



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0016975  
(43) 공개일자 2021년02월17일

(51) 국제특허분류(Int. Cl.)  
G10L 15/22 (2006.01) G10L 15/18 (2006.01)  
G10L 15/28 (2006.01)  
(52) CPC특허분류  
G10L 15/22 (2013.01)  
G10L 15/04 (2013.01)  
(21) 출원번호 10-2019-0095495  
(22) 출원일자 2019년08월06일  
심사청구일자 없음

(71) 출원인  
현대자동차주식회사  
서울특별시 서초구 현릉로 12 (양재동)  
기아자동차주식회사  
서울특별시 서초구 현릉로 12 (양재동)  
(72) 발명자  
이정업  
경기도 용인시 수지구 신봉2로 72 (신봉동, 신봉  
마을자이2차아파트) 210동 104호  
박영민  
경기도 군포시 송부로291번안길 3-35 (부곡동)  
503호  
김선아  
서울특별시 관악구 승방3나길 13 (남현동, 신우빌  
라) 이가채 103동 401호  
(74) 대리인  
특허법인세립

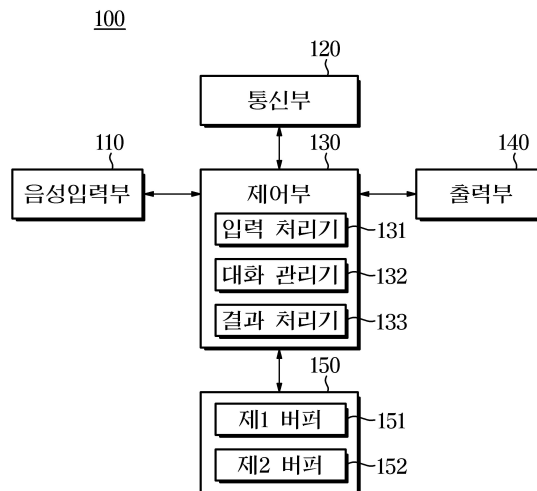
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 대화 처리 장치, 이를 포함하는 차량 및 대화 처리 방법

(57) 요약

개시된 발명의 일 실시예에 따른 대화 처리 장치는 사용자의 음성신호를 입력 받는 음성입력부; 입력된 음성신호를 저장하는 제1 버퍼; 출력부; 및 저장된 음성신호에 기초하여 발화 종료 시점을 감지하고, 발화 종료 시점 이전의 음성신호에 대응하는 제1 음성인식결과로부터 사용자의 의도 파악이 가능한지 여부에 기초하여 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하고, 제1 음성인식결과 또는 제2 음성인식결과 중 적어도 하나에 기초하여 결정된 사용자의 의도에 대응하는 응답을 출력하도록 출력부를 제어하는 제어부;를 포함한다.

대표도 - 도1



(52) CPC특허분류

*G10L 15/1822* (2013.01)

*G10L 15/28* (2013.01)

*G10L 25/87* (2013.01)

*G10L 2015/221* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

사용자의 음성신호를 입력 받는 음성입력부;

상기 입력된 음성신호를 저장하는 제1 버퍼;

출력부; 및

상기 저장된 음성신호에 기초하여 발화 종료 시점을 감지하고, 상기 발화 종료 시점 이전의 음성신호에 대응하는 제1 음성인식결과로부터 상기 사용자의 의도 파악이 가능한지 여부에 기초하여 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하고, 상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 결정된 상기 사용자의 의도에 대응하는 응답을 출력하도록 상기 출력부를 제어하는 제어부;를 포함하는 대화 처리 장치.

#### 청구항 2

제1항에 있어서,

상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 대화 처리 장치.

#### 청구항 3

제2항에 있어서,

제2버퍼;를 더 포함하고,

상기 제어부는,

상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 상기 제1 음성인식결과를 상기 제2 버퍼에 저장하는 대화 처리 장치.

#### 청구항 4

제2항에 있어서,

상기 제어부는,

상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 음성 인식 횟수에 기초하여 상기 제2 음성인식결과를 생성하는 대화 처리 장치.

#### 청구항 5

제4항에 있어서,

상기 제어부는,

상기 음성 인식 횟수가 미리 정해진 기준값 미만이면, 상기 발화 종료 시점 이후의 음성 신호에 기초하여 상기 제2 음성인식결과를 생성하는 대화 처리 장치.

#### 청구항 6

제4항에 있어서,

상기 제어부는,

상기 음성 인식 횟수가 미리 정해진 기준값 이상이면, 상기 제1 버퍼에 저장된 데이터를 삭제하고, 상기 사용자의 의도 파악이 불가능한 경우에 대응하는 응답을 생성하는 대화 처리 장치.

**청구항 7**

제1항에 있어서,

상기 제어부는,

상기 사용자의 음성 신호에 대응하는 응답이 출력되면, 음성 인식 횟수를 초기값으로 설정하는 대화 처리 장치.

**청구항 8**

제1항에 있어서,

상기 제어부는,

상기 제2 음성인식결과를 생성하면, 상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 상기 사용자의 의도를 결정하기 위한 의도 후보군을 결정하고, 상기 결정된 의도 후보군 중 어느 하나를 상기 사용자의 의도로 결정하는 대화 처리 장치.

**청구항 9**

제8항에 있어서,

상기 제어부는,

상기 의도 후보군에 대하여 정확도를 결정할 수 있고,

상기 의도 후보군 중 상기 정확도가 가장 높은 의도 후보를 상기 사용자의 의도로 결정하는 대화 처리 장치.

**청구항 10**

제1항에 있어서,

상기 제어부는,

상기 발화 종료 시점을 감지하면, 상기 제1 버퍼에 저장된 데이터를 삭제하고, 상기 발화 종료 시점 이후에 입력된 음성 신호를 상기 제1 버퍼에 저장하는 대화 처리 장치.

**청구항 11**

제1항에 있어서,

상기 제어부는,

상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 가능하면, 상기 제1 버퍼에 저장된 데이터를 삭제하는 대화 처리 장치.

**청구항 12**

사용자의 음성신호를 입력 받는 음성입력부;

상기 입력된 음성신호를 저장하는 제1 버퍼;

출력부; 및

상기 저장된 음성신호에 기초하여 발화 종료 시점을 감지하고, 상기 발화 종료 시점 이전의 음성신호에 대응하는 제1 음성인식결과로부터 상기 사용자의 의도 파악이 가능한지 여부에 기초하여 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하고, 상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 결정된 상기 사용자의 의도에 대응하는 응답을 출력하도록 상기 출력부를 제어하는 제어부;를 포함하는 차량.

**청구항 13**

제12항에 있어서,

상기 제어부는,

상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 차량.

**청구항 14**

제12항에 있어서,

상기 제어부는,

상기 제2 음성인식결과를 생성하면, 상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 상기 사용자의 의도를 결정하기 위한 의도 후보군을 결정하고, 상기 결정된 의도 후보군 중 어느 하나를 상기 사용자의 의도로 결정하는 차량.

**청구항 15**

사용자의 음성신호를 입력 받고;

상기 입력된 음성신호를 제1 버퍼에 저장하고;

상기 저장된 음성신호에 기초하여 발화 종료 시점을 감지하고;

상기 발화 종료 시점 이전의 음성신호에 대응하는 제1 음성인식결과로부터 상기 사용자의 의도 파악이 가능한지 여부에 기초하여 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하고; 및

상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 결정된 상기 사용자의 의도에 대응하는 응답을 출력하는 것;을 포함하는 대화 처리 방법.

**청구항 16**

제15항에 있어서,

상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 것은,

상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 것;을 포함하는 대화 처리 방법.

**청구항 17**

제16항에 있어서,

상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 것은,

상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 상기 제1 음성인식결과를 제2 버퍼에 저장하는 것;을 더 포함하는 대화 처리 방법.

**청구항 18**

제16항에 있어서,

상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 것은,

상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 음성 인식 횟수에 기초하여 상기 제2 음성인식결과를 생성하는 대화 처리 방법.

**청구항 19**

제18항에 있어서,

상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 것은,

상기 음성 인식 횟수가 미리 정해진 기준값 미만이면, 상기 발화 종료 시점 이후의 음성 신호에 기초하여 상기 제2 음성인식결과를 생성하는 것;을 포함하는 대화 처리 방법.

