	(19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2014-0089871 (43) 공개일자 2014년07월16일
(51) 국제특허분류(Int. Cl.) H04N 21/4227 (2011.01) H04N 21/23 (2011.01) G06F 3/16 (2006.01) (21) 출원번호 10-2013-0001829 (22) 출원일자 2013년01월07일 심사청구일자 없음	(71) 출원인 삼성전자주식회사 경기도 수원시 영통구 삼성로 129 (매탄동) (72) 발명자 정지혜 서울특별시 송파구 올림픽로 99 잠실엘스아파트 113-804 이청재 경기도 수원시 영통구 봉영로1517번길 73 벽적골 9단지아파트 905동 1501호 (뒷면에 계속) (74) 대리인 정홍식, 김태현, 이현수	

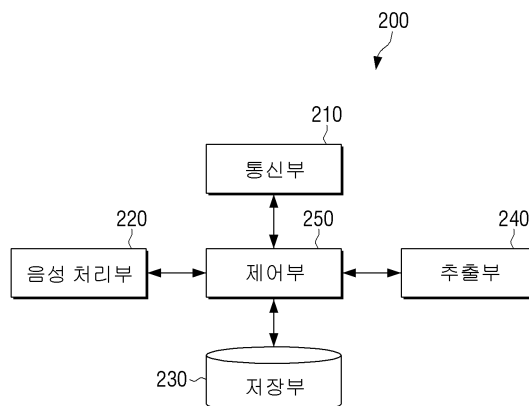
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 대화형 서버, 그 제어 방법 및 대화형 시스템

(57) 요약

대화형 서버, 그 제어 방법 및 대화형 시스템이 개시된다. 본 발명에 따른 대화형 서버는 디스플레이 장치와 통신을 수행하여 발화 음성 신호를 수신하는 통신부, 디스플레이 장치로부터 수신된 발화 음성 신호에 대한 발화 이력 정보를 저장하는 저장부, 수신된 발화 음성 신호로부터 발화 요소를 추출하는 추출부 및 저장부에 저장된 발화 이력 정보 및 추출된 발화 요소에 기초하여 응답 정보를 생성하고, 생성된 응답 정보를 디스플레이 장치로 전송하는 제어부를 포함한다. 이에 따라, 대화형 서버는 다양한 사용자의 발화 음성에 대해서 사용자의 의도를 파악하고, 그에 따른 응답 정보를 생성하여 디스플레이 장치로 제공할 수 있다.

대표도 - 도3



(72) 발명자

이혜정

서울특별시 송파구 강동대로 61-6

신용욱

경기도 오산시 경기대로 74 우림아파트 104동 130
3호

특허청구의 범위

청구항 1

디스플레이 장치와 통신을 수행하여 발화 음성 신호를 수신하는 통신부;

상기 발화 음성 신호가 수신되기 이전에 상기 디스플레이 장치로부터 수신된 발화 음성 신호에 대한 발화 이력 정보를 저장하는 저장부;

상기 수신된 발화 음성 신호로부터 발화 요소를 추출하는 추출부; 및

상기 저장부에 저장된 발화 이력 정보 및 상기 추출된 발화 요소에 기초하여 응답 정보를 생성하고, 상기 생성된 응답 정보를 디스플레이 장치로 전송하는 제어부;

를 포함하는 대화형 서버.

청구항 2

제 1 항에 있어서,

상기 제어부는,

상기 저장부에 상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 존재하면, 상기 발화 요소에 대응되는 실행어 및 명령어를 조합하여 상기 응답 정보를 생성하고,

상기 저장부에 상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 상기 저장부에 저장된 발화 이력 정보에 기초하여 상기 추출된 발화 요소에 대응되는 실행어 및 명령어를 추정하고, 추정된 실행어 및 명령어를 조합하여 상기 응답 정보를 생성하는 것을 특징으로 하는 대화형 서버.

청구항 3

제 2 항에 있어서,

상기 발화 음성 신호는 제1 발화 요소 및 제2 발화 요소를 포함하며,

상기 제어부는,

상기 저장부에 상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 상기 발화 이력 정보 중에서 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보를 획득하고, 상기 획득된 발화 이력 정보 내에서 상기 제1 발화 요소에 대응되는 실행어 및 상기 제2 발화 요소에 대응되는 명령어를 상기 발화 음성 신호에 대응되는 실행어 및 명령어로 추정하는 것을 특징으로 하는 대화형 서버.

청구항 4

제 3 항에 있어서,

상기 제어부는,

상기 저장부에 저장된 발화 이력 정보 중 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보가 저장되어 있지 않으면, 상기 응답 정보를 생성하지 않는 것을 특징으로 하는 대화형 서버.

청구항 5

제 3 항에 있어서,

상기 제어부는,

상기 발화 음성 신호 내에 제1 발화 요소 및 제2 발화 요소 중 적어도 하나가 포함되어 있지 않으면, 상기 발화 이력 정보 중에서 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보에 기초하여 상기 발화 음성 신호에 대응되는 응답 정보를 생성하는 것을 특징으로 하는 대화형 서버.

청구항 6

제 1 항에 있어서,
상기 수신된 발화 음성 신호를 텍스트 정보로 변환하는 음성 처리부;
를 더 포함하는 것을 특징으로 하는 대화형 서버.

청구항 7

대화형 서버의 제어 방법에 있어서,
디스플레이 장치로부터 발화 음성 신호를 수신하는 단계;
상기 수신된 발화 음성 신호로부터 발화 요소를 추출하는 단계;
기저장된 발화 이력 정보 및 상기 추출된 발화 요소에 기초하여 응답 정보를 생성하는 단계; 및
상기 생성된 응답 정보를 상기 디스플레이 장치로 전송하는 단계;를 포함하며,
상기 기저장된 발화 이력 정보는
상기 발화 음성 신호가 수신되기 이전에 상기 디스플레이 장치로부터 수신된 발화 음성 신호에 기초하여 생성된
정보인 것을 특징으로 하는 제어 방법.

청구항 8

제 7 항에 있어서,
상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 존재하는지 여부를 판단하는 단계;를 더 포함하며,
상기 생성하는 단계는,
상기 발화 요소에 대응되는 실행어 및 명령어가 존재하면, 상기 발화 요소에 대응되는 실행어 및 명령어를 조합
하여 응답 정보를 생성하고,
상기 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 상기 기저장된 발화 이력 정보에 기초하여 상기 추
출된 발화 요소에 대응되는 실행어 및 명령어를 추정하고, 상기 추정된 실행어 및 명령어를 조합하여 상기 응답
정보를 생성하는 것을 특징으로 하는 제어 방법.

청구항 9

제 8 항에 있어서,
상기 발화 음성 신호는 제1 발화 요소 및 제2 발화 요소를 포함하며,
상기 생성하는 단계는,
상기 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 상기 발화 이력 정보 중에서 상기 발화 음성 신호
가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보를 획득하고, 상기 획득된 발화 이력 정
보 내에서 상기 제1 발화 요소에 대응되는 실행어 및 상기 제2 발화 요소에 대응되는 명령어를 상기 발화 음성
신호에 대응되는 실행어 및 명령어로 추정하는 것을 특징으로 하는 제어 방법

청구항 10

제 9 항에 있어서,
상기 생성하는 단계는,
상기 발화 이력 정보 중 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력
정보가 저장되어 있지 않으면, 상기 응답 정보를 생성하지 않는 것을 특징으로 하는 제어 방법.

청구항 11

제 9 항에 있어서,

상기 생성하는 단계는,

상기 발화 음성 신호 내에 제1 발화 요소 및 제2 발화 요소 중 적어도 하나가 포함되어 있지 않으면, 상기 발화 이력 정보 중에서 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보에 기초하여 상기 발화 음성 신호에 대응되는 응답 정보를 생성하는 것을 특징으로 하는 제어 방법.

청구항 12

제 7 항에 있어서,

상기 수신된 발화 음성 신호를 텍스트 정보로 변환하는 단계;

를 더 포함하는 것을 특징으로 하는 대화형 서버.

청구항 13

대화형 시스템에 있어서,

사용자의 발화 음성이 입력되면, 입력된 발화 음성에 대응되는 기능을 수행하는 디스플레이 장치; 및

상기 디스플레이 장치로부터 발화 음성 신호가 수신되면, 상기 발화 음성 신호로부터 발화 요소를 추출하고, 기 저장된 발화 이력 정보 및 상기 추출된 발화 요소에 기초하여 응답 정보를 생성하여 상기 디스플레이 장치로 전송하는 대화형 서버;를 포함하며,

상기 디스플레이 장치는,

상기 대화형 서버로부터 수신한 응답 정보에 기초하여 상기 사용자 의도에 대응되는 기능을 수행하며,

상기 기저장된 발화 이력 정보는 상기 발화 음성 신호가 수신되기 이전에 상기 디스플레이 장치로부터 수신된 발화 음성 신호에 기초하여 생성된 정보인 것을 특징으로 하는 대화형 시스템.

청구항 14

제 13 항에 있어서,

상기 대화형 서버는,

상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 존재하면, 상기 발화 요소에 대응되는 실행어 및 명령어에 기초하여 상기 응답 정보를 생성하고,

상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 기저장된 발화 이력 정보에 기초하여 상기 추출된 발화 요소에 대응되는 실행어 및 명령어를 추정하고, 추정된 실행어 및 명령어를 조합하여 상기 응답 정보를 생성하는 것을 특징으로 하는 대화형 시스템.

명세서

기술분야

[0001] 본 발명은 대화형 서버, 그 제어 방법 및 대화형 시스템에 관한 것으로서, 보다 상세하게는 사용자의 다양한 발화에 대응되는 응답 정보를 제공하기 위한 대화형 서버, 그 제어 방법 및 대화형 시스템에 관한 것이다.

배경기술

[0002] 일반적으로, 대화형 시스템에서 음성 인식이 가능한 디스플레이 장치는 사용자의 발화 음성을 수집하고, 그 수집한 발화 음성을 네트워크를 통해 연결된 외부 서버로 전송한다. 발화 음성을 수신한 외부 서버는 발화 음성을 분석하여 사용자의 의도를 파악하고, 그에 따른 응답 정보를 생성하여 디스플레이 장치로 전송한다. 이에 따라, 디스플레이 장치는 외부 서버로부터 수신한 응답 정보에 기초하여 사용자의 발화 음성에 대응하는 기능을 실행하거나 정보를 제공할 수 있다.

[0003] 그러나, 이 같은 종래의 대화형 시스템은 사용자의 발화 음성을 분석하고, 그 분석 결과에 기초하여 사용자의 의도를 파악하는데 한계가 있다. 예를 들어, "제1 콘텐츠 보여줘"와 같이 지칭하는 대상이 명확한 발화 음성의

경우, 외부 서버는 이 같은 발화 음성을 분석하여 사용자의 의도를 올바르게 파악하고, 그에 따른 응답 정보를 생성하여 디스플레이 장치로 전송할 수 있다. 따라서, 디스플레이 장치는 응답 정보에 기초하여 사용자가 요청한 제1 콘텐츠를 디스플레이할 수 있다.

[0004] 그러나, "이거 보여줘"와 같이 사용자의 발화가 모호한 경우, 외부 서버는 이 같은 발화 음성으로부터 사용자의 의도를 명확하게 파악하지 못하는 문제가 있다. 다시 말해, 종래의 대화형 시스템은 기정의된 발화 음성에 대해서만 사용자의 의도를 파악하고 그에 따른 동작을 수행하거나 정보를 제공하기 때문에 사용자의 발화에 제약이 따르는 문제가 있다.

발명의 내용

[0005] 본 발명은 상술한 필요성에 따라 안출된 것으로, 본 발명의 목적은, 대화형 시스템에서 다양한 사용자의 발화에 대응되는 동작을 수행하도록 함을 목적으로 한다.

