 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 적어도 하나의 제1 기능에 대응하는 적어도 하나의 제1 보이스 커맨드를 상기 제1 보이스 엔진에 대응하는 제1 큐로 할당하도록, 구성될 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 적어도 하나의 제2 기능에 대응하는 적어도 하나의 제2 보이스 커맨드를, 또는 상기 제2 보이스 엔진에 대응하는 제2 큐로 할당하도록, 구성될 수 있다.

[251] 예를 들어, 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 제1 큐에 누적된(accumulated) 상기 적어도 하나의 제1 보이스 커맨드에 기반하여 상기 제1 보이스 엔진을 실행하도록, 구성될 수 있다.

[252] 예를 들어, 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 제2 큐에 누적된 상기 적어도 하나의 제2 보이스 커맨드를 나타내는 적어도 하나의 신호를, 상기 외부 전자 장치로 송신하도록, 구성될 수 있다.

[253] 예를 들어, 상기 전자 장치는, 메모리를 포함할 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 순서에 기반하여, 상기 메모리 내에 상기 적어도 하나의 결과를 포함하는 정보를 저장하도록, 구성될 수 있다.

[254] 예를 들어, 상기 정보는, 적어도 하나의 제2 오디오 신호 내 적어도 하나의 발언이 상기 순서에 기반하여 배열된 큐(queue)를 포함할 수 있다.

[255] 예를 들어, 상기 전자 장치는, 디스플레이를 포함할 수 있다. 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여, 상기 적어도 하나의 결과를 병렬로(in a parallel) 표시하기 위한 상이한 시각적 객체들을 표시하기 위하여 상기 디스플레이를 제어하도록, 구성될 수 있다.

[256] 상술한 바와 같은, 일 실시예에 따른, 전자 장치(electronic device)의 방법은, 상기 전자 장치의 마이크를 통하여 제1 오디오 신호를 수신하는 것에 기반하여, 상기 제1 오디오 신호 내 발언을 식별하는 동작을 포함할 수 있다. 상기 방법은, 상기 발언으로부터 복수의 보이스 커맨드들을 식별하는 것에 기반하여, 상기 전자 장치에 의해 실행되는 제1 보이스 엔진의 상태, 및 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 상태를 획득하는 동작을 포함할 수 있다. 상기 외부 전자 장치는 상기 전자 장치의 통신 회로를 통해 연결될 수 있다. 상기 방법은, 상기

제1 보이스 엔진의 상태 및 상기 제2 보이스 엔진의 상태에 기반하여, 상기 복수의 보이스 커맨드들을, 상기 제1 보이스 엔진에 의해 수행되는 적어도 하나의 제1 보이스 커맨드, 및 상기 제2 보이스 엔진에 의해 수행되는 제2 보이스 커맨드로 구분하기 위한 정보를 획득하는 동작을 포함할 수 있다. 상기 방법은, 상기 적어도 하나의 제1 보이스 커맨드에 대응하는 제1 기능을, 상기 제1 보이스 엔진을 이용하여 실행하고, 상기 적어도 하나의 제2 보이스 커맨드에 대응하는 제2 기능을 수행하기 위하여 외부 전자 장치로 요청하는 동작을 포함할 수 있다. 상기 방법은, 상기 전자 장치의 스피커를 통하여, 상기 발언에 의해 지시되는 상기 복수의 보이스 커맨드들의 순서에 기반하여, 상기 제1 기능, 및 상기 제2 기능을 실행한 적어도 하나의 결과를 표현한(representing) 적어도 하나의 제2 오디오 신호를 출력하는 동작을 포함할 수 있다.

[257] 예를 들어, 상기 제1 오디오 신호 내에서 상기 발언을 식별하는 동작은, 상기 외부 전자 장치로부터 상기 제1 오디오 신호에 대응하는 텍스트 데이터를 수신하는 동작을 포함할 수 있다. 상기 제1 오디오 신호 내에서 상기 발언을 식별하는 동작은, 상기 텍스트 데이터에 기반하여 상기 제1 오디오 신호 내 상기 발언을 식별하는 동작을 포함할 수 있다.

[258] 예를 들어, 상기 제1 보이스 엔진의 상태 및 상기 제2 보이스 엔진의 상태를 획득하는 동작은, 상기 전자 장치의 메모리 내에서, 상기 순서에 기반하는 상기 복수의 보이스 커맨드들의 수행을 촉발(triggering)하기 위한 지정된 발언을 식별하는 동작을 포함할 수 있다. 상기 제1 보이스 엔진의 상태 및 상기 제2 보이스 엔진의 상태를 획득하는 동작은, 상기 제1 오디오 신호 내 상기 발언이 상기 지정된 발언에 대응하는지 여부에 기반하여, 상기 복수의 보이스 커맨드들을 식별하는 동작을 포함할 수 있다.

[259] 예를 들어, 상기 제2 기능을 수행하기 위하여 상기 외부 전자 장치에 요청하는 동작은, 상기 제2 기능을 수행한 결과를 포함하는 정보를, 상기 외부 전자 장치로부터 수신하는 동작을 포함할 수 있다. 상기 제2 기능을 수행하기 위하여 상기 외부 전자 장치에 요청하는 동작은, 상기 전자 장치의 메모리 내에, 상기 순서에 기반하여 상기 상기 정보를 저장하는 동작을 포함할 수 있다.

[260] 예를 들어, 상기 제1 보이스 엔진의 상기 상태는 상기 제1 보이스 엔진이 복수의 기능들 중 하나를 수행할 수 있는지 여부를 나타내고, 상기 제2 보이스 엔진의 상기 상태는 상기 제2 보이스 엔진이 상기 복수의 기능들 중 어느 하나를 수행할 수 있는지 여부를 나타낼 수 있다.

[261] 예를 들어, 상기 적어도 하나의 제2 오디오 신호를 출력하는 동작은, 상기 제2 기능을 수행한 결과 및 상기 제1 기능을 수행한 결과를 보여주는 적어도 하나의 시각적 객체를 표시하기 위하여, 상기 전자 장치의 디스플레이를 제어하는 동작을 포함할 수 있다.

[262] 상술한 바와 같은, 일 실시예에 따른, 전자 장치의 방법은, 상기 전자 장치의 마이크를 통해 수신된 제1 오디오 신호를 식별하는 동작을 포함할 수 있다. 상기 방

법은, 상기 제1 오디오 신호가 복수의 기능들을 순차적으로 실행하기 위한 발언을 포함함을 식별하는 것에 기반하여, 상기 복수의 기능들로부터 적어도 하나의 제1 기능을 수행하는 동작을 포함할 수 있다. 상기 방법은, 상기 제1 오디오 신호가 상기 발언을 포함함을 식별하는 것에 기반하여, 상기 외부 전자 장치의 통신 회로를 통해 연결된 외부 전자 장치에서, 상기 복수의 기능들 중 적어도 하나의 제2 기능을 수행하는 동작을 포함할 수 있다. 상기 적어도 하나의 제2 기능은 상기 적어도 하나의 제1 기능과 다를 수 있다. 상기 방법은, 상기 전자 장치의 스피커를 통하여, 상기 발언과 관련된 발언에 기반하여 상기 적어도 하나의 제1 기능 및 상기 적어도 하나의 제2 기능을 수행한 적어도 하나의 결과를 표현하는 적어도 하나의 제2 오디오 신호를 순차적으로 출력하는 동작을 포함할 수 있다.

[263] 예를 들어, 상기 방법은, 상기 전자 장치에서 실행되는 제1 보이스 엔진의 상태 및 상기 외부 전자 장치에서 실행되는 제2 보이스 엔진의 상태를 획득하는 동작을 포함할 수 있다. 상기 방법은, 상기 제1 보이스 엔진의 상태 및 상기 제2 보이스 엔진의 상태에 기반하여, 상기 복수의 기능들을, 상기 제1 보이스 엔진에 의해 수행될 상기 적어도 하나의 제1 기능 및 상기 제2 보이스 엔진에 의해 수행될 상기 적어도 하나의 제2 기능으로 구분하는 동작을 포함할 수 있다.

[264] 예를 들어, 상기 제1 보이스 엔진의 상태는, 상기 제1 보이스 엔진이 상기 복수의 기능들 중 어느 하나를 실행할 수 있는지 여부를 나타낼 수 있다. 상기 제2 보이스 엔진의 상태는, 상기 제2 보이스 엔진이 상기 복수의 기능들 중 어느 하나를 실행할지 여부를 나타낼 수 있다.

[265] 예를 들어, 상기 구분하는 동작은, 상기 제1 보이스 엔진에 대응하는 제1 큐로, 상기 적어도 하나의 제1 기능에 대응하는 적어도 하나의 제1 보이스 커맨드를 할당하는 동작을 포함할 수 있다. 상기 구분하는 동작은, 상기 제2 보이스 엔진에 대응하는 제2 큐로, 상기 적어도 하나의 제2 기능에 대응하는 제2 보이스 커맨드를 할당하는 동작을 포함할 수 있다.

[266] 예를 들어, 상기 출력하는 동작은, 상기 적어도 하나의 제2 오디오 신호 내 적어도 하나의 발언의 큐에 기반하여, 상기 발언에 따라 상기 적어도 하나의 결과를 표현하는 상기 적어도 하나의 발언을 출력하는 동작을 포함할 수 있다.

[267] 본 문서에 개시된 다양한 실시예들에 따른 전자 장치는 다양한 형태의 장치가 될 수 있다. 전자 장치는, 예를 들면, 휴대용 통신 장치(예: 스마트폰), 컴퓨터 장치, 휴대용 멀티미디어 장치, 휴대용 의료 기기, 카메라, 웨어러블 장치, 또는 가전 장치를 포함할 수 있다. 본 문서의 실시예에 따른 전자 장치는 전술한 기기들에 한정되지 않는다.

[268] 본 문서의 다양한 실시예들 및 이에 사용된 용어들은 본 문서에 기재된 기술적 특징들을 특정한 실시예들로 한정하려는 것이 아니며, 해당 실시예의 다양한 변경, 균등물, 또는 대체물을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관련하여, 유사한 또는 관련된 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다. 아이템에 대응하는 명사의 단수 형은 관련된 문맥상 명백하게 다르게 지시

하지 않는 한, 상기 아이템 한 개 또는 복수 개를 포함할 수 있다. 본 문서에서, "A 또는 B", "A 및 B 중 적어도 하나", "A 또는 B 중 적어도 하나", "A, B 또는 C", "A, B 및 C 중 적어도 하나", 및 "A, B, 또는 C 중 적어도 하나"와 같은 문구들 각각은 그 문구들 중 해당하는 문구에 함께 나열된 항목들 중 어느 하나, 또는 그들의 모든 가능한 조합을 포함할 수 있다. "제 1", "제 2", 또는 "첫째" 또는 "둘째"와 같은 용어들은 단순히 해당 구성요소를 다른 해당 구성요소와 구분하기 위해 사용될 수 있으며, 해당 구성요소들을 다른 측면(예: 중요성 또는 순서)에서 한정하지 않는다. 어떤(예: 제 1) 구성요소가 다른(예: 제 2) 구성요소에, "기능적으로" 또는 "통신적으로"라는 용어와 함께 또는 이런 용어 없이, "커플드" 또는 "커넥티드"라고 언급된 경우, 그것은 상기 어떤 구성요소가 상기 다른 구성요소에 직접적으로(예: 유선으로), 무선으로, 또는 제 3 구성요소를 통하여 연결될 수 있다는 것을 의미한다.