**청구항 20**

제15항에 있어서,

상기 제2 음성인식결과를 생성하면, 상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 상기 사용자의 의도를 결정하기 위한 의도 후보군을 결정하고; 및

상기 결정된 의도 후보군 중 어느 하나를 상기 사용자의 의도로 결정하는 것;을 더 포함하는 대화 처리 방법.

### 발명의 설명

#### 기술 분야

[0001] 사용자와의 대화를 통하여 사용자의 의도를 파악하고, 사용자에게 필요한 정보나 서비스를 제공하는 대화 처리 장치, 이를 포함하는 차량 및 대화 처리 방법에 관한 것이다.

#### 배경 기술

[0002] 대화 처리 장치는 사용자와의 대화를 수행하는 장치로서, 사용자의 음성을 인식하고, 음성 인식 결과를 통하여 사용자의 의도를 파악하고, 사용자에게 필요한 정보나 서비스를 제공하기 위한 응답을 출력할 수 있다.

[0003] 한편, 사용자의 음성을 인식하는데 있어서, 사용자가 실제 발화한 구간을 결정하는 것이 필요하며, 이를 위해 사용자가 발화를 종료한 지점을 검출하는 것이 중요하다. 이를 EPD(End Point Detection)라 한다.

[0004] 종래 대화 처리 장치는 일정 시간 이상의 무음 구간이 검출되면, 사용자가 발화를 종료한 것으로 인식하며, 발화를 종료한 지점까지의 음성 데이터에 기초하여 음성 인식 결과를 생성하였다. 이 경우, 사용자가 발화를 종료할 의도가 없음에도 일정 시간 이상 발화를 하지 않으면, 이후 사용자가 입력한 음성은 음성 인식 결과를 생성하는데 고려되지 않아 사용자의 의도와는 상이한 응답이 출력되는 불편이 발생하였다.

### 발명의 내용

#### 해결하려는 과제

[0005] 일 측면은 할 수 있는 대화 처리 장치, 이를 포함하는 차량 및 대화 처리 방법을 제공한다.

#### 과제의 해결 수단

[0006] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 일 측면에 따른 대화 처리 장치는, 사용자의 음성신호를 입력 받는 음성입력부; 상기 입력된 음성신호를 저장하는 제1 버퍼; 출력부; 및 상기 저장된 음성신호에 기초하여 발화 종료 지점을 감지하고, 상기 발화 종료 지점 이전의 음성신호에 대응하는 제1 음성인식결과로부터 상기 사용자의 의도 파악이 가능한지 여부에 기초하여 상기 발화 종료 지점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하고, 상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 결정된 상기 사용자의 의도에 대응하는 응답을 출력하도록 상기 출력부를 제어하는 제어부;를 포함한다.

[0007] 또한, 상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 상기 발화 종료 지점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성할 수 있다.

[0008] 또한, 제2버퍼;를 더 포함하고, 상기 제어부는, 상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 상기 제1 음성인식결과를 상기 제2 버퍼에 저장할 수 있다.

[0009] 또한, 상기 제어부는, 상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 음성 인식 횟수에 기초하여 상기 제2 음성인식결과를 생성할 수 있다.

[0010] 또한, 상기 제어부는, 상기 음성 인식 횟수가 미리 정해진 기준값 미만이면, 상기 발화 종료 지점 이후의 음성신호에 기초하여 상기 제2 음성인식결과를 생성할 수 있다.

[0011] 또한, 상기 제어부는, 상기 음성 인식 횟수가 미리 정해진 기준값 이상이면, 상기 제1 버퍼에 저장된 데이터를 삭제하고, 상기 사용자의 의도 파악이 불가능한 경우에 대응하는 응답을 생성할 수 있다.

[0012] 또한, 상기 제어부는, 상기 사용자의 음성 신호에 대응하는 응답이 출력되면, 음성 인식 횟수를 초기값으로 설정할 수 있다.

[0013] 또한, 상기 제어부는, 상기 제2 음성인식결과를 생성하면, 상기 제1 음성인식결과 또는 상기 제2 음성인식결과

중 적어도 하나에 기초하여 상기 사용자의 의도를 결정하기 위한 의도 후보군을 결정하고, 상기 결정된 의도 후보군 중 어느 하나를 상기 사용자의 의도로 결정할 수 있다.

- [0014] 또한, 상기 제어부는, 상기 의도 후보군에 대하여 정확도를 결정할 수 있고, 상기 의도 후보군 중 상기 정확도가 가장 높은 의도 후보를 상기 사용자의 의도로 결정할 수 있다.
- [0015] 또한, 상기 제어부는, 상기 발화 종료 시점을 감지하면, 상기 제1 버퍼에 저장된 데이터를 삭제하고, 상기 발화 종료 시점 이후에 입력된 음성 신호를 상기 제1 버퍼에 저장할 수 있다.
- [0016] 또한, 상기 제어부는, 상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 가능하면, 상기 제1 버퍼에 저장된 데이터를 삭제할 수 있다.
- [0017] 다른 측면에 따른 차량은, 사용자의 음성신호를 입력 받는 음성입력부; 상기 입력된 음성신호를 저장하는 제1 버퍼; 출력부; 및 상기 저장된 음성신호에 기초하여 발화 종료 시점을 감지하고, 상기 발화 종료 시점 이전의 음성신호에 대응하는 제1 음성인식결과로부터 상기 사용자의 의도 파악이 가능한지 여부에 기초하여 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하고, 상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 결정된 상기 사용자의 의도에 대응하는 응답을 출력하도록 상기 출력부를 제어하는 제어부;를 포함한다.
- [0018] 또한, 상기 제어부는, 상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성할 수 있다.
- [0019] 또한, 상기 제어부는, 상기 제2 음성인식결과를 생성하면, 상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 상기 사용자의 의도를 결정하기 위한 의도 후보군을 결정하고, 상기 결정된 의도 후보군 중 어느 하나를 상기 사용자의 의도로 결정할 수 있다.
- [0020] 다른 측면에 따른 대화 처리 방법은 사용자의 음성신호를 입력 받고; 상기 입력된 음성신호를 제1 버퍼에 저장하고; 상기 저장된 음성신호에 기초하여 발화 종료 시점을 감지하고; 상기 발화 종료 시점 이전의 음성신호에 대응하는 제1 음성인식결과로부터 상기 사용자의 의도 파악이 가능한지 여부에 기초하여 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하고; 및 상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 결정된 상기 사용자의 의도에 대응하는 응답을 출력하는 것;을 포함한다.
- [0021] 또한, 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 것은, 상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 것;을 포함할 수 있다.
- [0022] 또한, 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 것은, 상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 상기 제1 음성인식결과를 제2 버퍼에 저장하는 것;을 더 포함할 수 있다.
- [0023] 또한, 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 것은, 상기 제1 음성인식결과로부터 상기 사용자의 의도 파악이 불가능하면, 음성 인식 횟수에 기초하여 상기 제2 음성인식결과를 생성할 수 있다.
- [0024] 또한, 상기 발화 종료 시점 이후의 음성신호에 대응하는 제2 음성인식결과를 생성하는 것은, 상기 음성 인식 횟수가 미리 정해진 기준값 미만이면, 상기 발화 종료 시점 이후의 음성 신호에 기초하여 상기 제2 음성인식결과를 생성하는 것;을 포함할 수 있다.
- [0025] 또한, 상기 제2 음성인식결과를 생성하면, 상기 제1 음성인식결과 또는 상기 제2 음성인식결과 중 적어도 하나에 기초하여 상기 사용자의 의도를 결정하기 위한 의도 후보군을 결정하고; 및 상기 결정된 의도 후보군 중 어느 하나를 상기 사용자의 의도로 결정하는 것;을 더 포함할 수 있다.

**발명의 효과**

- [0026] 일 측면에 따른 대화 처리 장치, 이를 포함하는 차량 및 대화 처리 방법에 따르면, 사용자의 음성을 인식하여 대화를 처리하는데 있어서 정확성이 증대될 수 있다. 동시에 사용자의 의도를 정확하게 파악할 수 있으므로, 사용자의 편의성이 증대될 수 있다.

**도면의 간단한 설명**

- [0027] 도 1은 일 실시예에 따른 대화 처리 장치의 제어 블록도이다.
- 도 2는 일 실시예에 따른 대화 처리 장치의 동작을 설명하기 위한 도면이다.
- 도 3은 일 실시예에 따른 대화 처리 장치의 동작을 설명하기 위한 도면이다.
- 도 4는 일 실시예에 따른 대화 처리 방법의 흐름도이다.
- 도 5는 다른 실시예에 따른 대화 처리 방법의 흐름도이다.