[0006] 이상과 같은 목적을 달성하기 위한 본 발명의 일 실시 예에 따른 대화형 서버는 디스플레이 장치와 통신을 수행하여 발화 음성 신호를 수신하는 통신부, 상기 디스플레이 장치로부터 수신된 발화 음성 신호에 대한 발화 이력 정보를 저장하는 저장부, 상기 수신된 발화 음성 신호로부터 발화 요소를 추출하는 추출부 및 상기 저장부에 저장된 발화 이력 정보 및 상기 추출된 발화 요소에 기초하여 응답 정보를 생성하고, 상기 생성된 응답 정보를 디스플레이 장치로 전송하는 제어부를 포함한다.

[0007] 그리고, 상기 제어부는, 상기 저장부에 상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 존재하면, 상기 발화 요소에 대응되는 실행어 및 명령어를 조합하여 상기 응답 정보를 생성하고, 상기 저장부에 상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 상기 저장부에 저장된 발화 이력 정보에 기초하여 상기 추출된 발화 요소에 대응되는 실행어 및 명령어를 추정하고, 추정된 실행어 및 명령어를 조합하여 상기 응답 정보를 생성할 수 있다.

[0008] 또한, 상기 발화 음성 신호는 제1 발화 요소 및 제2 발화 요소를 포함하며, 상기 제어부는, 상기 저장부에 상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 상기 발화 이력 정보 중에서 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보를 획득하고, 상기 획득된 발화 이력 정보 내에서 상기 제1 발화 요소에 대응되는 실행어 및 상기 제2 발화 요소에 대응되는 명령어를 상기 발화 음성 신호에 대응되는 실행어 및 명령어로 추정할 수 있다.

[0009] 그리고, 상기 제어부는, 상기 저장부에 저장된 발화 이력 정보 중 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보가 저장되어 있지 않으면, 상기 응답 정보를 생성하지 않을 수 있다.

[0010] 또한, 상기 제어부는, 상기 발화 음성 신호 내에 제1 발화 요소 및 제2 발화 요소 중 적어도 하나가 포함되어 있지 않으면, 상기 발화 이력 정보 중에서 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보에 기초하여 상기 발화 음성 신호에 대응되는 응답 정보를 생성할 수 있다.

[0011] 그리고, 상기 수신된 발화 음성 신호를 텍스트 정보로 변환하는 음성 처리부를 더 포함할 수 있다.

[0012] 한편, 본 발명의 일 실시 예에 따르면, 대화형 서버의 제어 방법에 있어서, 상기 방법은 디스플레이 장치로부터 발화 음성 신호를 수신하는 단계, 상기 수신된 발화 음성 신호로부터 발화 요소를 추출하는 단계, 상기 디스플레이 장치로부터 수신된 발화 음성 신호와 관련하여 기저장된 발화 이력 정보 및 상기 추출된 발화 요소에 기초하여 응답 정보를 생성하는 단계 및 상기 생성된 응답 정보를 상기 디스플레이 장치로 전송하는 단계를 포함한다.

[0013] 그리고, 상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 존재하는지 여부를 판단하는 단계를 더 포함하며, 상기 생성하는 단계는, 상기 발화 요소에 대응되는 실행어 및 명령어가 존재하면, 상기 발화 요소에 대응되는 실행어 및 명령어를 조합하여 응답 정보를 생성하고, 상기 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 상기 기저장된 발화 이력 정보에 기초하여 상기 추출된 발화 요소에 대응되는 실행어 및 명령어를 추정하고, 상기 추정된 실행어 및 명령어를 조합하여 상기 응답 정보를 생성할 수 있다.

[0014] 또한, 상기 발화 음성 신호는 제1 발화 요소 및 제2 발화 요소를 포함하며, 상기 생성하는 단계는, 상기 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 상기 발화 이력 정보 중에서 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보를 획득하고, 상기 획득된 발화 이력 정보 내에서

상기 제1 발화 요소에 대응되는 실행어 및 상기 제2 발화 요소에 대응되는 명령어를 상기 발화 음성 신호에 대응되는 실행어 및 명령어로 추정할 수 있다.

- [0015] 그리고, 상기 생성하는 단계는, 상기 발화 이력 정보 중 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보가 저장되어 있지 않으면, 상기 응답 정보를 생성하지 않을 수 있다.
- [0016] 또한, 상기 생성하는 단계는, 상기 발화 음성 신호 내에 제1 발화 요소 및 제2 발화 요소 중 적어도 하나가 포함되어 있지 않으면, 상기 발화 이력 정보 중에서 상기 발화 음성 신호가 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보에 기초하여 상기 발화 음성 신호에 대응되는 응답 정보를 생성할 수 있다.
- [0017] 그리고, 상기 수신된 발화 음성 신호를 텍스트 정보로 변환하는 단계를 더 포함할 수 있다.
- [0018] 한편, 본 발명의 일 실시 예에 따르면, 대화형 시스템은 사용자의 발화 음성이 입력되면, 입력된 발화 음성에 대응되는 기능을 수행하는 디스플레이 장치 및 상기 디스플레이 장치로부터 발화 음성 신호가 수신되면, 상기 발화 음성 신호로부터 발화 요소를 추출하고, 기저장된 발화 이력 정보 및 상기 추출된 발화 요소에 기초하여 응답 정보를 생성하여 상기 디스플레이 장치로 전송하는 대화형 서버;를 포함하며, 상기 디스플레이 장치는, 상기 대화형 서버로부터 수신한 응답 정보에 기초하여 상기 사용자 의도에 대응되는 기능을 수행할 수 있다.
- [0019] 그리고, 상기 대화형 서버는, 상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 존재하면, 상기 발화 요소에 대응되는 실행어 및 명령어에 기초하여 상기 응답 정보를 생성하고, 상기 추출된 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 기저장된 발화 이력 정보에 기초하여 상기 추출된 발화요소에 대응되는 실행어 및 명령어를 추정하고, 추정된 실행어 및 명령어를 조합하여 상기 응답 정보를 생성할 수 있다.
- [0020] 이상과 같이 본 발명의 다양한 실시 예에 따르면, 대화형 시스템에서 대화형 서버는 다양한 사용자의 발화 음성에 대해서 사용자의 의도를 파악하고, 그에 따른 응답 정보를 생성하여 디스플레이 장치로 제공할 수 있다.

도면의 간단한 설명

- [0021] 도 1은 본 발명의 일 실시예에 따른 사용자 발화 음성에 적합한 응답 정보를 제공하는 대화형 시스템의 제1 예시도,
- 도 2는 본 발명의 또다른 실시예에 따른 사용자 발화 음성에 적합한 응답 정보를 제공하는 대화형 시스템의 제2 예시도,
- 도 3은 본 발명의 일 실시예에 따른 대화형 서버의 블록도,
- 도 4는 본 발명의 일 실시예에 따른 디스플레이 장치에서 화면상에 표시된 객체의 표시 상태에 기초하여 발화 정보를 생성하는 예시도,
- 도 5는 본 발명의 일 실시예에 따른 디스플레이 장치의 블록도,
- 도 6은 본 발명의 일 실시예에 따른 대화형 서버의 제어 방법에 대한 흐름도,
- 도 7은 본 발명의 일 실시예에 따른 대화형 서버에서 응답 정보를 생성하는 방법의 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0022] 이하 첨부된 도면들을 참조하여 본 발명의 실시 예를 보다 상세하게 설명한다.
- [0023] 도 1은 본 발명의 일 실시예에 따른 사용자 발화 음성에 적합한 응답 정보를 제공하는 대화형 시스템의 제1 예시도이다.
- [0024] 도 1에 도시된 바와 같이, 본 발명의 일 실시 예에 따른 대화형 시스템은 디스플레이 장치(100) 및 대화형 서버(200)를 포함한다.
- [0025] 디스플레이 장치(100)는 인터넷이 가능한 장치로써, 스마트 TV, 스마트폰과 같은 휴대폰, 데스크탑 PC, 노트북, 네비게이션 등과 같은 다양한 전자 장치로 구현될 수 있다. 이 같은 디스플레이 장치(100)는 사용자의 발화 음성이 입력되면, 입력된 사용자의 발화 음성에 대응되는 동작을 수행한다. 구체적으로, 디스플레이 장치(100)는 사용자의 발화 음성에 대응되는 응답 메시지를 출력하거나 또는 사용자의 발화 음성에 대응되는 기능을 수행할 수 있다. 즉, 디스플레이 장치(100)는 입력된 발화 음성을 분석하여 내부적으로 수행가능한지 여부를 판단하고, 그 판단 결과에 따라, 입력된 발화 음성에 대응되는 기능을 수행하거나 또는 대화형 서버(200)로부터

수신한 응답 정보에 기초하여 동작을 수행할 수 있다.