[269] 본 문서의 다양한 실시예들에서 사용된 용어 "모듈"은 하드웨어, 소프트웨어 또는 펌웨어로 구현된 유닛을 포함할 수 있으며, 예를 들면, 로직, 논리 블록, 부품, 또는 회로와 같은 용어와 상호 호환적으로 사용될 수 있다. 모듈은, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는, 상기 부품의 최소 단위 또는 그 일부가 될 수 있다. 예를 들면, 일실시예에 따르면, 모듈은 ASIC(application-specific integrated circuit)의 형태로 구현될 수 있다.

[270] 본 문서의 다양한 실시예들은 기기(machine)(예: 전자 장치(101)) 의해 읽을 수 있는 저장 매체(storage medium)(예: 내장 메모리(136) 또는 외장 메모리(138))에 저장된 하나 이상의 명령어들을 포함하는 소프트웨어(예: 프로그램(140))로서 구현될 수 있다. 예를 들면, 기기(예: 전자 장치(101))의 프로세서(예: 프로세서(120))는, 저장 매체로부터 저장된 하나 이상의 명령어들 중 적어도 하나의 명령을 호출하고, 그것을 실행할 수 있다. 이것은 기기가 상기 호출된 적어도 하나의 명령어에 따라 적어도 하나의 기능을 수행하도록 운영되는 것을 가능하게 한다. 상기 하나 이상의 명령어들은 컴파일러에 의해 생성된 코드 또는 인터프리터에 의해 실행될 수 있는 코드를 포함할 수 있다. 기기로 읽을 수 있는 저장 매체는, 비일시적(non-transitory) 저장 매체의 형태로 제공될 수 있다. 여기서, '비일시적'은 저장 매체가 실재(tangible)하는 장치이고, 신호(signal)(예: 전자기파)를 포함하지 않는다는 것을 의미할 뿐이며, 이 용어는 데이터가 저장 매체에 반영구적으로 저장되는 경우와 임시적으로 저장되는 경우를 구분하지 않는다.

[271] 일실시예에 따르면, 본 문서에 개시된 다양한 실시예들에 따른 방법은 컴퓨터 프로그램 제품(computer program product)에 포함되어 제공될 수 있다. 컴퓨터 프로그램 제품은 상품으로서 판매자 및 구매자 간에 거래될 수 있다. 컴퓨터 프로그램 제품은 기기로 읽을 수 있는 저장 매체(예: compact disc read only memory(CD-ROM))의 형태로 배포되거나, 또는 어플리케이션 스토어(예: 플레이 스토어™)를 통해 또는 두 개의 사용자 장치들(예: 스마트 폰들) 간에 직접, 온라인으로 배포(예: 다운로드 또는 업로드)될 수 있다. 온라인 배포의 경우에, 컴퓨터

프로그램 제품의 적어도 일부는 제조사의 서버, 어플리케이션 스토어의 서버, 또는 중계 서버의 메모리와 같은 기기로 읽을 수 있는 저장 매체에 적어도 일시 저장되거나, 임시적으로 생성될 수 있다.

[272] 다양한 실시예들에 따르면, 상기 기술한 구성요소들의 각각의 구성요소(예: 모듈 또는 프로그램)는 단수 또는 복수의 개체를 포함할 수 있으며, 복수의 개체 중 일부는 다른 구성요소에 분리 배치될 수도 있다. 다양한 실시예들에 따르면, 전술한 해당 구성요소들 중 하나 이상의 구성요소들 또는 동작들이 생략되거나, 또는 하나 이상의 다른 구성요소들 또는 동작들이 추가될 수 있다. 대체적으로 또는 추가적으로, 복수의 구성요소들(예: 모듈 또는 프로그램)은 하나의 구성요소로 통합될 수 있다. 이런 경우, 통합된 구성요소는 상기 복수의 구성요소들 각각의 구성요소의 하나 이상의 기능들을 상기 통합 이전에 상기 복수의 구성요소들 중 해당 구성요소에 의해 수행되는 것과 동일 또는 유사하게 수행할 수 있다. 다양한 실시예들에 따르면, 모듈, 프로그램 또는 다른 구성요소에 의해 수행되는 동작들은 순차적으로, 병렬적으로, 반복적으로, 또는 휴리스틱하게 실행되거나, 상기 동작들 중 하나 이상이 다른 순서로 실행되거나, 생략되거나, 또는 하나 이상의 다른 동작들이 추가될 수 있다.

[273] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 어플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[274] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 컴퓨터 저장 매체 또는 장치에 구체화

(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[275] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 이때, 매체는 컴퓨터로 실행 가능한 프로그램을 계속 저장하거나, 실행 또는 다운로드를 위해 임시 저장하는 것일 수도 있다. 또한, 매체는 단일 또는 수 개의 하드웨어가 결합된 형태의 다양한 기록수단 또는 저장수단일 수 있는데, 어떤 컴퓨터 시스템에 직접 접속되는 매체에 한정되지 않고, 네트워크 상에 분산 존재하는 것일 수도 있다. 매체의 예시로는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등을 포함하여 프로그램 명령어가 저장되도록 구성된 것이 있을 수 있다. 또한, 다른 매체의 예시로, 어플리케이션을 유통하는 앱 스토어나 기타 다양한 소프트웨어를 공급 내지 유통하는 사이트, 서버 등에서 관리하는 기록매체 내지 저장매체도 들 수 있다.

[276] 이상과 같이 실시예들이 비록 한정된 실시예와 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.

[277] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 특허청구범위의 범위에 속한다.

청구범위

[청구항 1] 전자 장치(electronic device)에 있어서,
통신 회로;
마이크;
스피커;
적어도 하나의 인스트럭션를 저장하기 위한 메모리; 및
적어도 하나의 프로세서를 포함하고, 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여,
상기 마이크를 통하여 수신된 제1 오디오 신호를 식별하고;
복수의 기능들을 순차적으로 실행하기 위한 발언을 포함하는 제1 오디오 신호를 식별하는 것에 기반하여, 상기 복수의 기능들 중 적어도 하나의 제1 기능을 수행하고;
상기 제1 오디오 신호가 발언을 포함함을 식별하는 것에 기반하여, 상기 통신 회로를 통해 외부 전자 장치의 상기 복수의 기능들 중에서 적어도 하나의 제2 기능을 수행하고, 상기 적어도 하나의 제2 기능은, 상기 복수의 기능들 중에서, 상기 적어도 하나의 제1 기능과 상이함; 및
상기 스피커를 통하여, 상기 적어도 하나의 제1 기능, 및 상기 적어도 하나의 제2 기능을 수행한 적어도 하나의 결과를 표현하는(representing) 제2 오디오 신호들을, 상기 발언과 관련된 순서에 기반하여 순차적으로 출력하도록, 구성된,
전자 장치.

[청구항 2] 제1항에 있어서, 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여,
상기 전자 장치에서 실행되는 제1 보이스 엔진의 상태, 및 상기 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 상태를 획득하고;
상기 제1 보이스 엔진의 상태 및 상기 제2 보이스 엔진의 상태에 기반하여, 상기 복수의 기능들을, 상기 제1 보이스 엔진에 의해 수행될 상기 적어도 하나의 제1 기능 및 상기 제2 보이스 엔진에 의해 실행될 상기 적어도 하나의 제2 기능으로 구분하도록, 구성된,
전자 장치.

[청구항 3] 제2항에 있어서, 상기 제1 보이스 엔진의 상기 상태는 상기 제1 보이스 엔진이 상기 복수의 기능들 중 어느 하나를 수행할 수 있는 지 여부를 나타내고, 상기 제2 보이스 엔진의 상기 상태는 상기 제2 보이스 엔진이 상기 복수의 기능들 중 하나를 수행할 수 있는지 여부를 나타내는,
전자 장치.

[청구항 4] 제2항에 있어서, 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여,

상기 적어도 하나의 제1 기능에 대응하는 적어도 하나의 제1 보이스 커맨드를 상기 제1 보이스 엔진에 대응하는 제1 큐로 할당하고; 및
상기 적어도 하나의 제2 기능에 대응하는 적어도 하나의 제2 보이스 커맨드를, 또는 상기 제2 보이스 엔진에 대응하는 제2 큐로 할당하도록, 구성된,
전자 장치.

[청구항 5] 제4항에 있어서, 상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여,
상기 제1 큐에 누적된(accumulated) 상기 적어도 하나의 제1 보이스 커맨드에 기반하여 상기 제1 보이스 엔진을 실행하도록, 구성된,
전자 장치.

[청구항 6] 제4항에 있어서, 상기 프로세서는,
상기 제2 큐에 누적된 상기 적어도 하나의 제2 보이스 커맨드를 나타내는 적어도 하나의 신호를, 상기 외부 전자 장치로 송신하도록, 구성된,
전자 장치.

[청구항 7] 제1항에 있어서, 메모리를 더 포함하고,
상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여,
상기 순서에 기반하여, 상기 메모리 내에 상기 적어도 하나의 결과를 포함하는 정보를 저장하도록, 구성된,
전자 장치.

[청구항 8] 제7항에 있어서, 상기 정보는,
적어도 하나의 제2 오디오 신호 내 적어도 하나의 발언이 상기 순서에 기반하여 배열된 큐(queue)를 포함하는,
전자 장치.

[청구항 9] 제1항에 있어서, 디스플레이를 더 포함하고,
상기 적어도 하나의 프로세서는 상기 적어도 하나의 인스트럭션들을 실행하여,
상기 적어도 하나의 결과를 병렬로(in a parallel) 표시하기 위한 상이한 시각적 객체들을 표시하기 위하여 상기 디스플레이를 제어하도록, 구성된,
전자 장치.

[청구항 10] 전자 장치(electronic device)의 방법에 있어서,
상기 전자 장치의 마이크를 통하여 제1 오디오 신호를 수신하는 것에 기반하여, 상기 제1 오디오 신호 내 발언을 식별하는 동작;
상기 발언으로부터 복수의 보이스 커맨드들을 식별하는 것에 기반하여,
상기 전자 장치에 의해 실행되는 제1 보이스 엔진의 상태, 및 외부 전자 장치에 의해 실행되는 제2 보이스 엔진의 상태를 획득하는 동작, 상기 외부 전자 장치는 상기 전자 장치의 통신 회로를 통해 연결됨;

상기 제1 보이스 엔진의 상태 및 상기 제2 보이스 엔진의 상태에 기반하여, 상기 복수의 보이스 커맨드들을, 상기 제1 보이스 엔진에 의해 수행되는 적어도 하나의 제1 보이스 커맨드, 및 상기 제2 보이스 엔진에 의해 수행되는 제2 보이스 커맨드로 구분하기 위한 정보를 획득하는 동작;

상기 적어도 하나의 제1 보이스 커맨드에 대응하는 제1 기능을, 상기 제1 보이스 엔진을 이용하여 실행하고, 상기 적어도 하나의 제2 보이스 커맨드에 대응하는 제2 기능을 수행하기 위하여 외부 전자 장치로 요청하는 동작; 및

상기 전자 장치의 스피커를 통하여, 상기 발언에 의해 지시되는 상기 복수의 보이스 커맨드들의 순서에 기반하여, 상기 제1 기능, 및 상기 제2 기능을 실행한 적어도 하나의 결과를 표현한(representing) 적어도 하나의 제2 오디오 신호를 출력하는 동작을 포함하는,

방법.