**발명을 실시하기 위한 구체적인 내용**

- [0028] 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성요소를 지칭한다. 본 명세서가 실시예들의 모든 요소들을 설명하는 것은 아니며, 본 발명이 속하는 기술분야에서 일반적인 내용 또는 실시예들 간에 중복되는 내용은 생략한다. 명세서에서 사용되는 '부, 모듈, 부재, 블록'이라는 용어는 소프트웨어 또는 하드웨어로 구현될 수 있으며, 실시예들에 따라 복수의 '부, 모듈, 부재, 블록'이 하나의 구성요소로 구현되거나, 하나의 '부, 모듈, 부재, 블록'이 복수의 구성요소들을 포함하는 것도 가능하다.
- [0029] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 직접적으로 연결되어 있는 경우뿐 아니라, 간접적으로 연결되어 있는 경우를 포함하고, 간접적인 연결은 무선 통신망을 통해 연결되는 것을 포함한다.
- [0030] 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0031] 제 1, 제 2 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하기 위해 사용되는 것으로, 구성요소가 전술된 용어들에 의해 제한되는 것은 아니다.
- [0032] 단수의 표현은 문맥상 명백하게 예외가 있지 않는 한, 복수의 표현을 포함한다.
- [0033] 각 단계들에 있어 식별부호는 설명의 편의를 위하여 사용되는 것으로 식별부호는 각 단계들의 순서를 설명하는 것이 아니며, 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않는 이상 명기된 순서와 다르게 실시될 수 있다.
- [0034] 이하 첨부된 도면들을 참고하여 본 발명의 작용 원리 및 실시예들에 대해 설명한다.
- [0035] 도 1은 일 실시예에 따른 대화 처리 장치의 제어 블록도이다.
- [0036] 도 1을 참조하면, 일 실시예에 따른 대화 처리 장치(100)는 음성입력부(110), 통신부(120), 제어부(130), 출력부(140), 저장부(150)를 포함한다.
- [0037] 음성입력부(110)는 사용자의 명령을 음성으로 입력 받을 수 있다. 즉, 음성입력부(110)는 사용자의 음성 신호를 입력 받을 수 있다. 이를 위해, 음성입력부(110)는 음향을 입력 받아 전기적인 신호로 변환하여 출력하는 마이크로폰을 포함할 수 있다.
- [0038] 저장부(150)에는 대화 처리 장치(100)가 사용자의 음성에 대응하는 응답을 출력하기 위하여 직접적 또는 간접적으로 사용되는 각종 데이터가 저장될 수 있다.
- [0039] 또한, 저장부(150)는 제1 버퍼(151) 및 제2 버퍼(152)를 포함할 수 있다. 제1 버퍼(151)는 입력된 음성신호를 저장할 수 있고, 제2 버퍼(152)는 음성 인식 결과를 저장할 수 있다.
- [0040] 한편, 저장부는 캐쉬, ROM(Read Only Memory), PROM(Programmable ROM), EPROM(Erasable Programmable ROM), EEPROM(Electrically Erasable Programmable ROM) 및 플래시 메모리(Flash memory)와 같은 비휘발성 메모리 소자 또는 RAM(Random Access Memory)과 같은 휘발성 메모리 소자 또는 하드디스크 드라이브(HDD, Hard Disk Drive), CD-ROM과 같은 저장 매체 중 적어도 하나로 구현될 수 있으나 이에 한정되지는 않는다. 저장부(150)는 제어부(130)와 관련하여 후술할 프로세서와 별개의 칩으로 구현될 메모리일 수 있고, 프로세서와 단일 칩으로 구현될 수도 있다.
- [0041] 제어부(130)는 입력된 음성신호를 인식함으로써 음성 인식 결과를 생성하는 입력 처리기(131), 음성 인식 결과



에 기초하여 사용자의 의도를 파악하고, 이에 대응하는 액션을 결정하는 대화 관리자(132), 결정된 액션을 수행하기 위한 대화 응답을 생성하기 위한 결과 처리기(133)를 포함할 수 있다.

- [0042] 입력 처리기(131)는 입력된 사용자의 음성신호를 인식할 수 있고, 사용자의 음성신호를 텍스트 형태의 발화문으로 변환할 수 있다. 입력 처리기(131)는 이러한 발화문에 자연어 이해(Natural Language Understanding) 알고리즘을 적용하여 사용자의 의도를 파악할 수 있다.
- [0043] 이러한 텍스트 형태의 발화문 또는 사용자의 의도 중 적어도 하나는 음성 인식 결과로서 입력 처리기(131)에 의하여 출력될 수 있고, 입력 처리기(131)는 음성 인식 결과를 대화 관리자(132)로 전달할 수 있다.
- [0044] 이를 위해, 입력 처리기(131)는 음성 인식 모듈을 포함할 수 있으며, 입력된 음성을 처리하기 위한 동작을 수행하는 프로세서(미도시)로 구현될 수 있다.
- [0045] 이러한 입력 처리기(131)의 음성 처리 동작은 음성이 입력되는 것을 조건으로 수행될 수 있고, 음성이 입력되는 경우에도 특정한 조건을 만족하는 경우에만 수행될 수도 있다.
- [0046] 구체적으로, 입력 처리기(131)는 미리 정해진 호출어가 입력되거나 사용자로부터 음성 인식 시작 명령을 수신하는 경우에 전술한 음성 처리 동작을 수행할 수 있다.
- [0047] 또한, 입력 처리기(131)는 특정 구간에 대하여 입력된 음성 신호에 대응하는 음성 인식 결과를 생성할 수 있다. 구체적으로, 입력 처리기(131)는 발화 시작 시점으로부터 발화 종료 시점까지의 음성신호에 대응하는 음성인식 결과를 생성할 수 있다.
- [0048] 이를 위해, 입력 처리기(131)는 입력된 음성 신호에 기초하여 발화 시작 시점 또는 발화 종료 시점을 감지할 수 있다. 입력 처리기(131)는 미리 정해진 호출어가 입력되거나 사용자로부터 음성 인식 시작 명령을 수신하는 시점을 발화 시작 시점으로 결정할 수 있다. 이 때, 사용자는 미리 정해진 호출어를 발화하거나 별도의 버튼을 통하여 음성 인식 시작 명령을 입력할 수 있으며, 입력 처리기(131)는 발화 시작 시점부터 사용자의 음성을 인식할 수 있다.
- [0049] 또한, 입력 처리기(131)는 입력된 음성 신호에 미리 정해진 시간 이상의 무음 구간이 존재하는 경우, 사용자의 발화가 종료되었음을 확인할 수 있다. 이 때, 입력 처리기(131)는 무음 구간의 시작 시점부터 미리 정해진 시간 후의 시점을 발화 종료 시점으로 결정할 수 있다.
- [0050] 한편, 사용자가 발화를 종료할 의도가 없음에도 발화를 잠시 중단하는 경우, 발화를 잠시 중단하는 동안의 무음 구간에 의하여 사용자가 발화를 종료한 것으로 인식되고, 발화를 중단한 이후의 사용자의 음성 신호는 음성 인식 결과를 생성하는데 고려되지 않는 경우가 발생할 수 있다. 이 경우, 사용자가 발화를 중단한 이전의 음성 신호만을 기초로 사용자의 의도를 파악하기가 어려운 경우, 사용자의 정확한 의도 파악이 불가능하여 사용자의 상황에 적합한 대화 응답을 출력하기가 어려운 경우가 발생한다.
- [0051] 이를 위해, 입력 처리기(131)는 발화 종료 시점 이전의 음성 신호에 대응하는 사용자 의도 파악이 가능한지 여부에 기초하여, 발화 종료 시점 이후의 음성 신호를 인식할지 여부를 결정할 수 있다. 이에 대한 구체적인 설명은 후술한다.
- [0052] 한편, 입력 처리기(131)는 전술한 동작 외에도 입력된 음성 신호의 잡음을 제거하는 동작을 수행할 수 있고, 입력된 음성 신호에 대응하는 사용자를 인식할 수도 있다.
- [0053] 대화 관리자(132)는 입력 처리기(131)로부터 전달 받은 음성 인식 결과에 기초하여 사용자의 의도를 파악할 수 있고, 사용자의 의도에 대응되는 액션을 결정할 수 있다.
- [0054] 결과 처리기(133)는 대화 관리자(132)의 출력 결과에 따라 특정 서비스를 제공하거나 대화를 계속 이어 나가기 위한 시스템 발화를 출력할 수 있다. 결과 처리기(133)는 전달된 액션을 수행하기 위하여 필요한 대화 응답 및 명령어를 생성할 수 있고, 이러한 명령어를 출력할 수 있다. 대화 응답은 텍스트, 이미지 또는 오디오로 출력될 수 있고, 명령어가 출력되면 출력된 명령어에 대응되는 차량 제어, 외부 콘텐츠 제공 등의 서비스가 수행될 수 있다.
- [0055] 한편, 제어부(130)는 대화 처리 장치(100) 내 구성요소들의 동작을 제어하기 위한 알고리즘 또는 알고리즘을 재현한 프로그램에 대한 데이터를 저장하는 메모리(미도시), 및 메모리에 저장된 데이터를 이용하여 전술한 동작을 수행하는 프로세서(미도시)로 구현될 수 있다. 이때, 메모리와 프로세서는 각각 별개의 칩으로 구현될 수 있다. 또는, 메모리와 프로세서는 단일 칩으로 구현될 수도 있다.