- [0026] 예를 들어, "볼륨 올려"라는 발화 음성이 입력되면, 디스플레이 장치(100)는 기저장된 제어 정보 중 입력된 발화 음성에 대응되는 제어 정보에 기초하여 볼륨을 조정할 수 있다.
- [0027] 또다른 예를 들어, "오늘 날씨 어때?"라는 발화 음성이 입력되면, 디스플레이 장치(100)는 입력된 발화 음성에 대한 발화 음성 신호를 대화형 서버(200)로 전송한다. 여기서, 발화 음성은 아날로그 신호가 될 수 있다. 따라서, 디스플레이 장치(100)는 아날로그 신호인 발화 음성을 디지털 신호(이하 발화 음성이라 함)로 변환하여 대화형 서버(200)로 전송한다. 이후, 디스플레이 장치(100)는 대화형 서버(200)로부터 수신한 응답 정보에 기초하여 오늘 날씨에 대한 결과를 음성 또는 텍스트 형태의 영상으로 출력할 수 있다.
- [0028] 또다른 예를 들어, "○○○(컨텐츠명) 보여줘"라는 발화 음성이 입력되면, 디스플레이 장치(100)는 입력된 발화 음성을 대화형 서버(200)로 전송한다. 이후, 대화형 서버(200)로부터 응답 정보가 수신되면, 디스플레이 장치(100)는 수신한 응답 정보에 기초하여 ○○○(컨텐츠명)에 대한 검색 리스트를 디스플레이할 수 있다. 여기서, 검색 리스트는 ○○○(컨텐츠명)에 대한 회차별 컨텐츠 정보를 포함할 수 있다. 따라서, 사용자는 검색 리스트 상에 표시된 회차별 컨텐츠 정보를 확인하여 자신이 시청하고자하는 회차의 ○○○(컨텐츠명)를 선택하기 위한 발화를 할 수 있다. 예를 들어, 사용자는 검색 리스트 상에 표시된 회차별 컨텐츠 정보 중 첫번째 컨텐츠 정보에 해당하는 ○○○(컨텐츠명)를 시청하기 위하여 "첫번째 거 보여줘"라는 발화를 할 수 있다. 따라서, 디스플레이 장치(100)는 "첫번째 거 보여줘"라는 발화 음성을 입력받아 이를 대화형 서버(200)로 전송한다. 이후, 대화형 서버(200)로부터 수신된 발화 음성에 대응되는 응답 정보가 수신되면, 디스플레이 장치(100)는 수신된 응답 정보에 기초하여 첫번째 컨텐츠 정보에 해당하는 ○○○(컨텐츠명)를 컨텐츠 제공 서버(300)로부터 수신하여 디스플레이할 수 있다.
- [0029] 대화형 서버(200)는 디스플레이 장치(100)로부터 수신한 사용자의 발화 음성에 기초하여 사용자의 의도에 적합한 응답 정보를 제공한다. 구체적으로, 대화형 서버(200)는 디스플레이 장치(100)로 사용자의 발화 음성이 수신되면, 수신된 발화 음성으로부터 발화 요소를 추출하고, 그 추출된 발화 요소에 기초하여 사용자의 발화 음성 과 관련된 응답 정보를 생성하여 전송할 수 있다. 전술한 바와 같이, 디스플레이 장치(100)로부터 수신한 사용자의 발화 음성은 디지털 신호가 될 수 있다. 따라서, 대화형 서버(200)는 디지털 신호로 변환된 발화 음성이 수신되면, 이를 텍스트 정보로 생성한 후, 생성된 텍스트 정보를 분석하여 발화 요소를 추출하고, 추출된 발화 요소에 기초하여 사용자의 발화 음성에 대응되는 응답 정보를 생성할 수 있다. 디지털 신호로 변환된 발화 음성으로부터 텍스트 정보를 생성하는 것은 공지된 기술이기에 본 발명에서는 상세한 설명을 생략하도록 한다.
- [0030] 한편, 발화 요소는 사용자의 발화 음성 내에서 사용자가 요청한 동작을 수행하기 위한 핵심 키워드로써, 목적 영역(domain), 목적 기능(user action) 및 주요 특징(feature) 별로 추출된 발화 요소를 분류할 수 있다. 전술한 예와 같이, "오늘 날씨 어때?"라는 사용자의 발화 음성에 대한 텍스트 정보가 생성되면, 대화형 서버(200)는 "오늘", "날씨", "어때?"라는 발화 요소를 추출할 수 있다. 이후, 대화형 서버(200)는 추출된 발화 요소 중 "오늘" 및 "날씨"를 주요 특징(이하 실행 대상이라 함)에 대한 발화 요소(이하 제1 발화 요소라 함)로 분류하고, "어때?"를 목적 기능(이하 실행 명령이라 함)에 대한 발화 요소(이하 제2 발화 요소라 함)로 분류할 수 있다. 또한, 대화형 서버(200)는 추출된 발화 요소에 기초하여 목적 영역에 대한 발화 요소(이하 제3 발화 요소라 함)가 웹 서치라는 도메인에 속하는 것으로 분류할 수 있다. 이 같이, 사용자의 발화 음성에 대한 텍스트 정보로부터 제1 내지 제3 발화 요소가 분류되면, 대화형 서버(200)는 다양한 컨텐츠를 제공하는 외부 서버(400)로부터 날씨 정보를 제공받아, 이를 포함하는 응답 정보를 생성하여 디스플레이 장치(100)로 전송한다. 이에 따라, 디스플레이 장치(100)는 대화형 서버(200)로부터 수신한 응답 정보에 기초하여 오늘 날씨 정보를 음성 및 텍스트 중 적어도 하나를 통해 디스플레이할 수 있다.
- [0031] 한편, 대화형 서버(200)는 "첫번째 거 선택해줘"라는 발화 음성에 대한 텍스트 정보가 생성되면, 생성된 텍스트 정보로부터 "첫번째"라는 제1 발화 요소 및 "선택해줘"라는 제2 발화 요소를 추출할 수 있다. 이 같은 제1 및 제2 발화 요소가 추출되면, 대화형 서버(200)는 추출된 제1 및 제2 발화 요소에 기초하여 사용자 의도에 적합한 응답 정보를 생성할 수 있는지 여부를 판단한다. 실시예에 따라, 대화형 서버(200)는 추출된 제1 및 제2 발화 요소에 대응되는 정보가 데이터베이스 내에 존재하는지 여부를 따라 사용자 의도에 적합한 응답 정보의 생성 여부를 판단할 수 있다. 여기서, 제1 및 제2 발화 요소에 대응되는 정보는 제1 발화 요소에 대응되는 실행어 및 제2 발화 요소에 대응되는 명령어가 될 수 있다. 따라서, 대화형 서버(200)는 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 데이터베이스 내에 존재하는지 여부를 판단한다. 판단 결과, 추출된 제1 및 제2 발화 요소에 대응되는 정보가 데이터베이스 내에 존재하지 않으면, 대화형 서버(200)는 기저장된 발화 이력 정

보 중 "첫번째 거 선택해줘"라는 발화 음성이 수신되기 직전의 발화 이력 정보와 추출된 제1 및 제2 발화 요소에 기초하여 응답 정보를 생성하고, 이를 디스플레이 장치(100)로 전송할 수 있다. 예를 들어, "첫번째 거 선택해줘"라는 발화 음성이 수신되기 직전에 "○○○(컨텐츠명) 찾아줘"라는 발화 음성이 수신될 수 있으며, 이 같은 발화 음성에 대한 발화 이력 정보가 저장될 수 있다. 따라서, 대화형 서버(200)는 이전 발화 이력 정보에 기초하여 "○○○(컨텐츠명)"에 대한 컨텐츠 실행 관련 응답 정보를 생성하여 디스플레이 장치(100)로 전송할 수 있다. 따라서, 디스플레이 장치(100)는 대화형 서버(200)로부터 수신한 응답 정보에 기초하여 "○○○(컨텐츠명)"에 대한 첫번째 컨텐츠를 디스플레이할 수 있다.

[0032] 한편, 전술한 대화형 서버(200)는 디지털 신호로 변환된 사용자의 발화 음성을 텍스트 정보로 생성하는 제1 서버(10) 및 텍스트 정보로 생성된 발화 음성에 대응하는 응답 정보를 생성하는 제2 서버(20)를 포함할 수 있다. 이하에서는, 디스플레이 장치(100), 제1 및 제2 서버(10,20)를 통해 사용자의 발화 음성에 적합한 응답 정보를 제공하는 대화형 시스템에 대해서 상세히 설명하도록 한다.

[0033] 도 2는 본 발명의 또다른 실시예에 따른 사용자 발화 음성에 적합한 응답 정보를 제공하는 대화형 시스템의 제2 예시도이다.

[0034] 도 2에 도시된 바와 같이, 디스플레이 장치(100)는 사용자로부터 발화된 발화 음성이 입력되면, 입력된 발화 음성을 디지털 신호로 변환하여 제1 서버(10)로 전송한다. 디지털 신호로 변환된 발화 음성이 수신되면, 제1 서버(10)는 기저장된 다양한 발화 음성에 대한 특정 패턴에 따라, 사용자의 발화 음성에 대한 텍스트 정보를 생성하여 디스플레이 장치(100)로 전송한다.

[0035] 제1 서버(10)로부터 사용자의 발화 음성에 대한 텍스트 정보를 수신한 디스플레이 장치(100)는 사용자의 발화 음성에 대한 텍스트 정보를 제2 서버(20)로 전송한다. 사용자의 발화 음성에 대한 텍스트 정보를 수신한 제2 서버(20)는 수신한 텍스트 정보를 분석하여 발화 요소를 추출하고, 그 추출된 발화 요소에 기초하여 사용자의 발화 음성에 대응되는 동작을 수행하기 위한 응답 정보를 생성하여 디스플레이 장치(100)로 전송할 수 있다.

[0036] 지금까지, 본 발명에 따른 디스플레이 장치(100)와 대화형 서버(200)로 이루어진 대화형 시스템에서 사용자의 발화 음성에 대응하는 응답 정보를 제공하는 동작에 대해서 개략적으로 설명하였다. 이하에서는, 전술한 디스플레이 장치(100) 및 대화형 서버(200)의 각 구성에 대해서 상세히 설명하도록 한다.

[0037] 도 3은 본 발명의 일 실시예에 따른 대화형 서버의 블록도이다.

[0038] 도 3에 도시된 바와 같이, 대화형 서버(200)는 통신부(210), 음성 처리부(220), 저장부(230), 추출부(240) 및 제어부(250)를 포함한다.

[0039] 통신부(210)는 사용자의 발화 음성을 제공하는 디스플레이 장치(100)와 통신을 수행한다. 특히, 통신부(210)는 실행 대상을 나타내는 제1 발화 요소 및 실행 명령을 나타내는 제2 발화 요소를 포함하는 발화 음성에 대한 디지털 신호(이하 발화 음성이라 함)를 수신할 수 있다. 여기서, 실행 명령을 나타내는 제1 발화 요소는 사용자의 발화 음성 내에서 주요 특징(feature)으로 분류되는 발화 요소이며, 실행 명령을 나타내는 제2 발화 요소는 사용자의 발화 음성 내에서 목적 기능(user action)으로 분류되는 발화 요소이다. 예를 들어, "액션 영화 찾아줘"라는 발화 음성의 경우, "액션"은 주요 특징으로 분류되는 제1 발화 요소이며, "찾아줘"는 목적 기능으로 분류되는 제2 발화 요소가 될 수 있다.

[0040] 이 같은 제1 및 제2 발화 요소를 포함하는 발화 음성이 수신되면, 음성 처리부(220)는 수신된 발화 음성을 텍스트 정보로 변환한다. 실시예에 따라, 음성 처리부(220)는 STT(Speech to Text) 알고리즘을 이용하여 수신된 사용자의 발화 음성을 텍스트로 변환할 수 있다. 그러나, 본 발명은 이에 한정되지 않으며, 통신부(210)는 디스플레이 장치(100)로부터 사용자의 발화 음성에 대한 텍스트 정보를 수신할 수 있다. 이 경우, 디스플레이 장치(100)는 입력된 사용자의 발화 음성에 대한 텍스트 정보를 전술한 제1 서버(10)와 같은 음성 인식 장치(미도시)로부터 수신하여 대화형 서버(200)로 전송한다. 따라서, 통신부(210)는 디스플레이 장치(100)로부터 사용자의 발화 음성에 대한 텍스트 정보를 수신할 수 있다. 이 같이, 디스플레이 장치(100)로부터 사용자의 발화 음성에 대한 텍스트 정보를 수신할 경우, 전술한 음성 처리부(220)의 구성은 생략될 수 있다.

[0041] 한편, 사용자의 발화 음성이 텍스트 정보로 변환되거나 혹은 디스플레이 장치(100)로부터 사용자의 발화 음성에 대한 텍스트 정보가 수신되면, 추출부(240)는 텍스트 정보로 변환된 발화 음성으로부터 발화 요소를 추출한다. 전술한 바와 같이, 추출부(240)는 텍스트 정보로 변환된 발화 음성으로부터 주요 특징을 나타내는 제1 발화 요소 및 실행 명령을 나타내는 제2 발화 요소를 추출할 수 있다.