[청구항 11] 제10항에 있어서, 상기 제1 오디오 신호 내에서 상기 발언을 식별하는 동작은,

상기 외부 전자 장치로부터 상기 제1 오디오 신호에 대응하는 텍스트 데이터를 수신하는 동작;

상기 텍스트 데이터에 기반하여 상기 제1 오디오 신호 내 상기 발언을 식별하는 동작을 포함하는,

방법.

[청구항 12] 제10항에 있어서, 상기 제1 보이스 엔진의 상태 및 상기 제2 보이스 엔진의 상태를 획득하는 동작은,

상기 전자 장치의 메모리 내에서, 상기 순서에 기반하는 상기 복수의 보이스 커맨드들의 수행을 촉발(triggering)하기 위한 지정된 발언을 식별하는 동작; 및

상기 제1 오디오 신호 내 상기 발언이 상기 지정된 발언에 대응하는지 여부에 기반하여, 상기 복수의 보이스 커맨드들을 식별하는 동작을 포함하는,

방법.

[청구항 13] 제12한 항에 있어서, 상기 제2 기능을 수행하기 위하여 상기 외부 전자 장치에 요청하는 동작은,

상기 제2 기능을 수행한 결과를 포함하는 정보를, 상기 외부 전자 장치로부터 수신하는 동작;

상기 전자 장치의 메모리 내에, 상기 순서에 기반하여 상기 상기 정보를 저장하는 동작을 포함하는,

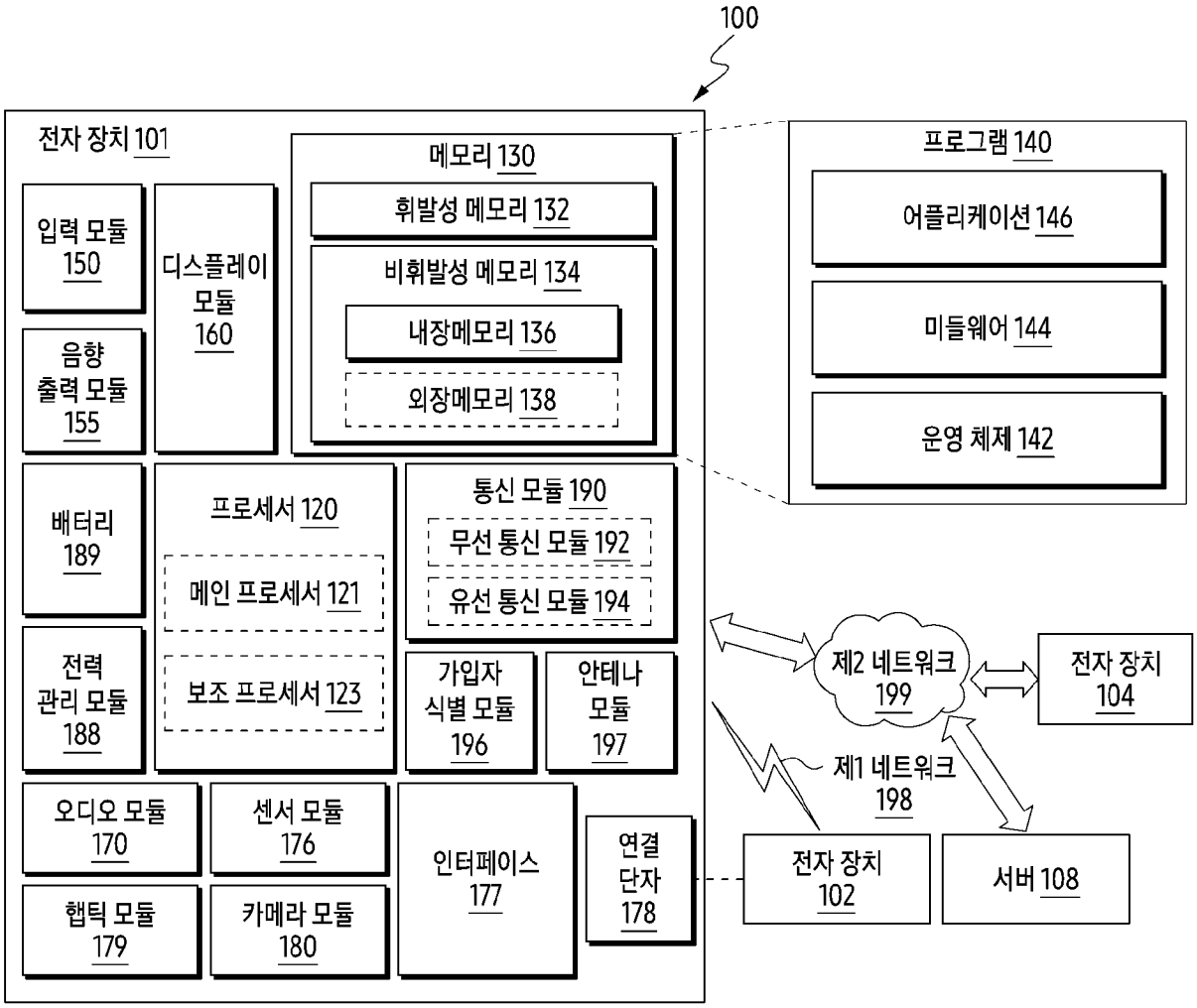
방법.

[청구항 14] 제10항에 있어서, 상기 제1 보이스 엔진의 상기 상태는 상기 제1 보이스 엔진이 복수의 기능들 중 하나를 수행할 수 있는지 여부를 나타내고, 상기

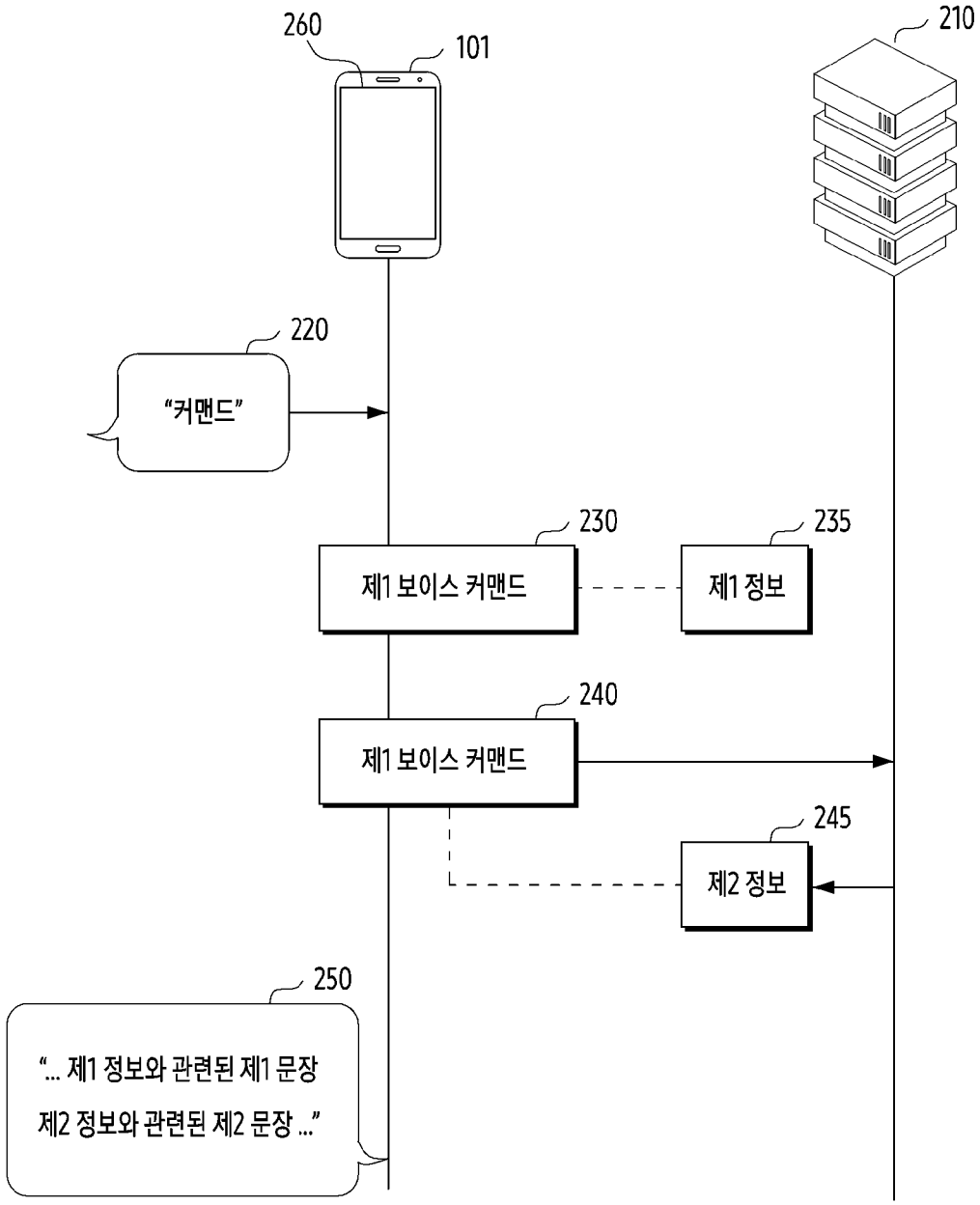
제2 보이스 엔진의 상기 상태는 상기 제2 보이스 엔진이 상기 복수의 기능들 중 어느 하나를 수행할 수 있는지 여부를 나타내는,
방법.

[청구항 15] 제10항에 있어서, 상기 적어도 하나의 제2 오디오 신호를 출력하는 동작은,
상기 제2 기능을 수행한 결과 및 상기 제1 기능을 수행한 결과를 보여주는 적어도 하나의 시각적 객체를 표시하기 위하여, 상기 전자 장치의 디스플레이를 제어하는 동작을 포함하는,
방법.