- [0056] 또는, 입력 처리기(131), 대화 관리기(132), 결과 처리기(133) 각각은 하나의 프로세서에 통합되어 구현될 수 있으며, 별도의 프로세서로 구별되어 구현될 수도 있다.
- [0057] 출력부(140)는 제어부(130)에 의하여 생성된 응답을 시각적 또는 청각적으로 출력할 수 있다. 이를 위해, 출력부(140)는 디스플레이(미도시) 또는 스피커(미도시)를 포함할 수 있다. 디스플레이(미도시) 및 스피커(미도시)는 사용자의 발화에 대한 응답, 사용자에 대한 질의, 또는 사용자에게 제공할 정보 뿐만 아니라 사용자에게 대한 의사 확인, 통화 상대방의 발화에 대한 응답을 시각적 또는 청각적으로 출력할 수 있다.
- [0058] 통신부(120)는 서버 등의 외부 장치와 통신을 수행할 수 있다. 이를 위해, 통신부(120)는 외부 장치와 통신을 가능하게 하는 하나 이상의 구성 요소를 포함할 수 있으며, 예를 들어 블루투스 모듈, 적외선 통신 모듈, RFID(Radio Frequency Identification) 통신 모듈 등의 근거리 통신 모듈, 캔(Controller Area Network; CAN) 통신 모듈, 지역 통신(Local Area Network; LAN) 모듈 등의 유선 통신 모듈 및 와이파이(Wifi) 모듈, 와이브로(Wireless broadband) 모듈 등의 무선 통신 모듈 중 적어도 하나를 포함할 수 있다.
- [0059] 한편, 통신부(120)가 무선 통신 모듈을 포함하도록 구현되는 경우, 무선 통신 모듈은 신호를 송신하는 안테나 및 송신기(Transmitter)를 포함하는 무선 통신 인터페이스를 포함할 수 있다. 또한, 무선 통신 모듈은 대화시스템(100)의 제어에 따라 무선 통신 인터페이스를 통해 대화시스템(100)으로부터 출력된 디지털 제어 신호를 아날로그 형태의 무선 신호로 변조하는 신호 변환 모듈을 더 포함할 수 있다.
- [0060] 무선 통신 모듈은 신호를 수신하는 안테나 및 수신기(Receiver)를 포함하는 무선 통신 인터페이스를 포함할 수 있다. 또한, 무선 통신 모듈은 무선 통신 인터페이스를 통하여 수신한 아날로그 형태의 무선 신호를 디지털 제어 신호로 복조하기 위한 신호 변환 모듈을 더 포함할 수 있다.
- [0061] 도 1에 도시된 대화 처리 장치(100)의 구성 요소들의 성능에 대응하여 적어도 하나의 구성요소가 추가되거나 삭제될 수 있다. 또한, 구성 요소들의 상호 위치는 시스템의 성능 또는 구조에 대응하여 변경될 수 있다는 것은 당해 기술 분야에서 통상의 지식을 가진 자에게 용이하게 이해될 것이다.
- [0062] 한편, 도 1에서 도시된 각각의 구성요소는 소프트웨어 및/또는 Field Programmable Gate Array(FPGA) 및 주문형 반도체(ASIC, Application Specific Integrated Circuit)와 같은 하드웨어 구성요소를 의미한다.
- [0063] 한편, 도 1에서 도시된 대화 처리 장치(100)의 전부 또는 일부 구성요소는 차량(100)에 포함될 수 있으며, 차량(100)의 운전자 및 동승자를 포함하는 사용자의 음성을 인식하고, 적합한 응답을 출력할 수 있다.
- [0064] 도 2는 일 실시예에 따른 대화 처리 장치의 동작을 설명하기 위한 도면이다.
- [0065] 도 2를 참조하면, 일 실시예에 따른 대화 처리 장치(100)는 사용자로부터 "산타페 ... 연비 알려줘"라는 음성 신호(Sin)를 입력 받을 수 있다. 이 경우, 사용자가 "산타페"를 발화한 이후 발화를 잠시 중단하고, "연비 알려줘"를 발화한 경우를 가정한다.
- [0066] 입력 처리기(131)는 사용자가 미리 정해진 호출어를 발화하거나 음성 인식 시작 명령을 입력한 시점(t1)을 발화 시작 시점으로 결정할 수 있고, 발화 시작 시점 이후 입력된 음성 신호를 제1 버퍼(151)에 저장할 수 있다.
- [0067] 입력 처리기(131)는 제1 버퍼(151)에 저장된 음성 신호에 기초하여 발화 종료 시점을 감지할 수 있다.
- [0068] 구체적으로, 저장된 음성 신호에 사용자가 음성 입력을 중단한 시점(t2)부터 미리 정해진 시간 이후의 시점(t3)까지의 무음 구간(TS1)이 존재하면, 입력 처리기(131)는 무음 구간(TS1)이 종료되는 시점을 발화 종료 시점으로 결정할 수 있다.
- [0069] 발화 종료 시점이 감지되면, 입력 처리기(131)는 발화 종료 시점 이전 구간(이하, 제1 구간(A)이라 한다)에 입력된 음성 신호를 인식할 수 있고, 이에 대한 음성 인식 결과로서 "산타페"라는 발화문(X[i])을 생성할 수 있다.
- [0070] 이 때, 입력 처리기(131)는 제1 버퍼(151)에 저장된 제1 구간(A)의 음성 신호를 삭제함으로써 제1 버퍼(151)를 초기화할 수 있다. 이와 함께, 입력 처리기(131)는 제1 구간(A)의 발화 종료 시점 이후에 입력되는 음성 신호를 제1 버퍼(151)에 저장할 수 있다.
- [0071] 또한, 입력 처리기(131)는 제1 구간(A)에 입력된 음성 신호에 대응하는 음성 인식 결과에 따라 사용자의 의도 파악이 가능한지 여부에 기초하여 발화 종료 시점 이후의 음성 신호에 대응하는 음성 인식 결과를 생성할 수 있다.