- [0042] 진술한 예와 같이, "액션 영화 찾아줘"라는 발화 음성의 경우, 추출부(240)는 "액션" 및 "영화"라는 제1 발화 요소와 "찾아줘"라는 제2 발화 요소를 추출할 수 있다. 또다른 예를 들어, "이거 실행해줘"라는 발화 음성의 경우, 추출부(240)는 "이거"라는 제1 발화 요소 및 "실행해줘"라는 제2 발화 요소를 추출할 수 있다.
- [0043] 이 같은 발화 요소가 추출되면, 제어부(250)는 저장부(230)에 저장된 발화 이력 정보 및 추출된 발화 요소에 기초하여 응답 정보를 생성하고, 생성된 응답 정보를 디스플레이 장치(100)로 전송한다. 여기서, 저장부(230)는 대화형 서버(200)를 동작시키기 위해 필요한 각종 프로그램 등이 저장되는 저장매체로써, 메모리, HDD(Hard Disk Drive) 등으로 구현될 수 있다. 예를 들어, 저장부(230)는 제어부(250)의 동작을 수행하기 위한 프로그램을 저장하는 ROM, 제어부(250)의 동작 수행에 따른 데이터를 일시적으로 저장하는 RAM 등을 구비할 수 있다. 뿐만 아니라, 저장부(230)는 각종 참조 데이터를 저장하는 EEROM(Electrically Erasable and Programmable ROM) 등을 더 구비할 수 있다. 특히, 저장부(230)는 디스플레이 장치(100)로부터 수신된 발화 음성에 대한 발화 이력 정보를 저장할 수 있다. 여기서, 발화 이력 정보는 디스플레이 장치(100)로부터 현재 발화 음성이 수신되기 이전에 수신된 발화 음성에 기초하여 생성된 정보이다. 구체적으로, 저장부(230)는 디스플레이 장치(100)로부터 수신된 발화 음성으로부터 생성된 텍스트 정보, 해당 발화 음성이 수신된 시간 정보, 해당 발화 음성에 기초하여 생성된 응답 정보 및 해당 응답 정보가 생성된 시간 정보 중 적어도 하나를 포함하는 발화 이력 정보를 저장할 수 있다.
- [0044] 따라서, 제어부(250)는 저장부(230)에 저장된 이전 발화 이력 정보 중 현재 발화 음성이 수신되기 이전의 발화 이력 정보와 수신된 발화 음성으로부터 추출된 발화 요소 중 적어도 하나에 기초하여 응답 정보를 생성하고, 이를 디스플레이 장치(100)로 전송할 수 있다. 구체적으로, 제어부(250)는 추출부(240)로부터 추출된 발화 요소에 대응되는 정보가 저장부(230)에 존재하는지 여부를 판단하여, 추출된 발화 요소에 대응되는 정보가 존재하면 추출된 발화 요소에 대응되는 정보에 기초하여 응답 정보를 생성한다. 여기서, 추출된 발화 요소에 대응되는 정보는 실행 대상을 나타내는 제1 발화 요소에 대응되는 실행어 및 실행 명령을 나타내는 제2 발화 요소에 대응되는 명령어 중 적어도 하나를 포함할 수 있다. 이 같은 실행어 및 명령어는 저장부(230)에 저장될 수 있다. 즉, 저장부(230)는 제1 발화 요소별로 정의된 실행어 및 제2 발화 요소별로 정의된 명령어를 더 저장할 수 있다. 예를 들어, "이거"라는 실행 대상을 간접적으로 나타내는 제1 발화 요소는 "\$this\$"라는 실행어로 정의될 수 있으며, "다음"이라는 실행 대상을 간접적으로 나타내는 제1 발화 요소는 "\$this\$"라는 실행어로 정의될 수 있으며, "세번째"라는 실행 대상을 간접적으로 나타내는 제1 발화 요소는 "\$3rd\$"라는 실행어로 정의될 수 있다. 또한, "○○○(컨텐츠명)"와 같이 실행 대상을 직접적으로 나타내는 제1 발화 요소는 "\$○○○(컨텐츠명)\$"라는 실행어로 정의될 수 있다. 한편, 실행 명령을 나타내는 제2 발화 요소는 실행 대상을 직접적으로 나타내는 제1 발화 요소 및 발화 음성에 대응되는 응답 정보를 제공할 실행 영역에 기초하여 컨텐츠 검색, 컨텐츠 추천, 장치 기능 실행, 장치 UI 제어 및 웹 검색 등으로 분류될 수 있으며, 각각의 분류된 타입에 대응하여 명령어가 정의될 수 있다. 따라서, 제어부(250)는 사용자의 발화 음성으로부터 발화 요소가 추출되면, 추출된 발화 요소에 대응되는 실행어 및 명령어가 저장부(230)에 저장되어 있는지 여부를 판단한다. 판단 결과, 추출된 발화 요소에 대응되는 실행어 및 명령어가 저장되어 있으면, 제어부(250)는 추출된 발화 요소에 대응되는 실행어 및 명령어를 획득하고, 획득한 실행어 및 명령어를 조합하여 응답 정보를 생성할 수 있다.
- [0045] 예를 들어, "○○○(컨텐츠명) 찾아줘"라는 발화 음성의 경우, 추출부(240)는 "○○○(컨텐츠명)"에 대한 제1 발화 요소 및 "찾아줘"라는 제2 발화 요소를 추출할 수 있다. 이 같이, "○○○(컨텐츠명) 찾아줘"라는 발화 음성으로부터 제1 및 제2 발화 요소가 추출되면, 제어부(250)는 추출된 제1 발화 요소에 기초하여 실행 영역이 컨텐츠인 것으로 판단한다. 또한, 제어부(250)는 저장부(230)에 저장된 제1 발화 요소별 정의된 실행어를 참조하여 "○○○(컨텐츠명)"에 대한 제1 발화 요소에 대응되는 실행어 "\$○○○(컨텐츠명)\$"를 획득할 수 있다. 또한, 제어부(250)는 컨텐츠로 판단된 실행 영역에 기초하여 "찾아줘"에 대한 제2 발화 요소의 타입이 컨텐츠 검색과 관련된 것으로 판단한다. 따라서, 제어부(250)는 저장부(230)에 저장된 제2 발화 요소별 정의된 명령어 중 컨텐츠 검색에 대응되는 명령어 "content_search"를 획득할 수 있다. 이후, 제어부(250)는 제1 및 제2 발화 요소에 대응되는 실행어 "\$○○○(컨텐츠명)\$" 및 명령어 "content_search"를 조합하여 "content_search(\$○○○(컨텐츠명)\$)"라는 실행 명령 스크립트를 생성할 수 있다. 이 같은 실행 명령 스크립트가 생성되면, 제어부(250)는 생성된 실행 명령 스크립트를 포함하는 응답 정보를 디스플레이 장치(100)로 전송한다. 이에 따라, 디스플레이 장치(100)는 대화형 서버(200)로부터 수신한 응답 정보에 포함된 실행 명령 스크립트에 기초하여 컨텐츠 제공 서버(300)로부터 "○○○(컨텐츠명)"에 대한 회차별 컨텐츠 정보를 수신하고, 수신된 회차별 컨텐츠 정보에 대한 컨텐츠 리스트를 화면상에 디스플레이할 수 있다.

- [0046] 그러나, 본 발명은 이에 한정되지 않으며, 제어부(250)는 "content_search(\$○○○(컨텐츠명)\$)"라는 실행 명령 스크립트에 기초하여 저장부(230)에 저장된 EPG 정보로부터 "○○○(컨텐츠명)"에 대한 회차별 컨텐츠 정보를 획득하거나 혹은 외부 서버(400)로부터 수신하여 디스플레이 장치(100)로 전송할 수 있다.
- [0047] 또다른 예를 들어, "취침예약 해취"라는 발화 음성의 경우, 추출부(240)는 "취침예약"에 대한 제1 발화 요소 및 "해취"라는 제2 발화 요소를 추출할 수 있다. 이 같이, "취침예약 해취"라는 발화 음성으로부터 제1 및 제2 발화 요소가 추출되면, 제어부(250)는 추출된 제1 발화 요소에 기초하여 실행 영역이 장치 제어인 것으로 판단한다. 또한, 제어부(250)는 저장부(230)에 저장된 제1 발화 요소별 정의된 실행어를 참조하여 "취침예약"에 대한 제1 발화 요소에 대응되는 실행어 "\$취침예약\$"를 획득할 수 있다. 또한, 제어부(250)는 장치 제어로 판단된 실행 영역에 기초하여 "해취"에 대한 제2 발화 요소의 타입이 장치 기능 실행에 해당하는 것으로 판단한다. 따라서, 제어부(250)는 저장부(230)에 저장된 제2 발화 요소별 정의된 명령어 중 장치 기능 실행에 대응되는 명령어 "device_execute"를 획득할 수 있다. 이후, 제어부(250)는 제1 및 제2 발화 요소에 대응되는 실행어 "\$취침예약\$" 및 명령어 "device_execute"를 조합하여 "device_execute(\$취침예약\$)"라는 실행 명령 스크립트를 생성할 수 있다. 이 같은 실행 명령 스크립트가 생성되면, 제어부(250)는 생성된 실행 명령 스크립트를 포함하는 응답 정보를 디스플레이 장치(100)로 전송한다. 이에 따라, 디스플레이 장치(100)는 대화형 서버(200)로부터 수신한 응답 정보에 포함된 실행 명령 스크립트에 기초하여 취침예약을 설정하고, 취침예약이 설정되었음을 안내하는 알림 메시지를 음성 및 텍스트 중 적어도 하나를 통해 출력할 수 있다.
- [0048] 또다른 예를 들어, "첫번째 거 선택해취"라는 발화 음성과 실행 정보를 포함하는 발화 정보가 수신되면, 추출부(240)는 "첫번째 거 선택해취"라는 발화 음성으로부터 "첫번째"라는 제1 발화 요소 및 "선택해취"라는 제2 발화 요소를 추출할 수 있다. 이 같이, "첫번째 거 선택해취"라는 발화 음성으로부터 제1 및 제2 발화 요소가 추출되면, 제어부(250)는 저장부(230)에 저장된 제1 발화 요소별 정의된 실행어를 참조하여 "첫번째"에 대한 제1 발화 요소에 대응되는 실행어 "\$1st\$"를 획득할 수 있다. 또한, 제어부(250)는 수신된 발화 정보에 포함된 실행 정보에 기초하여 실행 영역이 장치 UI 제어인 것으로 판단하고, 그 판단된 실행 영역에 기초하여 "선택해취"에 대한 제2 발화 요소의 타입이 장치 UI 제어에 해당하는 것으로 판단한다. 따라서, 제어부(250)는 저장부(230)에 저장된 제2 발화 요소별 정의된 명령어 중 장치 UI 제어에 대응되는 명령어 "device_select"를 획득할 수 있다. 한편, 제어부(250)는 "첫번째"에 대한 제1 발화 요소에 대응되는 실행어 "\$1st\$"가 획득되면, 디스플레이 장치(100)로부터 수신한 발화 정보에 포함된 실행 정보에 기초하여 기획된 실행어 "\$1st\$"에 대응되는 실행어로 변환한다. 이후, 제어부(250)는 기획된 실행어 "\$1st\$"에서 변환된 실행어 및 명령어를 조합하여 실행 명령 스크립트를 생성할 수 있다.
- [0049] 이하에서는 전술한 예와 같이, 디스플레이 장치(100)에서 발화 음성과 실행 정보를 포함하는 발화 정보를 대화형 서버(200)로 전송하는 동작 및 대화형 서버(200)에서 디스플레이 장치(100)로부터 수신된 발화 정보에 기초하여 응답 정보를 생성하는 동작에 대해서 구체적으로 설명하도록 한다.
- [0050] 도 4는 본 발명의 일 실시예에 따른 디스플레이 장치에서 화면상에 표시된 객체의 표시 상태에 기초하여 발화 정보를 생성하는 예시도이다.
- [0051] 도 4에 도시된 바와 같이, 디스플레이 장치(100)는 사용자가 요청한 채널을 통해 제1 컨텐츠(30)를 수신하여 디스플레이할 수 있다. 또한, 디스플레이 장치(100)는 사용자 명령에 기초하여 사용자가 요청한 제2 컨텐츠에 대한 컨텐츠 리스트(40)를 화면상에 디스플레이할 수 있다. 도시된 바와 같이, 컨텐츠 리스트(40) 상에는 지금까지 방송된 제2 컨텐츠에 대한 회차별 컨텐츠 정보(41~45)가 표시될 수 있다. 한편, 디스플레이 장치(100)는 기 설정된 우선 순위에 기초하여 컨텐츠 리스트(40) 상에 표시되는 회차별 컨텐츠 정보(41~45) 각각에 대한 식별 정보를 부여할 수 있다. 이 같이 제1 컨텐츠(30) 및 제2 컨텐츠에 대한 컨텐츠 리스트(40)가 화면상에 표시된 상태에서, 사용자는 컨텐츠 리스트(40) 상에 표시된 컨텐츠 정보(41~45)를 참조하여 특정 회차에 대응되는 제2 컨텐츠를 시청하기 위한 발화를 할 수 있다. 예를 들어, 사용자는 3회차 컨텐츠 정보(43)에 대응되는 제2 컨텐츠를 시청하기 위해 "세번째 거 선택해취"라는 발화를 할 수 있다. 이 같은 발화 음성이 입력되면, 디스플레이 장치(100)는 이전 발화 음성이 입력된 입력 시간을 기준으로 현재 입력된 발화 음성이 기설정된 시간 내에 입력되었는지 여부를 판단한다. 판단 결과, 기설정된 시간 내에 입력된 것으로 판단되면, 디스플레이 장치(100)는 "세번째 거 선택해취"라는 발화 음성을 대화형 서버(200)로 전송한다. 한편, 현재 입력된 발화 음성이 기설정된 시간 내에 입력되지 않은 것으로 판단되면, 디스플레이 장치(100)는 "세번째 거 선택해취"라는 발화 음성과 컨텐츠 리스트(40) 상에 표시된 회차별 컨텐츠 정보(41~45)에 부여된 식별 정보(전술한 실행 정보)를 포함하는 발화 정보를 대화형 서버(200)로 전송한다.