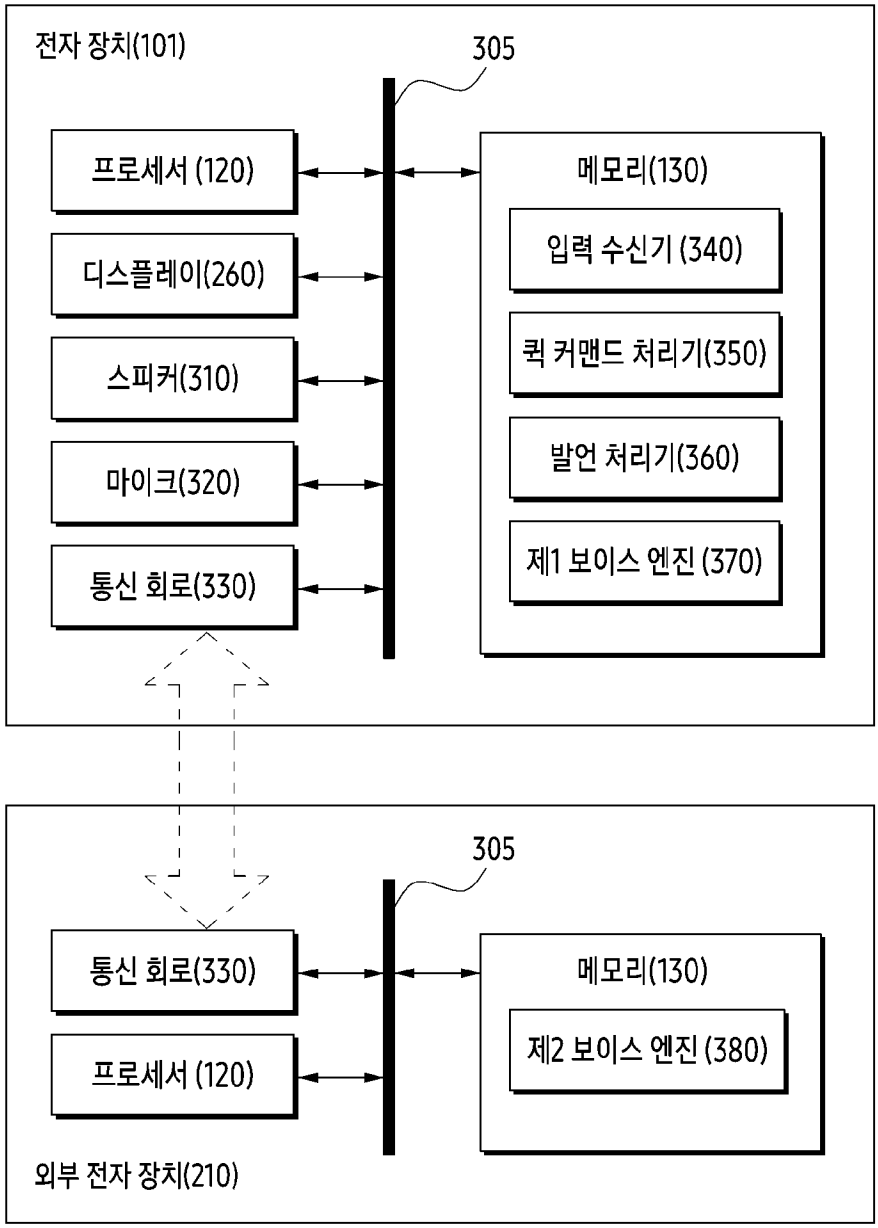
[도1]


100
전자 장치 101
입력 모듈 150
음향 출력 모듈 155
디스플레이 모듈 160
메모리 130
휘발성 메모리 132
비휘발성 메모리 134
내장메모리 136
외장메모리 138
프로그램 140
어플리케이션 146
미들웨어 144
운영 체제 142
배터리 189
프로세서 120
메인 프로세서 121
보조 프로세서 123
통신 모듈 190
무선 통신 모듈 192
유선 통신 모듈 194
전력 관리 모듈 188
가입자 식별 모듈 196
안테나 모듈 197
제2 네트워크 199
전자 장치 104
제1 네트워크 198
오디오 모듈 170
센서 모듈 176
인터페이스 177
연결 단자 178
햅틱 모듈 179
카메라 모듈 180
전자 장치 102
서버 108

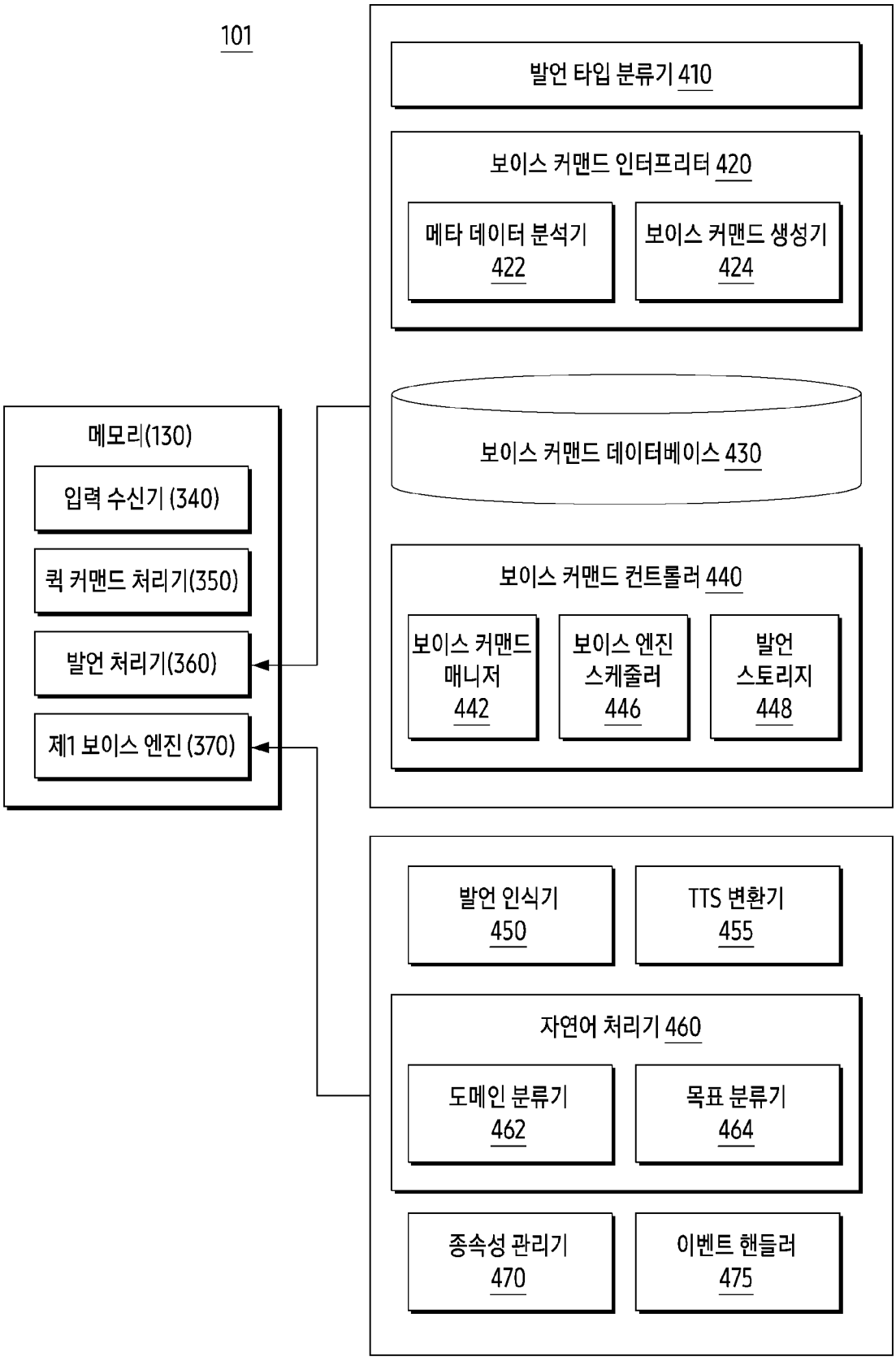
[도2]


260
101
210
220
"커맨드"
230
제1 보이스 커맨드
235
제1 정보
240
제1 보이스 커맨드
245
제2 정보
250
"... 제1 정보와 관련된 제1 문장
제2 정보와 관련된 제2 문장 ..."

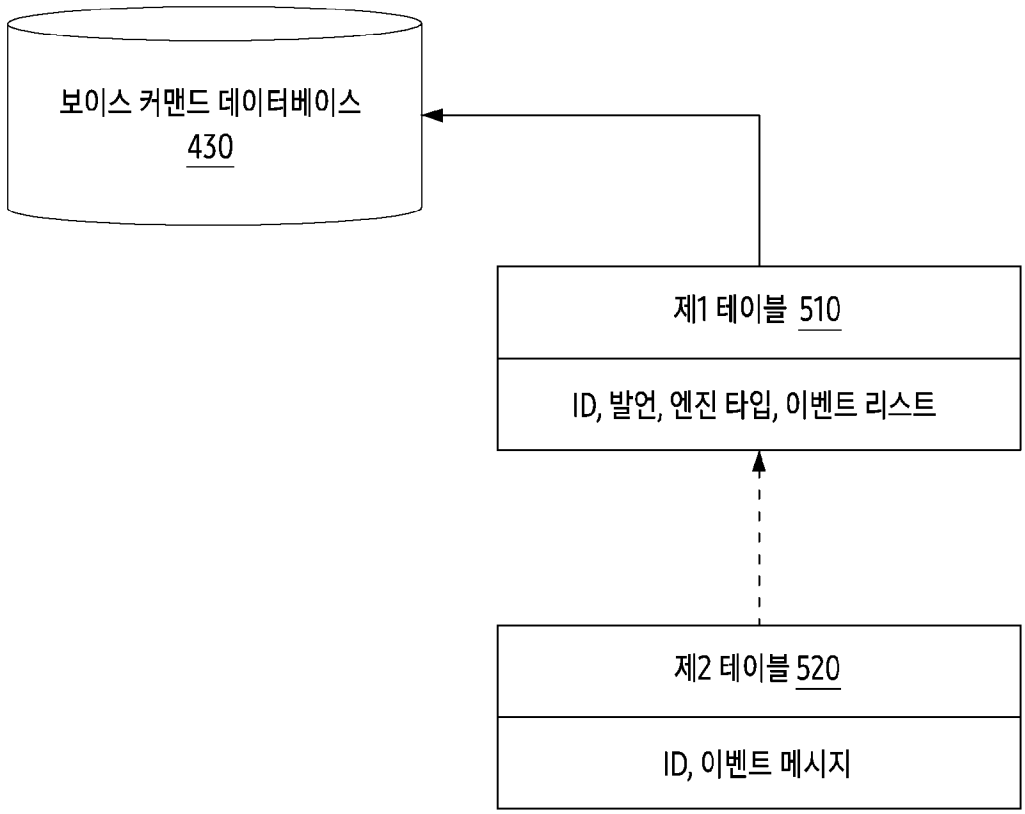
[도3]


전자 장치(101)
305
프로세서 (120)
디스플레이(260)
스피커(310)
마이크(320)
통신 회로(330)
메모리(130)
입력 수신기 (340)
퀵 커맨드 처리기(350)
발언 처리기(360)
제1 보이스 엔진 (370)
305
통신 회로(330)
프로세서 (120)
메모리(130)
제2 보이스 엔진 (380)
외부 전자 장치(210)