- [0072] 구체적으로, 제1 구간(A)에 입력된 음성 신호만으로 사용자의 의도 파악이 불가능한 경우, 입력 처리기(131)는 발화 종료 시점 이후의 구간(이하, 제2 구간(B)이라 한다)에서 입력된 음성 신호에 대응하는 음성 인식 결과를 생성할 수 있다.
- [0073] 이 때, 입력 처리기(131)는 제1 구간(A)의 음성 인식 결과를 제2 버퍼(152)에 저장할 수 있다.
- [0074] 이후, 또다른 무음 구간(TS2)이 존재함에 따라 제2 구간(B)에 대한 발화 종료 시점이 감지되면, 입력 처리기(131)는 제1 버퍼(151)에 저장된 제2 구간(B)에 입력된 음성 신호를 인식할 수 있고, 이에 대한 음성 인식 결과로서 "연비 알려줘"라는 발화문(X[i+1])을 생성할 수 있다.
- [0075] 이 때, 입력 처리기(131)는 제1 버퍼(151)에 저장된 제2 구간(B)의 음성 신호를 삭제함으로써 제1 버퍼(151)를 초기화할 수 있다. 이와 함께, 입력 처리기(131)는 제2 구간(B)의 발화 종료 시점 이후에 입력되는 음성 신호를 제1 버퍼(151)에 저장할 수 있다.
- [0076] 또한, 입력 처리기(131)는 제1 구간(A)의 음성 인식 결과 또는 제2 구간(B)의 음성 인식 결과 중 적어도 하나에 기초하여 사용자의 의도 파악이 가능한지 여부에 기초하여 발화 종료 시점 이후의 음성 신호에 대응하는 음성 인식 결과를 생성할 수 있다.
- [0077] 이 때, 제2 구간(B)에 대응하는 음성 인식 결과인 "연비 알려줘"라는 발화문(X[i+1])만으로는 사용자의 의도 파악이 불가능한 경우, 입력 처리기(131)는 제1 구간(A)의 음성 인식 결과인 "산타페"라는 발화문(X[i])과 제2 구간(B)의 음성 인식 결과 "연비 알려줘"라는 발화문(X[i+1])을 결합할 수 있고, 결합된 음성 인식 결과에 기초하여 사용자의 의도를 파악할 수 있다.
- [0078] 만약, 제1 구간(A)의 음성 인식 결과 및 제2 구간(B)의 음성 인식 결과의 결합에도 불구하고 사용자의 의도 파악이 불가능한 경우, 입력 처리기(131)는 제2 구간(B)의 발화 종료 시점 이후의 구간에서 입력된 음성 신호에 대응하는 음성 인식 결과를 생성할 수 있다. 이하 전술한 동작을 반복한다.
- [0079] 한편, 적어도 하나의 구간의 음성 인식 결과에 기초하여 사용자의 의도 파악이 가능한 경우, 입력 처리기(131)는 결정된 사용자의 의도를 대화 관리기(132)에 전달할 수 있다. 대화 관리기(132)는 사용자 의도에 대응하는 액션을 결정하고, 결과 처리기(133)는 대화 관리기(132)로부터 전달된 액션을 수행하기 위한 대화 응답을 생성할 수 있다. 이러한 결과 처리기(133)에 의한 응답은 출력부(140)에 의하여 출력될 수 있다.
- [0080] 또한, 적어도 하나의 구간의 음성 인식 결과에 기초하여 사용자의 의도 파악이 가능한 경우, 입력 처리기(131)는 음성 입력부(110)를 폐쇄 상태로 전환함으로써 입력되는 음성 신호를 차단할 수 있다. 또한, 입력 처리기(131)는 제1 버퍼(151) 또는 제2 버퍼(152) 중 적어도 하나에 저장된 데이터를 삭제함으로써 제1 버퍼(151) 또는 제2 버퍼(152) 중 적어도 하나를 초기화할 수 있다.
- [0081] 이처럼, 입력 처리기(131)는 사용자 의도 파악 가능 여부에 따라 발화 종료 시점 이후의 음성 신호에 대한 처리 동작을 수행할 수 있고, 발화 종료 시점 이후의 음성 인식 결과 뿐만 아니라 발화 종료 시점 이전의 음성 인식 결과를 사용자 의도 파악 동작의 제어 기초로 활용할 수 있다. 따라서, 정확한 사용자의 의도 파악이 가능하며, 사용자에게 적합한 응답의 출력이 가능함과 동시에 사용자의 편의성이 증대될 수 있다.
- [0082] 도 3은 다른 실시예에 따른 대화 처리 장치의 동작을 설명하기 위한 도면이다.
- [0083] 도 3을 참조하면, 일 실시예에 따른 대화 처리 장치(100)는 사용자로부터 "산타페 ... 아니, 소나타 연비 알려줘"라는 음성 신호(Sin')를 입력 받을 수 있다. 이 경우, 사용자가 "산타페"를 발화한 이후 발화를 잠시 중단하고, "아니, 소나타 연비 알려줘"를 발화한 경우를 가정한다.
- [0084] 도 2에서 전술한 바와 같이, 입력 처리기(131)는 사용자가 미리 정해진 호출어를 발화하거나 음성 인식 시작 명령을 입력한 시점(t1')을 발화 시작 시점으로 결정할 수 있고, 발화 시작 시점 이후 입력된 음성 신호를 제1 버퍼(151)에 저장할 수 있다.
- [0085] 입력 처리기(131)는 제1 버퍼(151)에 저장된 음성 신호에 기초하여 발화 종료 시점을 감지할 수 있다. 구체적으로, 저장된 음성 신호에 사용자가 음성 입력을 중단한 시점(t2')부터 미리 정해진 시간 이후의 시점(t3')까지의 무음 구간(TS3)이 존재하면, 입력 처리기(131)는 무음 구간(TS3)이 종료되는 시점을 발화 종료 시점으로 결정할 수 있다.
- [0086] 발화 종료 시점이 감지되면, 입력 처리기(131)는 발화 종료 시점 이전 구간(이하, 제1 구간(C)이라 한다)에 입력된 음성 신호를 인식할 수 있고, 이에 대한 음성 인식 결과로서 "산타페"라는 발화문(X[I'])을 생성할 수 있

다.

- [0087] 이 때, 입력 처리기(131)는 제1 버퍼(151)에 저장된 제1 구간(C)의 음성 신호를 삭제함으로써 제1 버퍼(151)를 초기화할 수 있다. 이와 함께, 입력 처리기(131)는 제1 구간(C)의 발화 종료 시점 이후에 입력되는 음성 신호를 제1 버퍼(151)에 저장할 수 있다.
- [0088] 또한, 입력 처리기(131)는 제1 구간(C)에 입력된 음성 신호에 대응하는 음성 인식 결과에 따라 사용자의 의도 파악이 가능한지 여부에 기초하여 발화 종료 시점 이후의 음성 신호에 대응하는 음성 인식 결과를 생성할 수 있다.
- [0089] 구체적으로, 제1 구간(C)에 입력된 음성 신호만으로 사용자의 의도 파악이 불가능한 경우, 입력 처리기(131)는 발화 종료 시점 이후의 구간(이하, 제2 구간(D)이라 한다)에서 입력된 음성 신호에 대응하는 음성 인식 결과를 생성할 수 있다.
- [0090] 이 때, 입력 처리기(131)는 제1 구간(C)의 음성 인식 결과를 제2 버퍼(152)에 저장할 수 있다.
- [0091] 이후, 또다른 무음 구간(TS4)이 존재함에 따라 제2 구간(D)에 대한 발화 종료 시점이 감지되면, 입력 처리기(131)는 제1 버퍼(151)에 저장된 제2 구간(D)에 입력된 음성 신호를 인식할 수 있고, 이에 대한 음성 인식 결과로서 "아니, 소나타 연비 알려줘"라는 발화문(X[I'+1])을 생성할 수 있다.
- [0092] 이 때, 입력 처리기(131)는 제1 버퍼(151)에 저장된 제2 구간(D)의 음성 신호를 삭제함으로써 제1 버퍼(151)를 초기화할 수 있다. 이와 함께, 입력 처리기(131)는 제2 구간(D)의 발화 종료 시점 이후에 입력되는 음성 신호를 제1 버퍼(151)에 저장할 수 있다.
- [0093] 또한, 입력 처리기(131)는 제1 구간(C)의 음성 인식 결과 또는 제2 구간(D)의 음성 인식 결과 중 적어도 하나에 기초하여 사용자의 의도 파악이 가능한지 여부에 기초하여 발화 종료 시점 이후의 음성 신호에 대응하는 음성 인식 결과를 생성할 수 있다.
- [0094] 이 때, 제2 구간(D)에 대응하는 음성 인식 결과인 "아니, 소나타 연비 알려줘"라는 발화문(X[I'+1])만으로 사용자의 의도 파악이 가능한 경우, 입력 처리기(131)는 대화 관리기(132)로 사용자의 의도를 전달할 수 있고, 대화 관리기(132)를 거쳐 결과 처리기(133)에 의하여 대화 응답이 생성될 수 있다.
- [0095] 또는, 입력 처리기(131)는 제2 구간(D)의 음성 인식 결과인 "아니, 소나타 연비 알려줘"라는 발화문(X[I'+1]) 또는 제1 구간(C)의 음성 인식 결과인 "싼타페"라는 발화문(X[I']) 중 적어도 하나에 기초하여 사용자의 의도를 결정하기 위한 의도 후보군을 결정할 수 있다. 입력 처리기(131)는 이러한 의도 후보군 중 어느 하나를 사용자의 의도로 결정할 수 있다.
- [0096] 구체적으로, 입력 처리기(131)는 의도 후보군에 대하여 정확도를 결정할 수 있고, 의도 후보군 중 정확도가 가장 높은 의도 후보를 사용자의 의도로 결정할 수 있다. 이 때, 의도 후보군에 대한 정확도는 확률값으로 산출될 수 있으며, 입력 처리기(131)는 확률값이 가장 높은 의도 후보를 사용자의 의도로 결정할 수 있다.
- [0097] 예를 들어, 입력 처리기(131)는 제2 구간(D)의 음성 인식 결과인 "아니, 소나타 연비 알려줘"라는 발화문(X[I'+1])에 기초하여 제1 의도 후보를 결정할 수 있다. 또한, 입력 처리기(131)는 제1 구간(C)과 제2 구간(D)의 음성 인식 결과를 결합한 "싼타페, 아니 소나타 연비 알려줘"라는 결과값에 기초하여 제2 의도 후보를 결정할 수 있다. 입력 처리기(131)는 제1 의도 후보 및 제2 의도 후보 중 정확도가 가장 높은 의도 후보를 사용자의 의도로 결정할 수 있다.
- [0098] 이후, 입력 처리기(131)는 결정된 사용자의 의도를 대화 관리기(132)에 전달함으로써 대화 관리기(132), 결과 처리기(133)를 거쳐 사용자의 의도에 부합하는 대화 응답을 생성할 수 있다.
- [0099] 또한, 사용자의 의도가 결정되면, 입력 처리기(131)는 음성 입력부(110)를 폐쇄 상태로 전환함으로써 입력되는 음성 신호를 차단할 수 있다. 또한, 입력 처리기(131)는 제1 버퍼(151) 또는 제2 버퍼(152) 중 적어도 하나에 저장된 데이터를 삭제함으로써 제1 버퍼(151) 또는 제2 버퍼(152) 중 적어도 하나를 초기화할 수 있다.
- [0100] 이처럼, 입력 처리기(131)는 발화 종료 시점으로 구분되는 적어도 하나의 음성 인식 구간의 음성 인식 결과를 조합하여 사용자의 의도 후보군을 결정하고, 정확도를 기초로 사용자의 최종 의도를 결정할 수 있다. 따라서, 정확한 사용자의 의도 파악이 가능하며, 사용자에게 적합한 응답의 출력이 가능함과 동시에 사용자의 편의성이 증대될 수 있다.