- [0052] 이 같은 발화 정보가 수신되면, 추출부(240)는 수신된 발화 정보에 포함된 발화 음성으로부터 "세번째"라는 제1 발화 요소 및 "선택해줘"라는 제2 발화 요소를 추출할 수 있다. 제1 및 제2 발화 요소가 추출되면, 제어부(250)는 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 저장부(230)에 저장되어 있는지 여부를 판단한다. 전술한 바와 같이, 저장부(230)는 제1 및 제2 발화 요소별 정의된 실행어 및 명령어를 저장할 수 있다. 따라서, 제어부(250)는 저장부(230)에 저장된 제1 발화 요소별 정의된 실행어를 참조하여 "세번째"에 대한 제1 발화 요소에 대응되는 실행어 "\$3rd\$"를 획득할 수 있다. 또한, 제어부(250)는 수신된 발화 정보에 포함된 식별 정보에 기초하여 실행 영역이 장치 UI 제어인 것으로 판단하고, 그 판단된 실행 영역에 기초하여 "선택해줘"에 대한 제2 발화 요소의 타입이 장치 UI 제어에 해당하는 것으로 판단한다. 따라서, 제어부(250)는 저장부(230)에 저장된 제2 발화 요소별 정의된 명령어 중 장치 UI 제어에 대응되는 명령어 "device_select"를 획득할 수 있다. 한편, 제어부(250)는 "첫번째"에 대한 제1 발화 요소에 대응되는 실행어 "\$3rd\$"가 획득되면, 디스플레이 장치(100)로부터 수신한 발화 정보에 포함된 식별 정보에 기초하여 기회득한 실행어 "\$3rd\$"에 대응되는 실행어로 변환한다. 예를 들어, 회차별 콘텐츠 정보(41~45) 중 "\$3rd\$"에 대응되는 콘텐츠 정보(43)의 식별 정보가 "#3"이면, 제어부(250)는 기회득한 실행어 "\$3rd\$"에서 변환된 실행어 "#3"으로 변환할 수 있다. 이후, 제어부(250)는 변환된 실행어 "#3"과 명령어 "device_select"를 조합하여 "device_select(#3)"이라는 실행 명령 스크립트를 생성하고, 이를 포함하는 응답 정보를 디스플레이 장치(100)로 전송할 수 있다. 이에 따라, 디스플레이 장치(100)는 대화형 서버(200)로부터 수신한 응답 정보에 포함된 실행 명령 스크립트에 기초하여 콘텐츠 정보(43)에 대응되는 회차의 제2 콘텐츠를 콘텐츠 제공 서버(300)로부터 수신하여 디스플레이할 수 있다.
- [0053] 이와 같이, 본 발명에 따른 제어부(250)는 사용자의 발화 음성으로부터 추출된 발화 요소에 대응되는 실행어 및 명령어가 존재하면, 발화 요소에 대응되는 실행어 및 명령어에 기초하여 사용자의 의도에 대응되는 응답 정보를 생성할 수 있다.
- [0054] 한편, 제어부(250)는 사용자의 발화 음성으로부터 추출된 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 저장부(230)에 기저장된 발화 이력 정보에 기초하여 추출된 발화 요소에 대응되는 실행어 및 명령어를 추정한다. 이후, 제어부(250)는 추정된 지시어 및 명령어를 조합하여 사용자의 의도에 적합한 응답 정보를 생성할 수 있다.
- [0055] 실시예에 따라, 제어부(250)는 추출된 발화 요소 즉, 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 부존재하면, 저장부(230)에 저장된 발화 이력 정보 중, 현재 발화 음성이 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보를 획득한다. 이후, 제어부(250)는 획득된 발화 이력 정보 내에서 제1 발화 요소에 대응되는 실행어 및 제2 발화 요소에 대응되는 명령어를 수신된 발화 음성에 대응되는 실행어 및 명령어로 추정할 수 있다. 한편, 제어부(250)는 저장부(230)에 저장된 발화 이력 정보 중, 현재 발화 음성이 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보가 저장되어 있지 않으면 응답 정보를 생성하지 않는다.
- [0056] 전술한 바와 같이, 저장부(230)는 디스플레이 장치(100)로부터 수신된 발화 음성으로부터 생성된 텍스트 정보, 해당 발화 음성이 수신된 시간 정보, 해당 발화 음성에 기초하여 생성된 응답 정보 및 해당 응답 정보가 생성된 시간 정보 중 적어도 하나를 포함하는 발화 이력 정보를 저장할 수 있다. 따라서, 제어부(250)는 현재 수신된 발화 음성으로부터 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어 중 적어도 하나가 저장부(230)에 저장되어 있지 않으면, 저장부(230)에 기저장된 발화 이력 정보에 기초하여 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어를 추정할 수 있다.
- [0057] 예를 들어, "첫번째 거 선택해줘"라는 발화 음성의 경우, 추출부(230)는 "첫번째"라는 제1 발화 요소와 "선택해줘"라는 제2 발화 요소를 추출할 수 있다. 이 같은 제1 및 제2 발화 요소가 추출되면, 제어부(250)는 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 저장부(230)에 저장되어 있는지 여부를 판단한다. 구체적으로, 제어부(250)는 저장부(230)에 저장된 제1 발화 요소별 정의된 실행어를 참조하여 "첫번째"라는 제1 발화 요소에 대응되는 실행어 "\$1st\$"를 획득할 수 있다. 한편, 제어부(250)는 추출된 발화 요소 중 제1 발화 요소에 기초하여 실행 영역 판단이 가능한지 여부를 판단한다. 판단 결과, "첫번째"에 대한 제1 발화 요소로부터 실행 영역이 불가능한 것으로 판단되면, 제어부(250)는 제2 발화 요소별 정의된 명령어를 저장하는 저장부(230)로부터 "선택해줘"라는 제2 발화 요소에 대응되는 명령어를 획득하지 못한다. 따라서, 제어부(250)는 "첫번째 거 선택해줘"라는 발화 음성으로부터 발화 음성에 대한 서비스를 제공하기 위한 실행 영역 및 실행 명령이 불명확한 것으로 판단한다. 이 같은 판단 결과에 따라, 제어부(250)는 저장부(230)에 저장된 발화 이력 정보 중 "첫번째 거 선택해줘"라는 발화 음성이 수신된 시점을 기준으로 기설정된 시간 범위 이내에 저장된 발화 이력 정보가 있는지 여부를 판단한다. 즉, 제어부(250)는 저장부(230)에 저장된 발화 이력 정보 중 "첫번째 거 선택해줘"라는 발화 음성이 수신된 시간과 가장 최근에 저장된 이전 발화 이력 정보의 시간 차이가 기설정된 시간 범위 내에

속하는지 여부를 판단한다. 판단 결과, 이전 발화 이력 정보가 기설정된 시간 범위 내에 속하지 않은 것으로 판단되면, 제어부(250)는 "첫번째 거 선택해줘"라는 발화 음성에 대한 응답 정보를 생성하지 않는다. 한편, 이전 발화 이력 정보가 기설정된 시간 범위 내에 속하는 것으로 판단되면, 추출부(240)는 이전 발화 이력 정보에 포함된 발화 음성의 텍스트 정보로부터 제1 및 제2 발화 요소를 추출한다.

[0058] 예를 들어, 이전 발화 이력 정보에 포함된 발화 음성이 "○○○(컨텐츠명) 찾아줘"이면, 추출부(240)는 "○○○(컨텐츠명)"에 대한 제1 발화 요소와 "찾아줘"라는 제2 발화 요소를 추출한다. 이후, 제어부(250)는 이전 발화 이력 정보에 포함된 발화 음성으로부터 추출된 제1 및 제2 발화 요소 중 "○○○(컨텐츠명)"에 대한 제1 발화 요소에 기초하여 "첫번째"에 대한 제1 발화 요소가 "○○○(컨텐츠명)"에 대한 제1 발화 요소인 것으로 추정한다. 이후, 제어부(250)는 추정된 제1 발화 요소에 기초하여 실행 영역이 장치 UI 제어인 것으로 판단하고, 저장부(230)에 저장된 제2 발화 요소별 정의된 명령어 중 장치 UI 제어에 대응되는 명령어 "device_select"를 획득할 수 있다. 이후, 제어부(250)는 획득한 제1 발화 요소에 대응되는 실행어 "\$1st\$"와 제2 발화 요소에 대응되는 명령어 "device_select"를 조합하여 "device_select(\$1st\$)"라는 실행 명령 스크립트를 생성할 수 있다. 이후, 제어부(250)는 생성된 실행 명령 스크립트를 포함하는 응답 정보를 디스플레이 장치(100)로 전송한다. 따라서, 디스플레이 장치(100)는 대화형 서버(200)로부터 수신한 응답 정보에 기초하여 "첫번째 컨텐츠를 선택하셨습니다."라는 응답 메시지를 음성 및 텍스트 중 적어도 하나를 출력할 수 있다. 이와 함께, 디스플레이 장치는 응답 정보에 포함된 실행 명령 스크립트에 기초하여 컨텐츠 제공 서버(300)로부터 첫번째 컨텐츠를 수신하여 디스플레이할 수 있다.

[0059] 한편, 본 발명의 추가적인 양상에 따라 제어부(250)는 수신된 발화 음성 내에 제1 발화 요소 및 제2 발화 요소 중 적어도 하나가 포함되어 있지 않으면, 저장부(230)에 저장된 발화 이력 정보 중 발화 음성이 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보에 기초하여 응답 정보를 생성할 수 있다.

[0060] 예를 들어, "그래"와 같은 발화 음성의 경우, 추출부(240)는 "그래"라는 발화 음성으로부터 제1 및 제2 발화 요소를 추출하지 못한다. 이 경우, 제어부(250)는 "그래"라는 발화 음성이 불명확한 것으로 판단하고, 저장부(230)에 저장된 발화 이력 정보 중 "그래"라는 발화 음성이 수신된 시점을 기준으로 기설정된 시간 범위 이내에 저장된 이전 발화 이력 정보를 획득한다. 예를 들어, 획득한 이전 발화 이력 정보에 포함된 발화 음성이 "TV 꺼줘"이면, 제어부(250)는 "TV"에 대한 제1 발화 요소와 "꺼줘"라는 제2 발화 요소를 추출한다. 이후, 제어부(250)는 추출된 제1 및 제2 발화 요소에 기초하여 현재 수신된 "그래"라는 발화 음성이 컨펌(confirm)인 것으로 추정한다. 이에 따라, 제어부(250)는 이전 발화 이력 정보로부터 추출된 "TV"에 대한 제1 발화 요소에 기초하여 실행 영역이 장치 제어인 것으로 판단한다. 또한, 제어부(250)는 저장부(230)에 저장된 제1 발화 요소별 정의된 실행어를 참조하여 "TV"에 대한 제1 발화 요소에 대응되는 실행어 "\$TV\$"를 획득할 수 있다. 또한, 제어부(250)는 장치 제어로 판단된 실행 영역에 기초하여 "꺼줘"에 대한 제2 발화 요소의 타입이 장치 기능 실행과 관련된 것으로 판단한다. 따라서, 제어부(250)는 저장부(230)에 제2 발화 요소별 정의된 명령어 중 장치 기능 실행에 대응되는 명령어를 획득한다. 이후, 제어부(250)는 장치 기능 실행에 대응하여 획득한 명령어 중 전원 제어에 대응되는 명령어 "device_exectue_turn off"를 획득할 수 있다. 이후, 제어부(250)는 제1 발화 요소에 대응되는 실행어 "\$TV\$"와 제2 발화 요소에 대응되는 명령어 "device_exectue_turn off"를 조합하여 "device_exectue_turn off(\$TV\$)"라는 실행 명령 스크립트를 생성한다. 이 같은 실행 명령 스크립트가 생성되면, 제어부(250)는 생성된 실행 명령 스크립트를 포함하는 응답 정보를 디스플레이 장치(100)로 전송한다. 이에 따라, 디스플레이 장치(100)는 대화형 서버(200)로부터 수신한 응답 정보에 포함된 실행 명령 스크립트에 기초하여 TV 전원을 오픈시킬 수 있다.