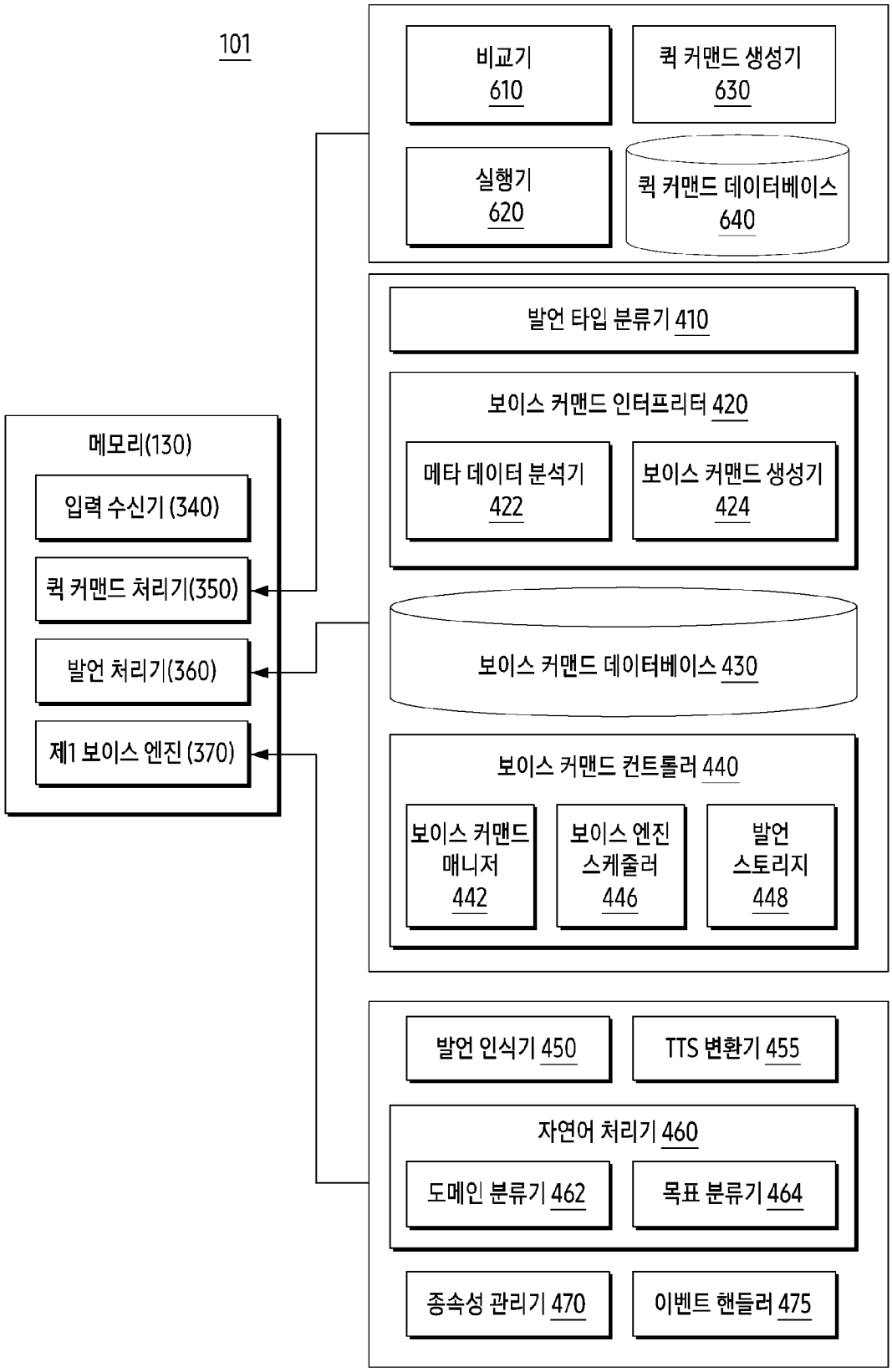
[도4]


101
발언 타입 분류기 410
보이스 커맨드 인터프리터 420
메타 데이터 분석기 422
보이스 커맨드 생성기 424
보이스 커맨드 데이터베이스 430
보이스 커맨드 컨트롤러 440
보이스 커맨드 매니저 442
보이스 엔진 스케줄러 446
발언 스토리지 448
메모리(130)
입력 수신기 (340)
퀵 커맨드 처리기(350)
발언 처리기(360)
제1 보이스 엔진 (370)
발언 인식기 450
TTS 변환기 455
자연어 처리기 460
도메인 분류기 462
목표 분류기 464
종속성 관리기 470
이벤트 핸들러 475

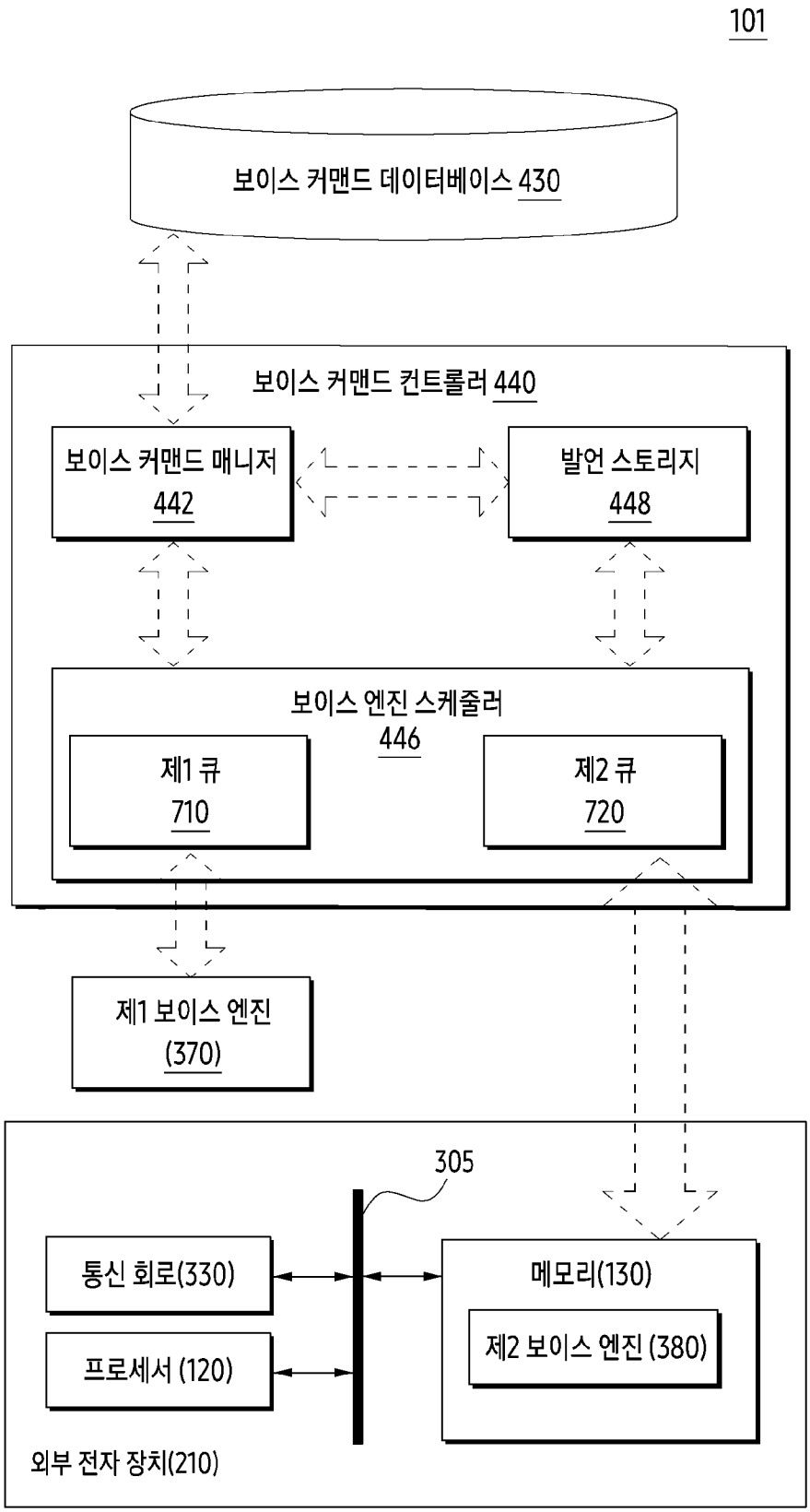
[도5]


보이스 커맨드 데이터베이스
430
제1 테이블 510
ID, 발언, 엔진 타입, 이벤트 리스트
제2 테이블 520
ID, 이벤트 메시지

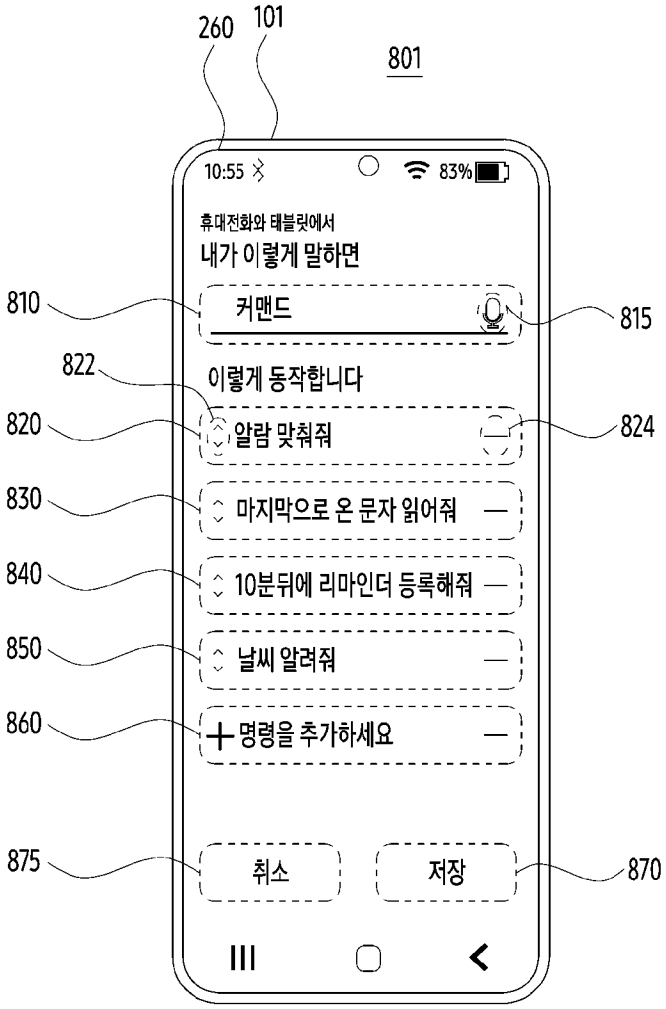
[도6]