- [0101] 도 4는 일 실시예에 따른 대화 처리 방법의 흐름도이다.
- [0102] 도 4를 참조하면, 일 실시예에 따른 대화 처리 장치(100)는 호출 명령이 인식되는지 여부를 확인할 수 있다(401). 이 때, 호출 명령은 미리 정해진 호출어로 설정될 수 있으며, 사용자는 미리 정해진 호출어를 발화하거나 버튼의 조작 등의 음성 인식 시작 명령을 입력함으로써 호출 명령을 내릴 수 있다.
- [0103] 호출 명령이 인식되는 경우(401의 예), 대화 처리 장치(100)는 입력된 음성 신호를 제1 버퍼(151)에 저장할 수 있다(402). 이 때, 대화 처리 장치(100)는 실시간으로 입력된 음성 신호를 제1 버퍼(151)에 저장할 수 있다.
- [0104] 대화 처리 장치(100)는 제1 버퍼에 저장된 음성 신호에 기초하여 음성 인식 결과를 생성할 수 있다(403).
- [0105] 대화 처리 장치(100)는 발화 종료 시점이 감지되는지 여부를 확인할 수 있다(404). 구체적으로, 저장된 음성 신호에 사용자가 음성 입력을 중단한 시점부터 미리 정해진 시간 이후의 시점까지의 무음 구간이 존재하면, 대화 처리 장치(100)는 무음 구간이 종료되는 시점을 발화 종료 시점으로 결정할 수 있다.
- [0106] 발화 종료 시점이 감지되는 경우(404의 예), 대화 처리 장치(100)는 제1 버퍼(151)에 저장된 음성 신호(n번째 음성 인식 구간에서 입력된 음성 신호)를 삭제함으로써 제1 버퍼(151)를 초기화할 수 있고(405), 발화 종료 시점 이후 입력된 음성 신호(n+1번째 음성 인식 구간에서 입력된 음성 신호)를 제1 버퍼(151)에 저장할 수 있다(406).
- [0107] 이후, 대화 처리 장치(100)는 발화 종료 시점 이전의 음성 신호에 기초하여 생성된 음성 인식 결과를 이용하여 사용자의 의도 파악이 가능한지 여부를 확인할 수 있다(407). 이 때, 음성 인식 결과는 단계 403에서 생성된 음성 인식 결과, 즉 n번째 음성 인식 구간에서 입력된 음성 신호에 대한 음성 인식 결과를 의미할 수 있다.
- [0108] 사용자의 의도 파악이 불가능한 경우(407의 아니오), 대화 처리 장치(100)는 생성된 음성 인식 결과를 제2 버퍼(152)에 저장할 수 있다(410). 즉, 대화 처리 장치(100)는 단계 403에서 생성된 음성 인식 결과(n번째 음성 인식 구간에서 입력된 음성 신호에 대한 음성 인식 결과)를 제2 버퍼(152)에 저장할 수 있다. 이후, 대화 처리 장치(100)는 제1 버퍼(151)에 저장된 음성 신호(n+1번째 음성 인식 구간에서 입력된 음성 신호)에 기초하여 음성 인식 결과를 생성할 수 있고(403), 제1 버퍼(151)에 저장된 음성 신호는 발화 종료 시점 이후 입력된 음성 신호(즉, n+1번째 음성 인식 구간에서 입력된 음성 신호)를 의미할 수 있다.
- [0109] 이후, 대화 처리 장치(100)는 n+1번째 음성 인식 구간의 발화 종료 시점이 감지되는지 여부를 확인할 수 있고(404), 발화 종료 시점이 감지되면(404의 예), 전술한 405, 406 단계를 수행할 수 있다. 이후, 대화 처리 장치(100)는 n+1번째 음성 인식 구간의 음성 인식 결과 또는 제2 버퍼(152)에 저장된 n번째 음성 인식 구간의 음성 인식 결과 중 적어도 하나에 기초하여 사용자의 의도 파악이 가능한지 여부를 확인할 수 있다(407). 이하, 전술한 과정을 반복할 수 있다.
- [0110] 다른 예로, 사용자의 의도 파악이 가능한 경우(407의 예), 대화 처리 장치(100)는 음성 신호 입력을 차단하고 제1 버퍼(151)를 초기화할 수 있다(408). 구체적으로, 대화 처리 장치(100)는 음성 입력부(110)를 폐쇄 상태로 전환함으로써 입력되는 음성 신호를 차단할 수 있고, 제1 버퍼(151)에 저장된 데이터를 삭제함으로써 제1 버퍼(151)를 초기화할 수 있다. 이 때, 대화 처리 장치(100)는 제2 버퍼(152)에 저장된 데이터를 삭제함으로써 제2 버퍼(152)를 초기화할 수 있다.
- [0111] 대화 처리 장치(100)는 사용자 의도에 대응하는 응답을 생성하고, 출력할 수 있다(409).
- [0112] 이처럼, 입력 처리기(131)는 사용자 의도 파악 가능 여부에 따라 발화 종료 시점 이후의 음성 신호에 대한 처리 동작을 수행할 수 있고, 발화 종료 시점 이후의 음성 인식 결과 뿐만 아니라 발화 종료 시점 이전의 음성 인식 결과를 사용자 의도 파악 동작의 제어 기초로 활용할 수 있다. 따라서, 정확한 사용자의 의도 파악이 가능하며, 사용자에게 적합한 응답의 출력이 가능함과 동시에 사용자의 편의성이 증대될 수 있다.
- [0113] 한편, 도 4에서는 403 단계 이후 404 단계가 수행되는 것으로 서술하였으나, 403 단계와 404 단계가 동시에 수행되는 것이 가능하고, 404 단계 이후 403 단계가 수행되는 것 또한 가능하다. 다만, 404단계 이후 403 단계가 수행되는 경우는 발화 종료 시점이 감지되면(404의 예), 제1 버퍼(151)에 저장된 음성 신호에 기초하여 음성 인식 결과를 생성(403)하는 경우를 의미한다.
- [0114] 도 5는 다른 실시예에 따른 대화 처리 방법의 흐름도이다.
- [0115] 도 5를 참조하면, 일 실시예에 따른 대화 처리 장치(100)는 호출 명령이 인식되는지 여부를 확인할 수 있고(501), 호출 명령이 인식되는 경우(501의 예), 입력된 음성 신호를 제1 버퍼(151)에 저장할 수 있다(502). 이

때, 대화 처리 장치(100)는 실시간으로 입력된 음성 신호를 제1 버퍼(151)에 저장할 수 있다.