[0061] 지금까지, 본 발명에 따른 대화형 서버(200)에 각 구성에 대해서 상세히 설명하였다. 이하에서는, 사용자의 발화 음성을 입력받는 디스플레이 장치(100)의 각 구성에 대해서 상세히 설명하도록 한다.

[0062] 도 5는 본 발명의 일 실시예에 따른 디스플레이 장치의 블록도이다.

[0063] 도 5에 도시된 바와 같이, 디스플레이 장치(100)는 입력부(110), 통신부(120), 디스플레이부(130), 제어부(140) 및 저장부(150)를 포함한다.

[0064] 입력부(110)는 사용자로부터 발화된 발화 음성을 입력받는다. 구체적으로, 입력부(110)는 아날로그 형태의 사용자의 발화 음성이 입력되면, 입력된 사용자 발화 음성을 샘플링하여 디지털 신호로 변환한다. 이때, 입력부(110)는 입력된 사용자의 발화 음성에 노이즈(예를 들어, 에어컨 소리, 청소기 소리 등)가 있는 경우, 노이즈를 제거한 후, 노이즈가 제거된 사용자의 발화 음성을 디지털 신호로 변환함이 바람직하다. 뿐만 아니라, 입력부(110)는 다양한 사용자 조작을 입력받아 제어부(130)로 전달하는 기능을 수행할 수 있다. 이 경우, 입력부

(110)는 터치패드(Touch Pad) 혹은 각종 기능키, 숫자키, 특수키, 문자키 등을 구비한 키패드(Key Pad) 또는 터치 스크린(Touch Screen)을 통해 사용자 조작 명령을 입력받을 수 있다.

[0065] 통신부(120)는 입력부(110)를 통해 입력된 사용자의 발화 음성을 서버 장치(이하 대화형 서버라 함)로 전송하고, 전송된 발화 음성에 대응되는 응답 정보를 수신한다. 여기서, 통신부(120)는 근거리 무선 통신 모듈(미도시), 무선 통신 모듈(미도시) 등과 같은 다양한 통신 모듈을 포함할 수 있다. 여기서, 근거리 무선 통신 모듈(미도시)은 근거리에 위치한 대화형 서버(200) 및 콘텐츠를 제공하는 외부 서버(미도시)와 무선 통신을 수행하는 통신 모듈로써, 예를 들어, 블루투스, 지그비 등이 될 수 있다. 무선 통신 모듈(미도시)은 와이파이(WiFi), IEEE 등과 같은 무선 통신 프로토콜에 따라 외부 네트워크에 연결되어 통신을 수행하는 모듈이다. 이 밖에 무선 통신 모듈은 3G(3rd Generation), 3GPP(3rd Generation Partnership Project), LTE(Long Term Evolution) 등과 같은 다양한 이동 통신 규격에 따라 이동 통신 망에 접속하여 통신을 수행하는 이동 통신 모듈을 더 포함할 수도 있다.

[0066] 디스플레이부(130)는 액정 표시 장치(Liquid Crystal Display, LCD), 유기 전기 발광 다이오드(Organic Light Emitting Display, OLED) 또는 플라즈마 표시 패널(Plasma Display Panel, PDP) 등으로 구현되어, 디스플레이 장치(100)를 통해 제공 가능한 다양한 디스플레이 화면을 제공할 수 있다. 특히, 디스플레이부(161) 대화형 서버(200)로부터 수신한 응답 정보에 기초하여 사용자의 발화 음성에 대응되는 콘텐츠 혹은 콘텐츠 정보를 디스플레이한다.

[0067] 제어부(140)는 디스플레이 장치(100)의 구성을 전반적으로 제어한다. 특히, 제어부(140)는 대화형 서버(200)로부터 실행어 및 명령어의 조합으로 생성된 실행 명령 스크립트를 포함하는 응답 정보가 수신되면, 수신된 응답 정보에 포함된 실행 명령 스크립트에 기초하여 동작을 수행한다.

[0068] 전술한 예와 같이, "○○○(콘텐츠명) 찾아줘"라는 사용자의 발화 음성에 대한 응답 정보가 대화형 서버(200)로부터 수신되면, 제어부(140)는 수신된 응답 정보에 포함된 실행 명령 스크립트에 기초하여 통신부(120)를 통해 콘텐츠 제공 서버(300)로부터 "○○○(콘텐츠명)"에 대한 회차별 콘텐츠 정보를 수신하고, 수신된 회차별 콘텐츠 정보에 대한 콘텐츠 리스트를 디스플레이부(130)를 통해 화면상에 디스플레이할 수 있다. 구체적으로, 대화형 서버(200)로부터 수신되는 응답 정보에는 "content_search(\$○○○(콘텐츠명)\$)"라는 실행 명령 스크립트가 포함될 수 있다. 따라서, 제어부(140)는 "content_search(\$○○○(콘텐츠명)\$)"라는 실행 명령 스크립트를 해석하여, 실행어는 "\$○○○(콘텐츠명)\$"이며, 명령어는 "content_search"라고 구분할 수 있다. 이 같이, 실행 명령 스크립트로부터 지시자 및 명령어가 구분되면, 제어부(140)는 구분된 지시자에 기초하여 콘텐츠 제공 서버(300)로부터 "○○○(콘텐츠명)"에 대한 회차별 콘텐츠 정보를 수신하고, 수신된 회차별 콘텐츠 정보에 대한 콘텐츠 리스트가 화면상에 디스플레이되도록 디스플레이부(130)를 제어할 수 있다. 그러나, 본 발명은 이에 한정되지 않으며, 통신부(120)는 대화형 서버(200)로부터 "○○○(콘텐츠명)"에 대한 회차별 콘텐츠 정보를 포함하는 응답 정보를 수신할 수 있다. 이 경우, 제어부(140)는 대화형 서버(200)로부터 수신된 응답 정보에 포함된 "○○○(콘텐츠명)"에 대한 회차별 콘텐츠 정보를 디스플레이부(130)를 통해 화면상에 디스플레이할 수 있다.

[0069] 전술한 또다른 예와 같이, "취침예약 해줘"라는 사용자의 발화 음성에 대한 응답 정보가 대화형 서버(200)로부터 수신되면, 제어부(140)는 수신된 응답 정보에 포함된 실행 명령 스크립트에 기초하여 기설정된 시간으로 취침예약을 자동 설정할 수 있다. 구체적으로, 대화형 서버(200)로부터 수신되는 응답 정보에는 "device_execute(\$취침예약\$)"라는 실행 명령 스크립트가 포함될 수 있다. 따라서, 제어부(140)는 "device_execute(\$취침예약\$)"라는 실행 명령 스크립트를 해석하여, 실행어는 "\$\$취침예약\$\$"이며, 명령어는 "device_execute"라고 구분할 수 있다. 이 같이, 실행 명령 스크립트로부터 지시자 및 명령어가 구분되면, 제어부(140)는 구분된 지시자에 기초하여 기설정된 시간으로 취침예약을 자동 설정할 수 있다. 이 같은 취침예약이 자동 설정되면, 제어부(140)는 수신된 응답 정보에 기초하여 "취침예약 설정이 완료되었습니다."라는 텍스트 형태의 알림 메시지를 디스플레이부(130)를 통해 화면상에 출력할 수 있다.

[0070] 한편, 도 4에서 설명한 바와 같이, 디스플레이부(130)는 제어부(140)의 제어 명령에 따라, 사용자가 요청한 채널을 통해 제1 콘텐츠(30)를 수신하여 디스플레이할 수 있다. 또한, 디스플레이부(130)는 사용자 명령에 기초하여 사용자가 요청한 제2 콘텐츠에 대한 콘텐츠 리스트(40)를 화면상에 디스플레이할 수 있다. 이 같이, 제2 콘텐츠에 대한 회차별 콘텐츠 정보(41~45)를 포함하는 콘텐츠 리스트(40)가 화면상에 디스플레이되면, 제어부(140)는 기설정된 우선 순위에 기초하여 회차별 콘텐츠 정보(41~45) 각각에 대한 식별 정보를 부여할 수 있다. 이 같이, 제1 콘텐츠(30) 및 제2 콘텐츠에 대한 콘텐츠 리스트(40)가 화면상에 디스플레이된 상태에서, 사용자로부터 특정 회차에 대응되는 제2 콘텐츠를 시청하기 위한 발화 음성이 입력될 수 있다. 예를 들어, 사용자는

3회차 콘텐츠 정보(43)에 대응되는 제2 콘텐츠를 시청하기 위해 "세번째 거 선택해줘"라는 발화를 할 수 있다. 이 같은 발화 음성이 입력되면, 제어부(140)는 "세번째 거 선택해줘"라는 발화 음성에 대한 발화 이력 정보가 저장부(150)에 저장되어 있는지 여부를 판단한다. 여기서, 발화 이력 정보는 이전 사용자로부터 입력된 발화 음성 및 해당 발화 음성이 입력된 시간 정보를 포함할 수 있다. 판단 결과, 발화 이력 정보가 저장되어 있지 않을 경우, 제어부(140)는 "세번째 거 선택해줘"라는 발화 음성과 회차별 콘텐츠 정보(41~45)에 부여된 식별 정보를 포함하는 발화 정보를 대화형 서버(200)로 전송할 수 있다.

[0071] 한편, 저장부(150)에 발화 이력 정보가 저장되어 있으면, 제어부(140)는 저장된 발화 이력 정보에 포함된 시간 정보에 기초하여 가장 최근의 발화 이력 정보를 획득한다. 이후, 제어부(140)는 획득한 발화 이력 정보에 포함된 시간 정보를 기준으로 "세번째 거 선택해줘"라는 발화 음성이 기설정된 시간 내에 입력되었는지 여부를 판단한다. 판단 결과, 기설정된 시간 내에 입력된 것으로 판단되면, 제어부(140)는 통신부(120)를 통해 "세번째 거 선택해줘"라는 발화 음성을 대화형 서버(200)로 전송한다. 한편, 현재 입력된 발화 음성이 기설정된 시간 내에 입력되지 않은 것으로 판단되면, 제어부(140)는 "세번째 거 선택해줘"라는 발화 음성과 콘텐츠 리스트(40)상에 표시된 회차별 콘텐츠 정보(41~45)에 부여된 식별 정보(전송한 실행 정보)를 포함하는 발화 정보를 대화형 서버(200)로 전송한다.

[0072] 이후, 대화형 서버(200)로부터 "세번째 거 선택해줘"라는 발화 음성에 대한 응답 정보가 수신되면, 제어부(140)는 수신된 응답 정보에 포함된 실행 명령 스크립트에 기초하여 회차별 콘텐츠 정보(41~45) 중 3회차 콘텐츠 정보(43)에 대응되는 제2 콘텐츠를 콘텐츠 제공 서버(300)로부터 수신하여 디스플레이부(130)를 통해 디스플레이할 수 있다.

[0073] 지금까지, 본 발명에 따른 사용자의 발화 음성을 인식하고, 인식된 발화 음성에 대응되는 응답 정보에 기초하여 동작을 수행하는 디스플레이 장치(100)의 각 구성에 대해서 상세히 설명하였다. 이하에서는, 사용자의 발화 음성에 대응되는 응답 정보를 제공하는 대화형 서버(200) 및 응답 정보에 기초하여 동작을 수행하는 디스플레이 장치(100)의 제어 방법에 대해서 상세히 설명하도록 한다.