101
비교기 610
퀵 커맨드 생성기 630
실행기 620
퀵 커맨드 데이터베이스 640
발언 타입 분류기 410
보이스 커맨드 인터프리터 420
메타 데이터 분석기 422
보이스 커맨드 생성기 424
보이스 커맨드 데이터베이스 430
보이스 커맨드 컨트롤러 440
보이스 커맨드 매니저 442
보이스 엔진 스케줄러 446
발언 스토리지 448
발언 인식기 450
TTS 변환기 455
자연어 처리기 460
도메인 분류기 462
목표 분류기 464
종속성 관리기 470
이벤트 핸들러 475
메모리(130)
입력 수신기 (340)
퀵 커맨드 처리기(350)
발언 처리기(360)
제1 보이스 엔진 (370)

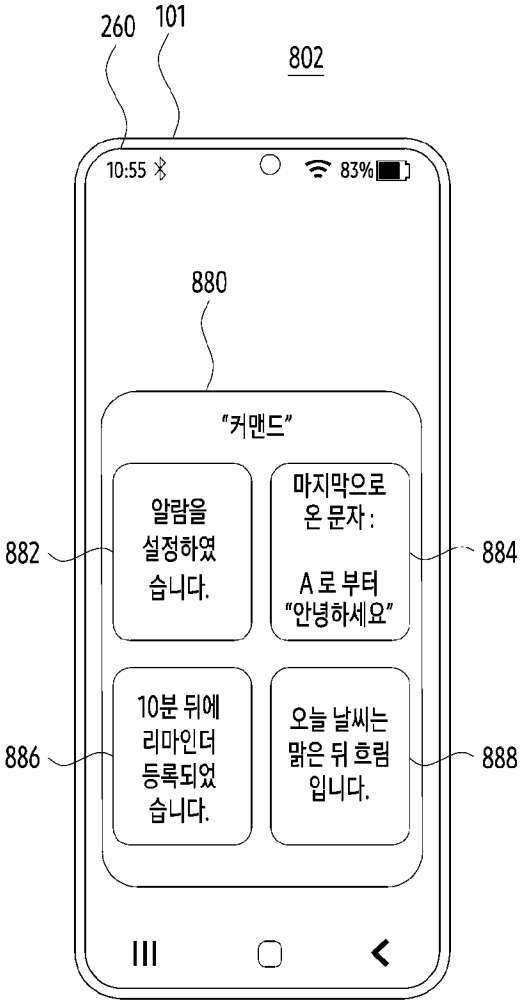
[도7]


101
보이스 커맨드 데이터베이스 430
보이스 커맨드 컨트롤러 440
보이스 커맨드 매니저 442
발언 스토리지 448
보이스 엔진 스케줄러 446
제1 큐 710
제2 큐 720
제1 보이스 엔진 (370)
305
통신 회로(330)
프로세서 (120)
메모리(130)
제2 보이스 엔진 (380)
외부 전자 장치(210)

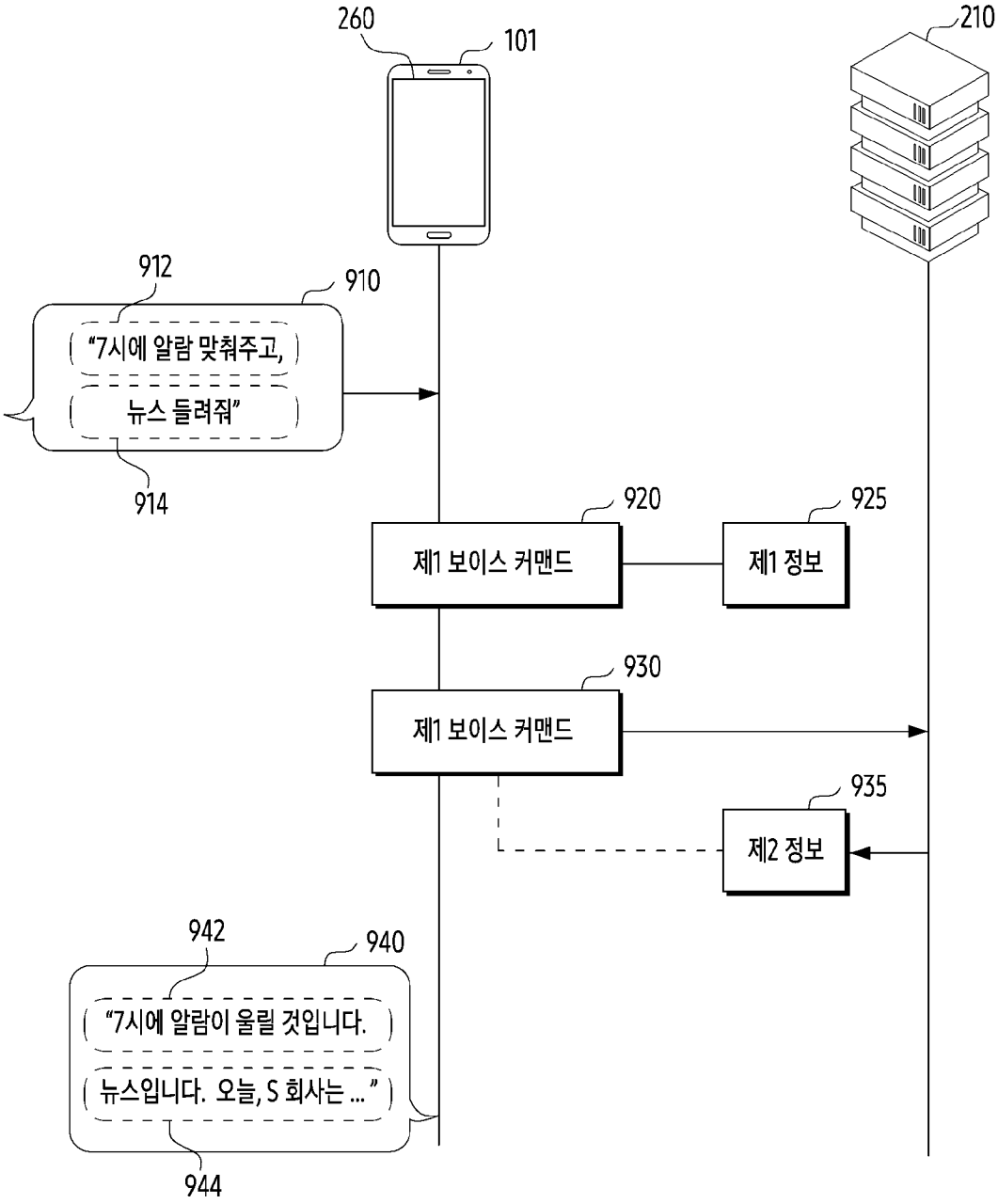
[도8a]


260
101
801
10:55
83%
휴대전화와 태블릿에서
내가 이렇게 말하면
810
커맨드
815
822
이렇게 동작합니다
820
알람 맞춰줘
824
830
마지막으로 온 문자 읽어줘
840
10분뒤에 리마인더 등록해줘
850
날씨 알려줘
860
명령을 추가하세요
875
취소
저장
870

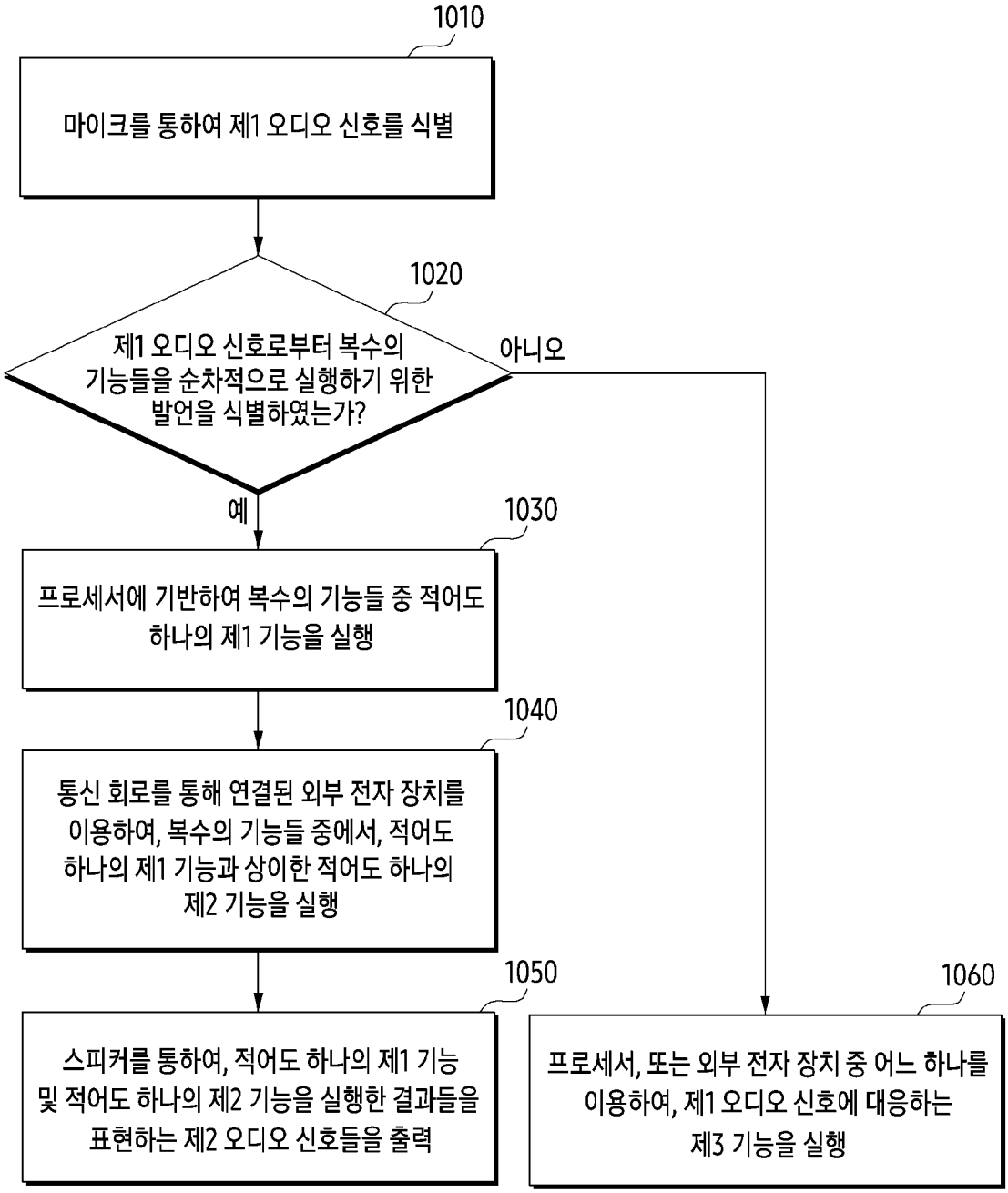
[도8b]


260
101
802
10:55
83%
880
"커맨드"
알람을 설정하였습니다.
882
마지막으로 온 문자 : A 로 부터 "안녕하세요"
884
10분 뒤에 리마인더 등록되었습니다.
886
오늘 날씨는 맑은 뒤 흐림 입니다.
888