- [0116] 대화 처리 장치(100)는 제1 버퍼에 저장된 음성 신호에 기초하여 음성 인식 결과를 생성할 수 있다(503).
- [0117] 대화 처리 장치(100)는 발화 종료 시점이 감지되는지 여부를 확인할 수 있다(504).
- [0118] 발화 종료 시점이 감지되는 경우(504의 예), 대화 처리 장치(100)는 제1 버퍼(151)에 저장된 음성 신호(n번째 음성 인식 구간에서 입력된 음성 신호)를 삭제함으로써 제1 버퍼(151)를 초기화할 수 있고(505), 발화 종료 시점 이후 입력된 음성 신호(n+1번째 음성 인식 구간에서 입력된 음성 신호)를 제1 버퍼(151)에 저장할 수 있다(506).
- [0119] 이후, 대화 처리 장치(100)는 발화 종료 시점 이전의 음성 신호에 기초하여 생성된 음성 인식 결과를 이용하여 사용자의 의도 파악이 가능한지 여부를 확인할 수 있다(507). 이 때, 음성 인식 결과는 단계 503에서 생성된 음성 인식 결과, 즉 n번째 음성 인식 구간에서 입력된 음성 신호에 대한 음성 인식 결과를 의미할 수 있다.
- [0120] 사용자의 의도 파악이 불가능한 경우(507의 아니오), 대화 처리 장치(100)는 음성 인식 결과에 결과값이 존재하는지 여부를 확인할 수 있다(510).
- [0121] 이 때, 대화 처리 장치(100)는 음성 인식 결과에 발화문이 존재하는 경우, 결과값이 존재하는 것으로 확인할 수 있다. 즉, 사용자가 발화를 하지 않는 경우 등으로 음성 인식 결과에 발화문이 생성되지 않은 경우, 대화 처리 장치(100)는 음성 인식 결과에 결과값이 존재하지 않는 것으로 확인할 수 있다.
- [0122] 음성 인식 결과에 결과값이 존재하는 경우(510의 예), 대화 처리 장치(100)는 생성된 음성 인식 결과를 제2 버퍼(152)에 저장할 수 있다(510). 즉, 대화 처리 장치(100)는 단계 503에서 생성된 음성 인식 결과(n번째 음성 인식 구간에서 입력된 음성 신호에 대한 음성 인식 결과)를 제2 버퍼(152)에 저장할 수 있다. 이후, 대화 처리 장치(100)는 제1 버퍼(151)에 저장된 음성 신호(n+1번째 음성 인식 구간에서 입력된 음성 신호)에 기초하여 음성 인식 결과를 생성할 수 있고(503), 제1 버퍼(151)에 저장된 음성 신호는 발화 종료 시점 이후 입력된 음성 신호(즉, n+1번째 음성 인식 구간에서 입력된 음성 신호)를 의미할 수 있다.
- [0123] 이후, 대화 처리 장치(100)는 n+1번째 음성 인식 구간의 발화 종료 시점이 감지되는지 여부를 확인할 수 있고(504), 발화 종료 시점이 감지되면(504의 예), 전술한 505, 506 단계를 수행할 수 있다. 이후, 대화 처리 장치(100)는 n+1번째 음성 인식 구간의 음성 인식 결과 또는 제2 버퍼(152)에 저장된 n번째 음성 인식 구간의 음성 인식 결과 중 적어도 하나에 기초하여 사용자의 의도 파악이 가능한지 여부를 확인할 수 있다(407). 이하, 전술한 과정을 반복할 수 있다.
- [0124] 다른 예로, 음성 인식 결과에 결과값이 존재하지 않는 경우(510의 아니오), 대화 처리 장치(100)는 음성 인식 횟수가 기준값 이상인지 여부를 확인할 수 있다(512). 이 때, 음성 인식 횟수는 음성 인식 결과를 생성한 횟수를 의미할 수 있다. 또한, 음성 인식 횟수의 기준값은 저장부(150)의 저장 용량을 고려한 최대 음성 인식 가능 횟수를 의미할 수 있다.
- [0125] 음성 인식 횟수가 기준값 미만이면(512의 아니오), 대화 처리 장치(100)는 제1 버퍼(151)에 저장된 음성 신호에 기초하여 음성 인식 결과를 생성할 수 있고(503), 제1 버퍼(151)에 저장된 음성 신호는 발화 종료 시점 이후 입력된 음성 신호(즉, n+1번째 음성 인식 구간에서 입력된 음성 신호)를 의미할 수 있다. 이후, 전술한 과정을 반복할 수 있다.
- [0126] 다른 예로, 음성 인식 횟수가 기준값 이상이거나(512의 예) 사용자의 의도 파악이 가능한 경우(507의 예), 대화 처리 장치(100)는 음성 신호 입력을 차단하고 제1 버퍼(151)를 초기화할 수 있다(408). 구체적으로, 대화 처리 장치(100)는 음성 입력부(110)를 폐쇄 상태로 전환함으로써 입력되는 음성 신호를 차단할 수 있고, 제1 버퍼(151)에 저장된 데이터를 삭제함으로써 제1 버퍼(151)를 초기화할 수 있다. 이 때, 대화 처리 장치(100)는 제2 버퍼(152)에 저장된 데이터를 삭제함으로써 제2 버퍼(152)를 초기화할 수 있다.
- [0127] 대화 처리 장치(100)는 사용자 의도에 대응하는 응답을 생성하고, 출력할 수 있다(509). 이 때, 대화 처리 장치(100)는 음성 인식 횟수를 초기값으로 설정할 수 있다. 음성 인식 횟수가 초기값으로 설정된 이후에 호출 명령이 인식되는 경우, 대화 처리 장치(100)는 1번째 음성 인식 구간에 대한 음성 인식 결과를 생성할 수 있다.
- [0128] 이처럼, 입력 처리기(131)는 사용자 의도 파악 가능 여부에 따라 발화 종료 시점 이후의 음성 신호에 대한 처리 동작을 수행할 수 있고, 발화 종료 시점 이후의 음성 인식 결과 뿐만 아니라 발화 종료 시점 이전의 음성 인식 결과를 사용자 의도 파악 동작의 제어 기초로 활용할 수 있다. 따라서, 정확한 사용자의 의도 파악이 가능하며,

사용자에게 적합한 응답의 출력이 가능함과 동시에 사용자의 편의성이 증대될 수 있다.