[0074] 도 6은 본 발명의 일 실시예에 따른 대화형 서버의 제어 방법에 대한 흐름도이다.

[0075] 도 6에 도시된 바와 같이, 대화형 서버(200)는 디스플레이 장치(100)로부터 발화 음성 신호(이하 발화 음성이라 함)를 수신한다(S610). 여기서, 사용자의 발화 음성은 아날로그 형태의 음성 신호에서 디지털 신호로 변환된 음성 신호이다. 이 같은 발화 음성이 수신되면, 대화형 서버(200)는 수신된 발화 음성을 텍스트 정보로 변환한다(S620). 실시예에 따라, 대화형 서버(200)는 STT(Speech to Text) 알고리즘을 이용하여 사용자의 발화 음성을 텍스트로 변환할 수 있다. 그러나, 본 발명은 이에 한정되지 않으며, 대화형 서버(200)는 디스플레이 장치(100)로부터 사용자의 발화 음성에 대한 텍스트 정보를 수신할 수 있다. 이 경우, 디스플레이 장치(100)는 입력된 사용자의 발화 음성에 대한 텍스트 정보를 전송한 제1 서버(10)와 같은 ASR 서버로부터 수신하여 대화형 서버(200)로 전송한다. 따라서, 대화형 서버(200)는 디스플레이 장치(100)로부터 사용자의 발화 음성에 대한 텍스트 정보를 수신할 수 있다.

[0076] 이 같은 사용자의 발화 음성이 텍스트 정보로 변환되거나 혹은 디스플레이 장치(100)로부터 사용자의 발화 음성에 대한 텍스트 정보가 수신되면, 대화형 서버(200)는 텍스트 정보로 변환된 발화 음성으로부터 발화 요소를 추출한다(S630). 즉, 대화형 서버(200)는 텍스트 정보로 변환된 발화 음성으로부터 주요 특징을 나타내는 제1 발화 요소 및 실행 명령을 나타내는 제2 발화 요소를 추출한다. 이 같은 제1 및 제2 발화 요소 중 적어도 하나의 발화 요소가 추출되면, 대화형 서버(200)는 추출된 발화 요소에 기초하여 사용자의 의도에 적합한 응답 정보의 생성 여부를 판단한다(S640). 판단 결과, 추출된 발화 요소에 기초하여 응답 정보 생성이 불가능한 것으로 판단되면, 대화형 서버(200)는 사용자의 발화 음성으로부터 응답 정보를 생성하지 않는다. 한편, 판단 결과, 추출된 발화 요소에 기초하여 응답 정보 생성이 가능한 것으로 판단되면, 대화형 서버(200)는 사용자의 발화 음성으로부터 응답 정보를 생성한다. 이후, 대화형 서버(200)는 생성된 응답 정보를 디스플레이 장치(100)로 전송한다(S650).

[0077] 구체적으로, 대화형 서버(200)는 사용자의 발화 음성으로부터 제1 및 제2 발화 요소가 추출되면, 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 기저장되어 있는지 여부를 판단한다. 여기서, 제1 발화 요소는 주요 특징을 나타내는 발화 요소이며, 제2 발화 요소는 실행 명령을 나타내는 발화 요소이다. 따라서, 대화형 서버(200)는 주요 특징을 나타내는 제1 발화 요소에 대응되는 실행어 및 실행 명령을 나타내는 제2 발화 요소에 대응되는 명령어가 기저장되어 있는지 여부를 판단한다. 판단 결과, 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 저장되어 있으면, 대화형 서버(200)는 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및

명령어의 조합하여 실행 명령 스크립트를 생성한다. 한편, 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 저장되어 있지 않으면, 대화형 서버(200)는 기저장된 발화 이력 정보에 기초하여 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어 추정이 가능한지 여부를 판단한다. 판단 결과, 기저장된 발화 이력 정보에 기초하여 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어 추정이 가능하면, 대화형 서버(200)는 추정된 실행어 및 명령어를 조합하여 실행 명령 스크립트를 생성한다. 한편, 기저장된 발화 이력 정보에 기초하여 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어 추정이 불가능하면, 대화형 서버(200)는 사용자의 발화 음성에 대한 응답 정보를 생성하지 않는다. 이와 같이, 본 발명에 따른 대화형 서버(200)는 사용자의 발화 음성으로부터 추출된 발화 요소 및 기저장된 발화 이력 정보에 기초하여 사용자의 발화 음성에 대한 응답 정보 생성 여부를 판단하고, 그 판단 결과에 따라 생성된 응답 정보를 디스플레이 장치(100)로 전송할 수 있다.

[0078] 이하에서는, 도 7를 통해 대화형 서버(200)에서 사용자의 발화 음성으로부터 추출된 발화 요소 및 기저장된 발화 이력 정보에 기초하여 응답 정보를 생성하는 방법에 대해서 보다 구체적으로 설명하도록 한다.

[0079] 도 7은 본 발명의 일 실시예에 따른 대화형 서버에서 응답 정보를 생성하는 방법의 흐름도이다.

[0080] 도 7에 도시된 바와 같이, 전술한 단계 S620을 통해 사용자의 발화 음성이 텍스트 정보로 변환되거나, 혹은 디스플레이 장치(200)로부터 텍스트 정보로 변환된 발화 음성이 수신될 수 있다. 이 같이, 사용자의 발화 음성이 텍스트 정보로 변환되거나 혹은 수신되면, 대화형 서버(200)는 텍스트 정보로 변환된 발화 음성으로부터 발화 요소를 추출한다. 발화 요소가 추출되면, 대화형 서버(200)는 추출된 발화 요소를 분석하여 주요 특징을 나타내는 제1 발화 요소와 실행 명령을 나타내는 제2 발화 요소가 있는지 여부를 판단한다(S710). 판단 결과, 추출된 발화 요소가 제1 발화 요소 및 제2 발화 요소 중 적어도 하나의 발화 요소를 포함하지 않는 것으로 판단되면, 대화형 서버(200)는 후술할 단계 S740을 통해 이전 발화 이력 정보가 있는지 여부를 판단한다.

[0081] 한편, 사용자의 발화 음성으로부터 제1 및 제2 발화 요소가 추출되면, 대화형 서버(200)는 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 존재하는지 여부를 판단한다(S720). 판단 결과, 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 존재하면, 대화형 서버(200)는 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어를 조합하여 응답 정보를 생성한다(S730). 한편, 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 존재하지 않으면, 대화형 서버(200)는 사용자의 발화 음성이 수신되기 이전에 수신된 발화 음성에 대한 발화 이력 정보가 존재하는지 여부를 판단한다(S740). 판단 결과, 발화 이력 정보가 존재하지 않으면, 대화형 서버(200)는 사용자의 발화 음성에 대한 응답 정보를 생성하지 않는다. 한편, 발화 이력 정보가 존재하면, 대화형 서버(200)는 사용자의 발화 음성이 수신된 시점을 기준으로 기설정된 시간 범위 이내의 발화 이력 정보를 획득한다(S750). 이후, 대화형 서버(200)는 획득한 발화 이력 정보 내에서 사용자의 발화 음성으로부터 추출된 제1 발화 요소에 대응되는 실행어 및 제2 발화 요소에 대응되는 명령어를 사용자의 발화 음성에 대응되는 실행어 및 명령어로 추정한다(S760). 이 같이, 사용자의 발화 음성에 대응되는 실행어 및 명령어가 추정되면, 대화형 서버(200)는 추정된 실행어 및 명령어를 조합하여 응답 정보를 생성한다(S770).

[0082] 먼저, 사용자의 발화 음성으로부터 추출된 발화 요소가 제1 및 제2 발화 요소 중 적어도 하나를 포함하지 않는 경우, 대화형 서버(200)에서 사용자의 발화 음성에 대한 응답 정보를 생성하는 방법에 대해서 설명하도록 한다.

[0083] 예를 들어, "그래"와 같은 발화 음성의 경우, 대화형 서버(200)는 "그래"라는 발화 음성으로부터 제1 및 제2 발화 요소를 추출하지 못한다. 이 경우, 대화형 서버(200)는 "그래"라는 발화 음성이 불명확한 것으로 판단하고, 기저장된 발화 이력 정보 중 "그래"라는 발화 음성이 수신된 시점을 기준으로 기설정된 시간 범위 이내에 저장된 이전 발화 이력 정보를 획득한다. 예를 들어, 획득한 이전 발화 이력 정보에 포함된 발화 음성이 "TV 꺼줘"이면, 대화형 서버(200)는 "TV"에 대한 제1 발화 요소와 "꺼줘"라는 제2 발화 요소를 추출한다. 이후, 대화형 서버(200)는 추출된 제1 및 제2 발화 요소에 기초하여 현재 수신된 "그래"라는 발화 음성이 컨펌(confirm)인 것으로 추정한다. 이에 따라, 대화형 서버(200)는 이전 발화 이력 정보로부터 추출된 "TV"에 대한 제1 발화 요소에 기초하여 실행 영역이 장치 제어인 것으로 판단한다. 또한, 대화형 서버(200)는 기저장된 제1 발화 요소별 정의된 실행어를 참조하여 "TV"에 대한 제1 발화 요소에 대응되는 실행어 "\$TV\$"를 획득할 수 있다. 또한, 대화형 서버(200)는 장치 제어로 판단된 실행 영역에 기초하여 "꺼줘"에 대한 제2 발화 요소의 타입이 장치 기능 실행과 관련된 것으로 판단한다. 따라서, 대화형 서버(200)는 기저장된 제2 발화 요소별 정의된 명령어 중 장치 기능 실행에 대응되는 명령어를 획득한다. 이후, 대화형 서버(200)는 장치 기능 실행에 대응하여 획득한 명령어 중 전원 제어에 대응되는 명령어 "device_exectue_turn off"를 획득할 수 있다. 이후, 대화형 서버(200)는 제1 발화 요소에 대응되는 실행어 "\$TV\$"와 제2 발화 요소에 대응되는 명령어 "device_exectue_turn

off"를 조합하여 "device_exectue_turn off(\$TV\$)"라는 실행 명령 스크립트를 생성한다. 이 같은 실행 명령 스크립트가 생성되면, 대화형 서버(200)는 생성된 실행 명령 스크립트를 포함하는 응답 정보를 생성할 수 있다.

[0084] 이하에서는, 사용자의 발화 음성으로부터 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 존재하는 경우, 대화형 서버(200)에서 사용자의 발화 음성에 대한 응답 정보를 생성하는 방법에 대해서 설명하도록 한다.

[0085] 한편, "○○○(컨텐츠명) 보여줘"라는 발화 음성이 입력되면, 대화형 서버(200)는 입력된 발화 음성으로부터 "○○○(컨텐츠명)"에 대한 제1 발화 요소 및 "찾아줘"라는 제2 발화 요소를 추출할 수 있다. 이 같은 제1 및 제2 발화 요소가 추출되면, 대화형 서버(200)는 추출된 제1 발화 요소에 기초하여 실행 영역이 컨텐츠인 것으로 판단한다. 또한, 대화형 서버(200)는 기저장된 제1 발화 요소별 정의된 실행어 중 "○○○(컨텐츠명)"에 대한 제1 발화 요소에 대응되는 실행어 "\$○○○(컨텐츠명)\$"를 획득할 수 있다. 또한, 대화형 서버(200)는 컨텐츠로 판단된 실행 영역에 기초하여 "찾아줘"에 대한 제2 발화 요소의 타입이 컨텐츠 검색과 관련된 것으로 판단한다. 따라서, 대화형 서버(200)는 기저장된 제2 발화 요소별 정의된 명령어 중 컨텐츠 검색에 대응되는 명령어 "content_search"를 획득할 수 있다.

[0086] 이와 같이, 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 획득되면, 대화형 서버는 획득한 실행어 "\$○○○(컨텐츠명)\$"와 명령어 "content_search"를 조합하여 "content_search(\$○○○(컨텐츠명))"라는 실행 명령 스크립트를 생성한다. 이 같은 실행 명령 스크립트가 생성되면, 대화형 서버(200)는 생성된 실행 명령 스크립트를 포함하는 응답 정보를 생성할 수 있다.