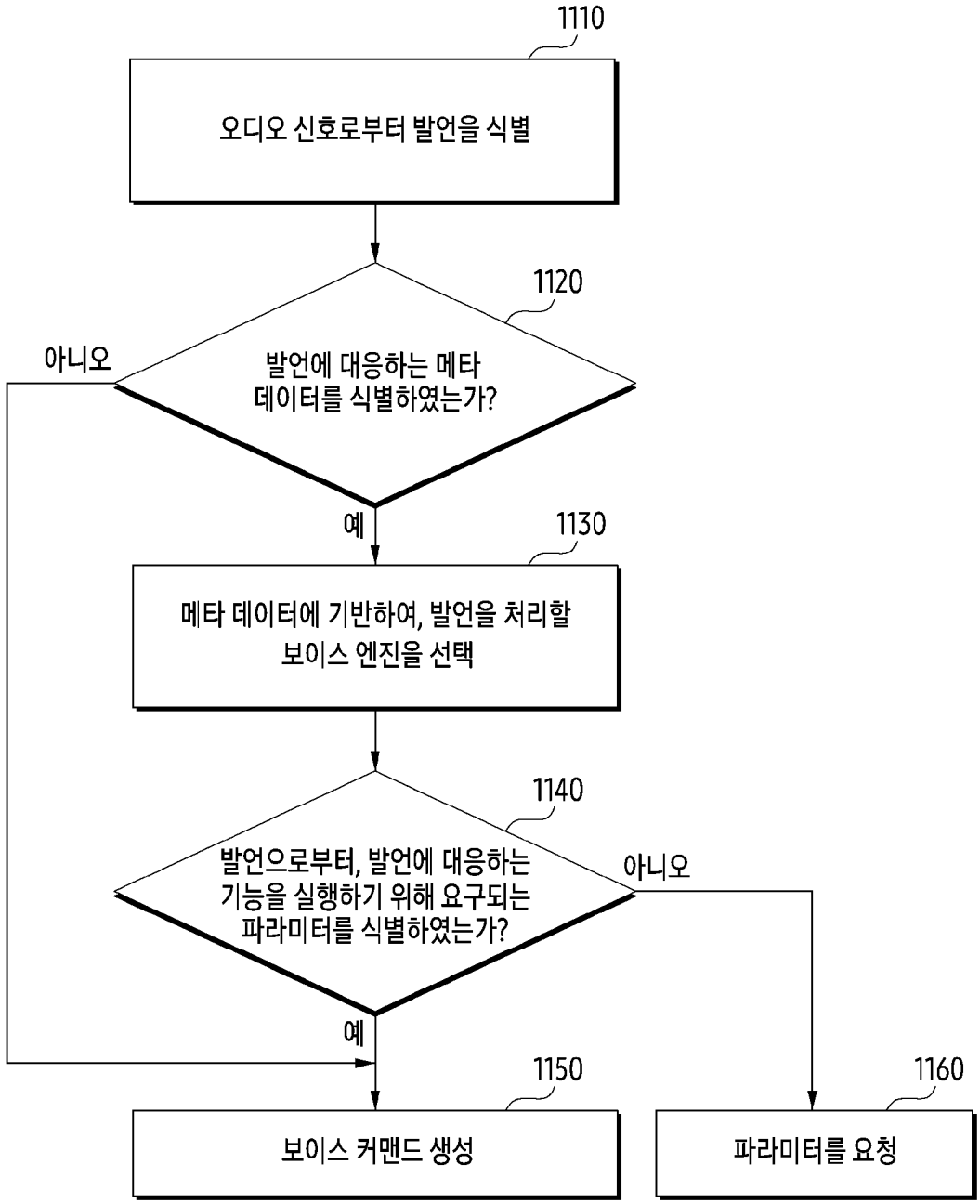
[도9]


260
101
210
912
910
"7시에 알람 맞춰주고,
뉴스 들려줘"
914
920
925
제1 보이스 커맨드
제1 정보
930
제1 보이스 커맨드
935
제2 정보
942
940
"7시에 알람이 울릴 것입니다.
뉴스입니다. 오늘, S 회사는 ..."
944

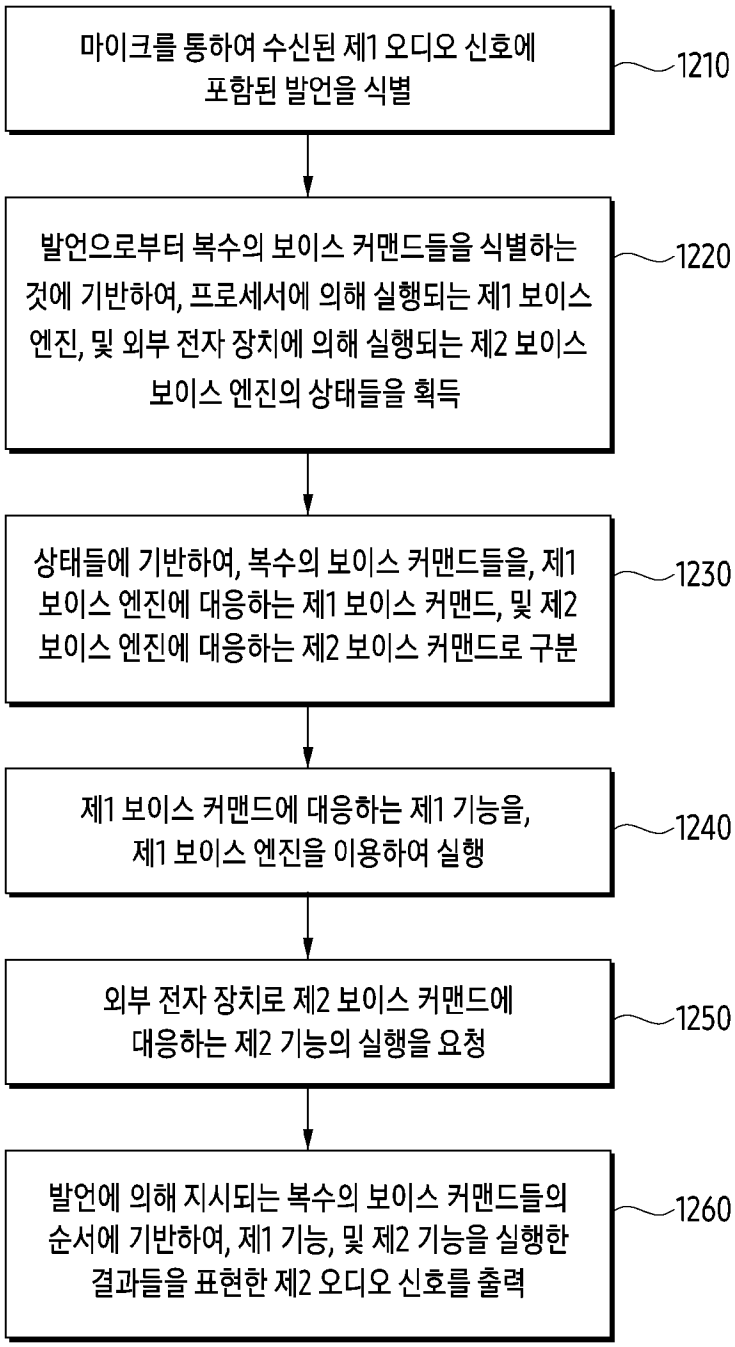
[도10]


1010
마이크를 통하여 제1 오디오 신호를 식별
1020
제1 오디오 신호로부터 복수의 기능들을 순차적으로 실행하기 위한 발언을 식별하였는가?
아니오
예
1030
프로세서에 기반하여 복수의 기능들 중 적어도 하나의 제1 기능을 실행
1040
통신 회로를 통해 연결된 외부 전자 장치를 이용하여, 복수의 기능들 중에서, 적어도 하나의 제1 기능과 상이한 적어도 하나의 제2 기능을 실행
1050
스피커를 통하여, 적어도 하나의 제1 기능 및 적어도 하나의 제2 기능을 실행한 결과들을 표현하는 제2 오디오 신호들을 출력
1060
프로세서, 또는 외부 전자 장치 중 어느 하나를 이용하여, 제1 오디오 신호에 대응하는 제3 기능을 실행

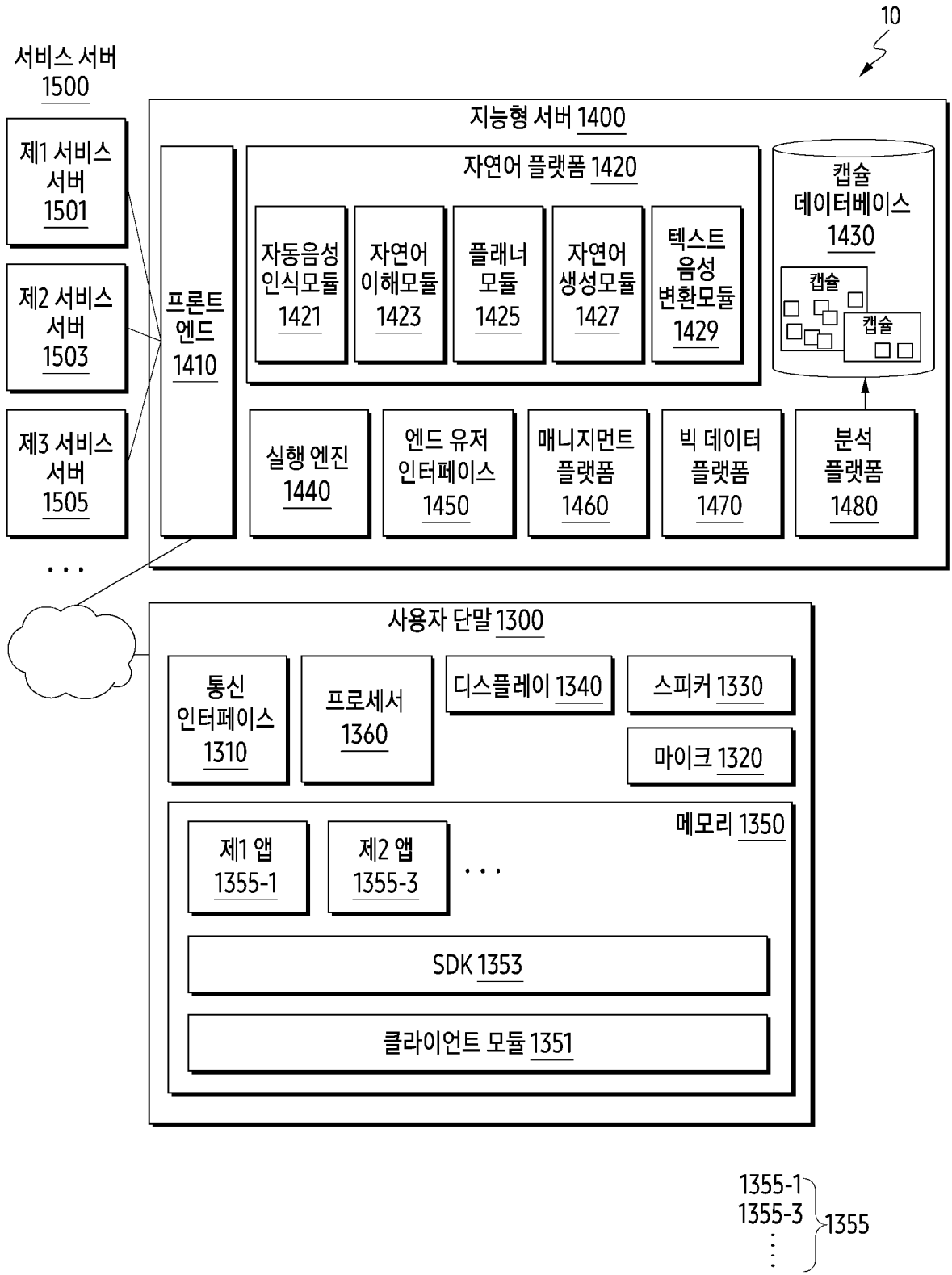
[도11]