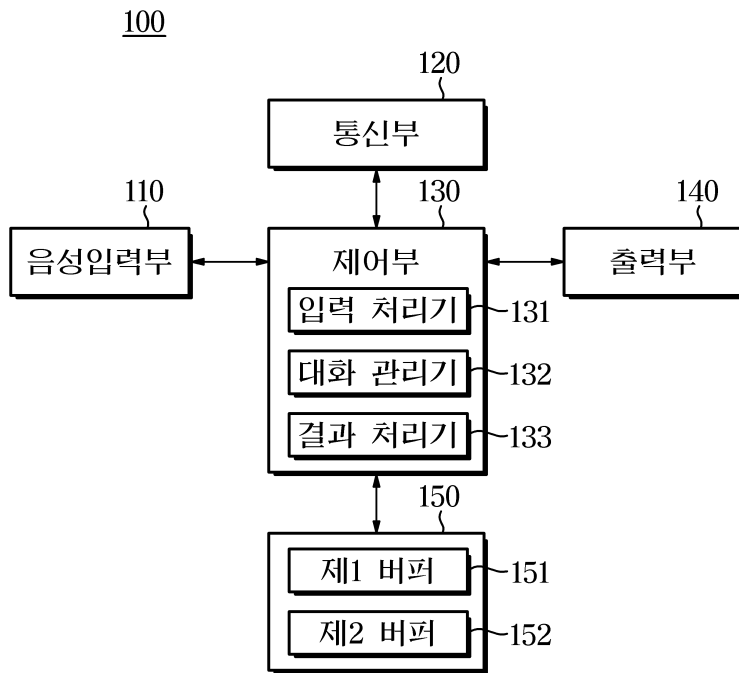
- [0129] 또한, 음성 인식 횟수에 따라 사용자가 입력할 음성에 대한 음성 인식 결과를 생성하므로, 저장 용량을 고려한 효율적인 음성 인식을 수행할 수 있다.
- [0130] 한편, 도 5에서는 503 단계 이후 504 단계가 수행되는 것으로 서술하였으나, 503 단계와 504 단계가 동시에 수행되는 것이 가능하고, 504 단계 이후 503 단계가 수행되는 것 또한 가능하다. 다만, 504단계 이후 503 단계가 수행되는 경우는 발화 종료 시점이 감지되면(504의 예), 제1 버퍼(151)에 저장된 음성 신호에 기초하여 음성 인식 결과를 생성(503)하는 경우를 의미한다.
- [0131] 한편, 개시된 실시예들은 컴퓨터에 의해 실행 가능한 명령어를 저장하는 기록매체의 형태로 구현될 수 있다. 명령어는 프로그램 코드의 형태로 저장될 수 있으며, 프로세서에 의해 실행되었을 때, 프로그램 모듈을 생성하여 개시된 실시예들의 동작을 수행할 수 있다. 기록매체는 컴퓨터로 읽을 수 있는 기록매체로 구현될 수 있다.
- [0132] 컴퓨터가 읽을 수 있는 기록매체로는 컴퓨터에 의하여 해독될 수 있는 명령어가 저장된 모든 종류의 기록 매체를 포함한다. 예를 들어, ROM(Read Only Memory), RAM(Random Access Memory), 자기 테이프, 자기 디스크, 플래쉬 메모리, 광 데이터 저장장치 등이 있을 수 있다.
- [0133] 이상에서와 같이 첨부된 도면을 참조하여 개시된 실시예들을 설명하였다. 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고도, 개시된 실시예들과 다른 형태로 본 발명이 실시될 수 있음을 이해할 것이다. 개시된 실시예들은 예시적인 것이며, 한정적으로 해석되어서는 안 된다.
- [0134] 이상에서와 같이 첨부된 도면을 참조하여 개시된 실시예들을 설명하였다. 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고도, 개시된 실시예들과 다른 형태로 본 발명이 실시될 수 있음을 이해할 것이다. 개시된 실시예들은 예시적인 것이며, 한정적으로 해석되어서는 안 된다.

**부호의 설명**

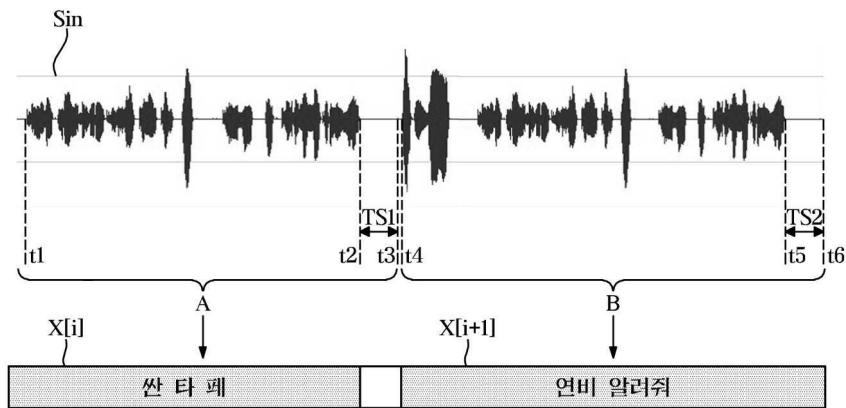
- [0135] 100: 대화 처리 장치
- 110: 음성 입력부
- 120: 통신부
- 130: 제어부
- 131: 입력 처리기
- 132: 대화 관리기
- 133: 결과 처리기
- 140: 출력부
- 150: 저장부
- 151: 제1 버퍼
- 152: 제2 버퍼
- 200: 차량

도면

도면1

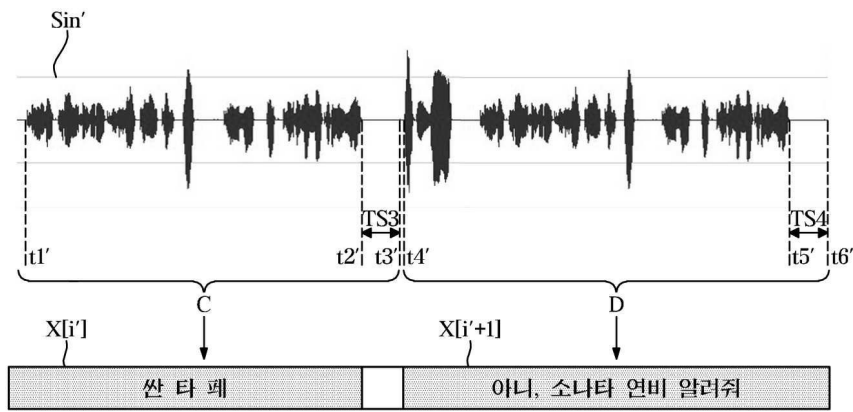


도면2

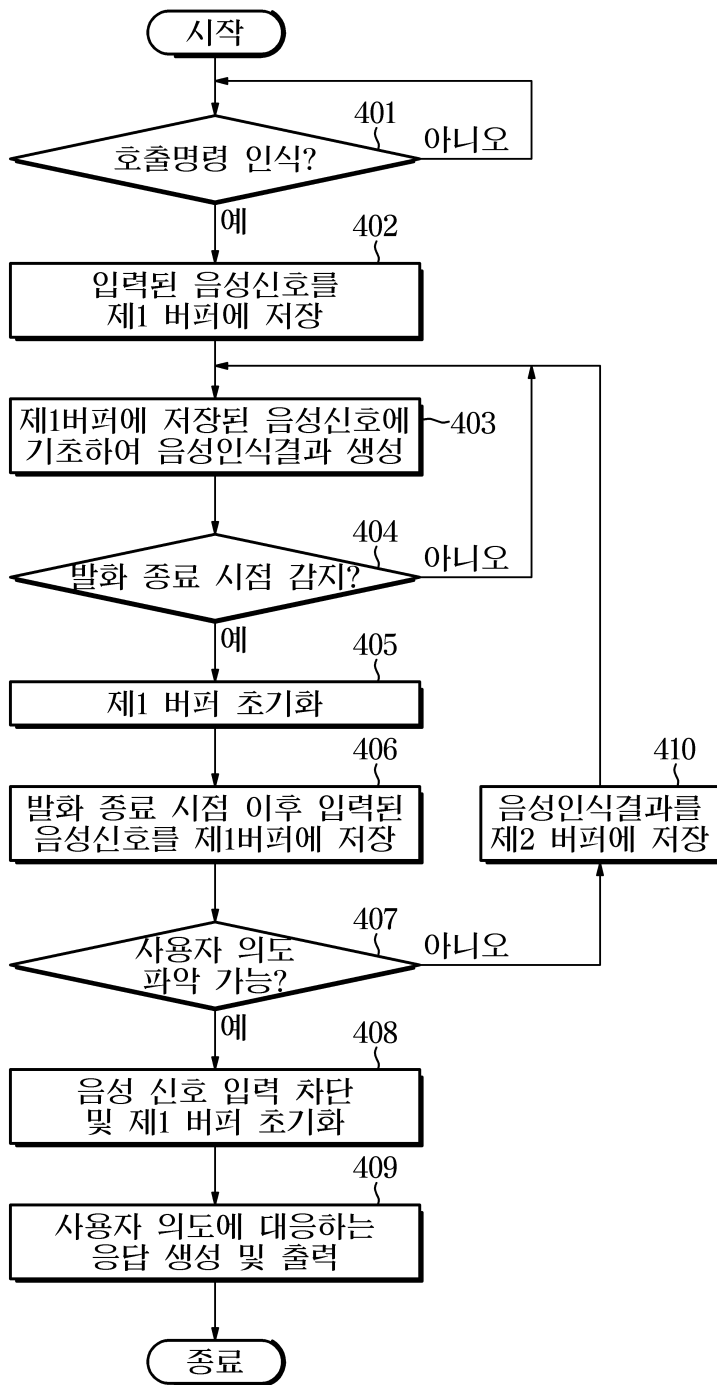




도면3



도면4



도면5