[0087] 이하에서는, 사용자의 발화 음성으로부터 추출된 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 부존재하는 경우, 대화형 서버(200)에서 사용자의 발화 음성에 대한 응답 정보를 생성하는 방법에 대해서 설명하도록 한다.

[0088] 예를 들어, "첫번째 거 선택해줘"라는 발화 음성의 경우, 대화형 서버(200)는 "첫번째"라는 제1 발화 요소와 "선택해줘"라는 제2 발화 요소를 추출할 수 있다. 이 같은 제1 및 제2 발화 요소가 추출되면, 대화형 서버(200)는 제1 및 제2 발화 요소에 대응되는 실행어 및 명령어가 기저장되어 있는지 여부를 판단한다. 구체적으로, 대화형 서버(200)는 기저장된 제1 발화 요소별 정의된 실행어를 참조하여 "첫번째"라는 제1 발화 요소에 대응되는 실행어 "\$1st\$"를 획득할 수 있다. 한편, 대화형 서버(200)는 추출된 제1 발화 요소에 기초하여 실행 영역 판단이 가능한지 여부를 판단한다. 판단 결과, "첫번째"에 대한 제1 발화 요소로부터 실행 영역이 불가능한 것으로 판단되면, 대화형 서버(200)는 기저장된 제2 발화 요소별 정의된 명령어로부터 "선택해줘"라는 제2 발화 요소에 대응되는 명령어를 획득하지 못한다. 따라서, 대화형 서버(200)는 "첫번째 거 선택해줘"라는 발화 음성으로부터 발화 음성에 대한 서비스를 제공하기 위한 실행 영역 및 실행 명령이 불명확한 것으로 판단한다. 이 같은 판단 결과에 따라, 대화형 서버(200)는 기저장된 발화 이력 정보 중 "첫번째 거 선택해줘"라는 발화 음성이 수신된 시점을 기준으로 기설정된 시간 범위 이내에 저장된 발화 이력 정보가 있는지 여부를 판단한다. 즉, 대화형 서버(200)는 기저장된 발화 이력 정보 중 "첫번째 거 선택해줘"라는 발화 음성이 수신된 시간과 가장 최근에 저장된 이전 발화 이력 정보의 시간 차이가 기설정된 시간 범위 내에 속하는지 여부를 판단한다. 판단 결과, 이전 발화 이력 정보가 기설정된 시간 범위 내에 속하지 않은 것으로 판단되면, 대화형 서버(200)는 "첫번째 거 선택해줘"라는 발화 음성에 대한 응답 정보를 생성하지 않는다. 한편, 이전 발화 이력 정보가 기설정된 시간 범위 내에 속하는 것으로 판단되면, 대화형 서버(200)는 이전 발화 이력 정보에 포함된 발화 음성의 텍스트 정보로부터 제1 및 제2 발화 요소를 추출한다.

[0089] 예를 들어, 이전 발화 이력 정보에 포함된 발화 음성이 "○○○(컨텐츠명) 찾아줘"이면, 대화형 서버(200)는 "○○○(컨텐츠명)"에 대한 제1 발화 요소와 "찾아줘"라는 제2 발화 요소를 추출한다. 이후, 대화형 서버(200)는 이전 발화 이력 정보에 포함된 발화 음성으로부터 추출된 제1 및 제2 발화 요소 중 "○○○(컨텐츠명)"에 대한 제1 발화 요소에 기초하여 "첫번째"에 대한 제1 발화 요소가 "○○○(컨텐츠명)"에 대한 제1 발화 요소인 것으로 추정한다. 이후, 대화형 서버(200)는 추정된 제1 발화 요소에 기초하여 실행 영역이 장치 UI 제어인 것으로 판단하고, 기저장된 제2 발화 요소별 정의된 명령어 중 장치 UI 제어에 대응되는 명령어 "device_select"를 획득할 수 있다. 이후, 대화형 서버(200)는 기획득한 제1 발화 요소에 대응되는 실행어 "\$1st\$"와 제2 발화 요소에 대응되는 명령어 "device_select"를 조합하여 "device_select(\$1st\$)"라는 실행 명령 스크립트를 생성한다. 이 같은 실행 명령 스크립트가 생성되면, 대화형 서버(200)는 생성된 실행 명령 스크립트를 포함하는 응답 정보를 생성할 수 있다.

[0090] 이와 같이, 본 발명에 따른 대화형 서버(200)는 사용자의 발화 음성을 분석하여 사용자의 발화 의도를

파악하고, 그 파악된 발화 의도에 기초하여 응답 정보 생성 여부를 판단한다. 판단 결과, 사용자의 발화 의도에 기초하여 응답 정보 생성이 불가능하면, 대화형 서버(200)는 사용자의 이전 발화 음성에 기초하여 현재 사용자의 발화 음성에 대한 응답 정보를 생성할 수 있다. 따라서, 사용자는 특정 발화에 국한되지 않고 자신이 표현하고자 하는 방식으로 발화하여 그에 따른 서비스를 제공받을 수 있다.

[0091] 이제까지 본 발명에 대하여 그 바람직한 실시예들을 중심으로 살펴보았다.

[0092] 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

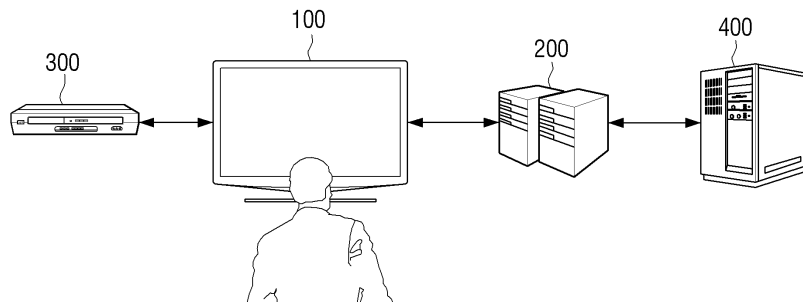
부호의 설명

[0093]

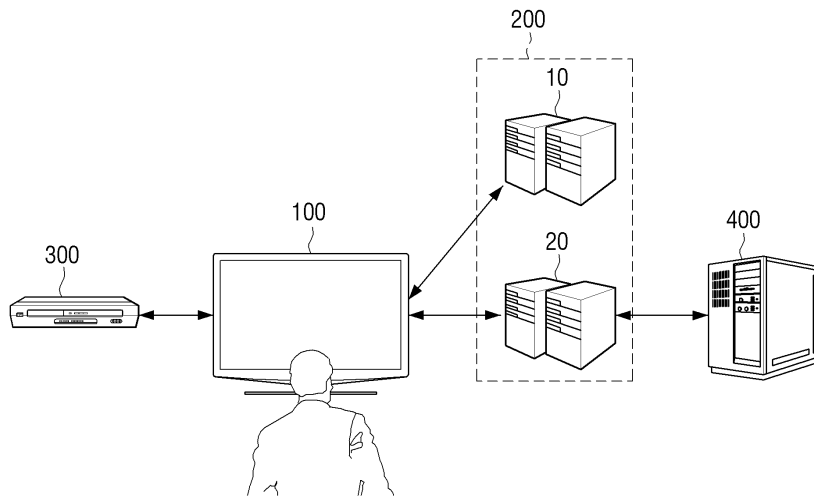
10 : 제1 서버	20 : 제2 서버
100 : 디스플레이 장치	110 : 입력부
120,210 : 통신부	130 : 디스플레이부
140,250 : 제어부	150,230 : 저장부
200 : 대화형 서버	220 : 음성 처리부
240 : 추출부	300 : 콘텐츠 제공 서버
400 : 외부 서버	

도면

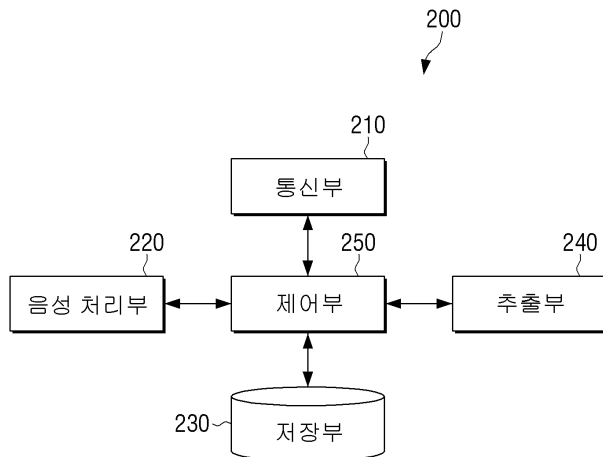
도면1



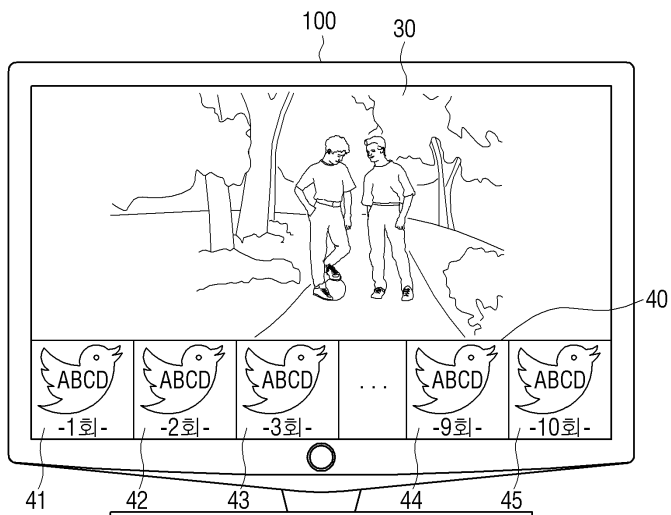
도면2



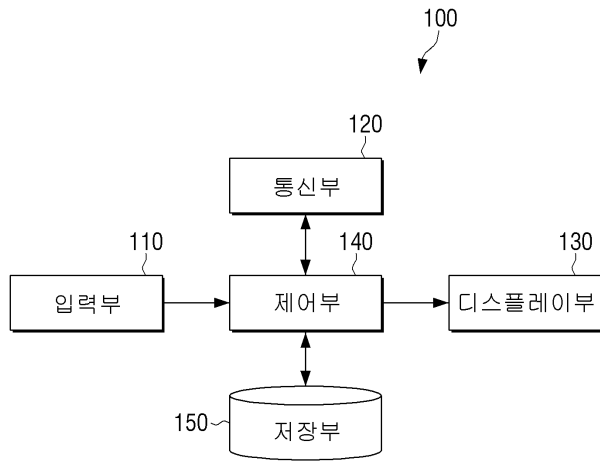
도면3



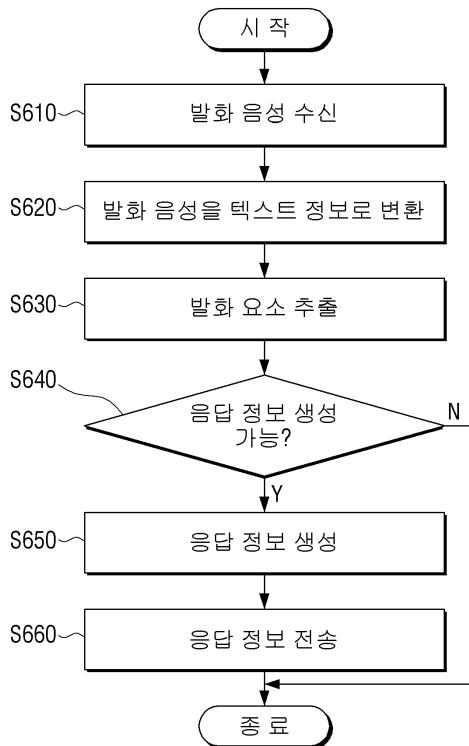
도면4



도면5



도면6



도면7