1110
오디오 신호로부터 발언을 식별
1120
발언에 대응하는 메타 데이터를 식별하였는가?
아니오
예
1130
메타 데이터에 기반하여, 발언을 처리할 보이스 엔진을 선택
1140
발언으로부터, 발언에 대응하는 기능을 실행하기 위해 요구되는 파라미터를 식별하였는가?
아니오
예
1150
보이스 커맨드 생성
1160
파라미터를 요청

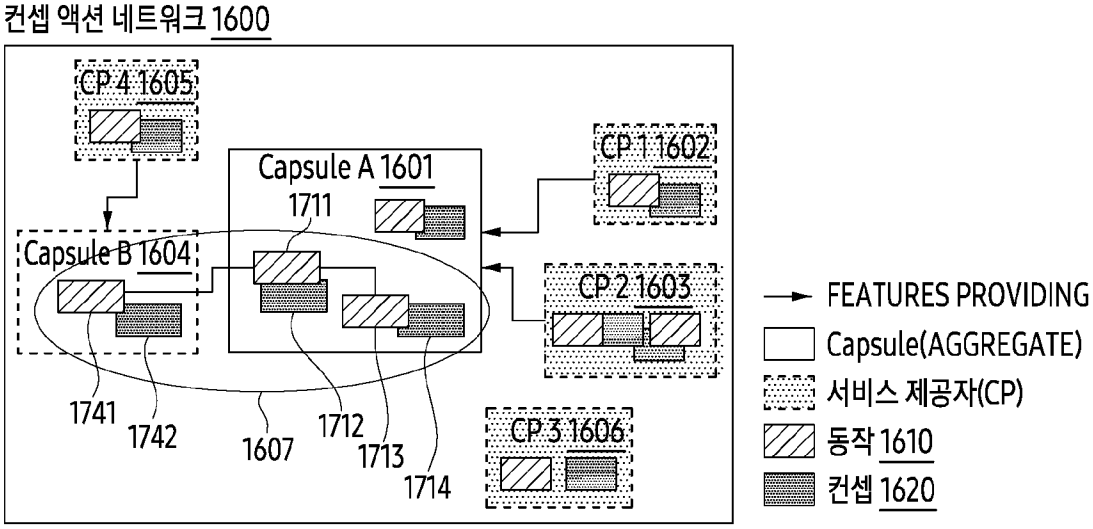
[도12]


마이크를 통하여 수신된 제1 오디오 신호에 포함된 발언을 식별
1210
발언으로부터 복수의 보이스 커맨드들을 식별하는 것에 기반하여, 프로세서에 의해 실행되는 제1 보이스 엔진, 및 외부 전자 장치에 의해 실행되는 제2 보이스 보이스 엔진의 상태들을 획득
1220
상태들에 기반하여, 복수의 보이스 커맨드들을, 제1 보이스 엔진에 대응하는 제1 보이스 커맨드, 및 제2 보이스 엔진에 대응하는 제2 보이스 커맨드로 구분
1230
제1 보이스 커맨드에 대응하는 제1 기능을, 제1 보이스 엔진을 이용하여 실행
1240
외부 전자 장치로 제2 보이스 커맨드에 대응하는 제2 기능의 실행을 요청
1250
발언에 의해 지시되는 복수의 보이스 커맨드들의 순서에 기반하여, 제1 기능, 및 제2 기능을 실행한 결과들을 표현한 제2 오디오 신호를 출력
1260

[도13]


10
서비스 서버
1500
제1 서비스 서버
1501
제2 서비스 서버
1503
제3 서비스 서버
1505
지능형 서버 1400
프론트 엔드
1410
자연어 플랫폼 1420
자동음성 인식모듈
1421
자연어 이해모듈
1423
플래너 모듈
1425
자연어 생성모듈
1427
텍스트 음성 변환모듈
1429
캡슐 데이터베이스
1430
캡슐
캡슐
실행 엔진
1440
엔드 유저 인터페이스
1450
매니지먼트 플랫폼
1460
빅 데이터 플랫폼
1470
분석 플랫폼
1480
사용자 단말 1300
통신 인터페이스
1310
프로세서
1360
디스플레이 1340
스피커 1330
마이크 1320
메모리 1350
제1 앱
1355-1
제2 앱
1355-3
SDK 1353
클라이언트 모듈 1351
1355-1
1355-3
1355

[도14]


컨셉 액션 네트워크 1600
CP 4 1605
Capsule A 1601
1711
CP 1 1602
Capsule B 1604
CP 2 1603
1741
1742
1607
1712
1713
1714
CP 3 1606
FEATURES PROVIDING
Capsule(AGGREGATE)
서비스 제공자(CP)
동작 1610
컨셉 1620

[도15]


1300
1510
NOV
1 2 3
4 5 6 7 8 9 10
11 12 13 14 15 16 17
18 19 20 21 22 23 24
25 26 27 28
1511
이번주 일정 알려줘!
1513
1300
1520
NOV

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2023/015158

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/16(2006.01)i; **G06F 3/0484**(2013.01)i; **G06F 3/0481**(2013.01)i; **G10L 15/02**(2006.01)i; **G10L 15/22**(2006.01)i; **G06F 3/048**(2006.01)i; **G10L 15/26**(2006.01)i; **G10L 25/51**(2013.01)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)

G06F 3/16(2006.01); B25J 11/00(2006.01); B25J 13/00(2006.01); G06F 3/041(2006.01); G10L 15/06(2006.01); G10L 15/16(2006.01); G10L 15/183(2013.01); G10L 15/26(2006.01); G10L 15/30(2013.01); G10L 15/34(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: 발화(utterance), 음성인식(voice recognition), 기능(function), 순차적(sequential), 동시 처리(parallel processing)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2018-0116726 A (SAMSUNG ELECTRONICS CO., LTD.) 25 October 2018 (2018-10-25) See paragraphs [0006]-[0040], [0101] and [0176]; and figures 2-6.	1-15
Y	KR 10-2026479 B1 (DIALOGUE DESIGN AGENCY INC.) 30 September 2019 (2019-09-30) See paragraphs [0035]-[0041]; and figures 1-2.	1-15
Y	KR 10-2017-0087207 A (SAMSUNG ELECTRONICS CO., LTD.) 28 July 2017 (2017-07-28) See paragraph [0041]; and figure 3.	4-6,8
A	KR 10-2019-0035454 A (KT CORPORATION) 03 April 2019 (2019-04-03) See paragraphs [0037]-[0044]; and figure 5.	1-15
A	US 2020-0126538 A1 (GOOGLE LLC) 23 April 2020 (2020-04-23) See paragraphs [0039] and [0076]; and figure 6.	1-15

☐ 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	Date of mailing of the international search report
22 January 2024	**22 January 2024**

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office **Government Complex-Daejeon Building 4, 189 Cheongsa-ro, Seo-gu, Daejeon 35208**	
Facsimile No. **+82-42-481-8578**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2023/015158

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
KR 10-2018-0116726 A	25 October 2018	KR 10-2369083 B1	02 March 2022
		US 2020-0075008 A1	05 March 2020
		WO 2018-194267 A1	25 October 2018
KR 10-2026479 B1	30 September 2019	None	
KR 10-2017-0087207 A	28 July 2017	KR 10-2392113 B1	29 April 2022
		US 10229681 B2	12 March 2019
		US 2017-0206900 A1	20 July 2017
KR 10-2019-0035454 A	03 April 2019	KR 10-2549204 B1	30 June 2023
US 2020-0126538 A1	23 April 2020	US 11145293 B2	12 October 2021
		US 11335333 B2	17 May 2022
		US 2020-0027444 A1	23 January 2020
		US 2022-0005465 A1	06 January 2022

국 제 조 사 보 고 서

국제출원번호
PCT/KR2023/015158

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06F 3/16(2006.01)i; **G06F 3/0484**(2013.01)i; **G06F 3/0481**(2013.01)i; **G10L 15/02**(2006.01)i; **G10L 15/22**(2006.01)i; **G06F 3/048**(2006.01)i; **G10L 15/26**(2006.01)i; **G10L 25/51**(2013.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06F 3/16(2006.01); B25J 11/00(2006.01); B25J 13/00(2006.01); G06F 3/041(2006.01); G10L 15/06(2006.01); G10L 15/16(2006.01); G10L 15/183(2013.01); G10L 15/26(2006.01); G10L 15/30(2013.01); G10L 15/34(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: 발화(utterance), 음성인식(voice recognition), 기능(function), 순차적(sequential), 동시 처리(parallel processing)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2018-0116726 A (삼성전자주식회사) 2018.10.25 단락 [0006]-[0040], [0101], [0176]; 및 도면 2-6	1-15
Y	KR 10-2026479 B1 (주식회사 다이얼로그디자인에이전시) 2019.09.30 단락 [0035]-[0041]; 및 도면 1-2	1-15
Y	KR 10-2017-0087207 A (삼성전자주식회사) 2017.07.28 단락 [0041]; 및 도면 3	4-6,8
A	KR 10-2019-0035454 A (주식회사 케이티) 2019.04.03 단락 [0037]-[0044]; 및 도면 5	1-15
A	US 2020-0126538 A1 (GOOGLE LLC) 2020.04.23 단락 [0039], [0076]; 및 도면 6	1-15

☐ 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 문헌과조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.
"&" 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일	국제조사보고서 발송일
2024년01월22일(22.01.2024)	**2024년01월22일(22.01.2024)**

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**

국 제 조 사 보 고 서
대응특허에 관한 정보

국제출원번호
PCT/KR2023/015158

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2018-0116726 A	2018/10/25	KR 10-2369083 B1	2022/03/02
		US 2020-0075008 A1	2020/03/05
		WO 2018-194267 A1	2018/10/25
KR 10-2026479 B1	2019/09/30	없음	
KR 10-2017-0087207 A	2017/07/28	KR 10-2392113 B1	2022/04/29
		US 10229681 B2	2019/03/12
		US 2017-0206900 A1	2017/07/20
KR 10-2019-0035454 A	2019/04/03	KR 10-2549204 B1	2023/06/30
US 2020-0126538 A1	2020/04/23	US 11145293 B2	2021/10/12
		US 11335333 B2	2022/05/17
		US 2020-0027444 A1	2020/01/23
		US 2022-0005465 A1	2022/01/06