



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년05월18일
 (11) 등록번호 10-1622111
 (24) 등록일자 2016년05월12일

(51) 국제특허분류(Int. Cl.)
 G10L 15/22 (2006.01) G10L 15/183 (2013.01)
 G10L 15/24 (2006.01)
 (21) 출원번호 10-2009-0122975
 (22) 출원일자 2009년12월11일
 심사청구일자 2014년11월19일
 (65) 공개번호 10-2011-0066357
 (43) 공개일자 2011년06월17일
 (56) 선행기술조사문헌
 US20070050191 A1*
 KR1020040050928 A*
 KR1020060094788 A*
 *는 심사관에 의하여 인용된 문헌

(73) 특허권자
 삼성전자 주식회사
 경기도 수원시 영통구 삼성로 129 (매탄동)
 (72) 발명자
 장준원
 경기도 용인시 기흥구 구성로 105-15, 동일하이빌 116동 1501호 (언남동)
 한우섭
 경기도 용인시 수지구 수지로113번길 15, 신LG2차 빌리지 221동 1804호 (성북동)
 (74) 대리인
 특허법인세림

전체 청구항 수 : 총 22 항

심사관 : 정성윤

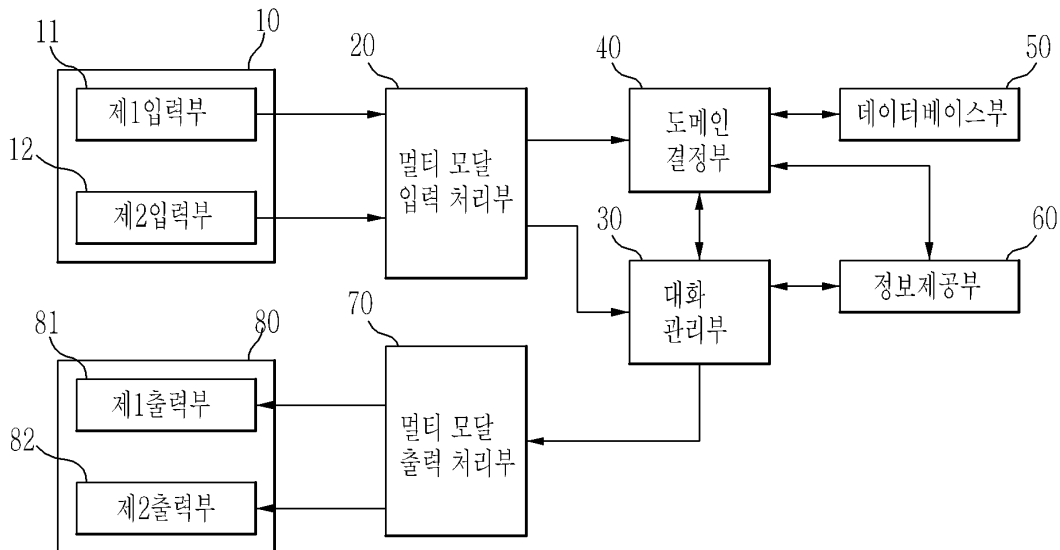
(54) 발명의 명칭 **대화 시스템 및 그의 대화 방법**

(57) 요약

본 발명은 사용자의 음성과 비음성 정보가 입력되는 멀티 모달 입력부; 적어도 하나의 음성 및 비음성 정보가 조합되어 미리 저장된 복수 시추에이션 케이스를 가지고, 멀티 모달 입력부로부터 입력된 음성 및 비음성 정보를 기초로 미리 저장된 복수 시추에이션 케이스의 적합도를 각각 산출하여 현재 도메인을 결정하는 도메인 결정부;

(뒷면에 계속)

대표도 - 도1



현재 도메인에 대응하는 응답을 선택하는 대화 관리부; 응답을 출력하는 멀티 모달 출력부를 포함한다.

본 발명은 도메인 결정에 반영된 복수 정보 조합으로 이루어진 시추에이션 케이스와 현재 정보, 음성 인식 결과를 이용하여 도메인을 결정함으로써 도메인 결정 정확도를 높일 수 있어, 이에 따라 사용자와의 자연스러운 대화를 수행할 수 있다. 또한 도메인 결정 정확도를 상승시켜 사용자의 발화 의도를 파악하기 위한 검색 공간을 줄이기 때문에 사용자의 의도를 보다 정확하게 파악할 수 있어 사용자 발화 의도에 대응하는 적절한 서비스를 수행할 수 있고, 이에 따라 사용자의 서비스 만족도를 높일 수 있다.

명세서

청구범위

청구항 1

사용자의 음성 정보 및 비음성 정보가 입력되는 멀티 모달 입력부;

상기 입력된 음성 정보 및 비음성 정보를 조합하여 시추에이션 케이스를 생성하고, 상기 생성된 시추에이션 케이스와 미리 저장된 복수 시추에이션 케이스를 각각 비교하여 상기 복수 시추에이션 케이스 각각의 적합도를 산출하고, 상기 산출된 적합도에 기초하여 현재 도메인을 결정하는 도메인 결정부;

상기 현재 도메인에 대응하는 응답을 선택하는 대화 관리부; 및

상기 응답을 출력하는 멀티 모달 출력부를 포함하되,

상기 미리 저장된 복수 시추에이션 케이스 각각은 적어도 하나의 음성 정보 및 비음성 정보의 조합으로 이루어지고,

상기 복수 시추에이션 케이스 각각의 적합도는 상기 생성된 시추에이션 케이스의 조합 정보와 상기 복수 시추에이션 케이스의 조합 정보 간의 일치 정도인 대화 시스템.

청구항 2

삭제

청구항 3

제 1 항에 있어서, 상기 도메인 결정부는,

상기 출력된 응답에 대응하는 사용자 응답을 입력받아 상기 사용자의 만족도를 판단하고, 상기 생성된 시추에이션 케이스에 상기 만족도를 반영하여 업데이트하여 저장하는 대화 시스템.

청구항 4

제 1 항에 있어서,

상기 복수 시추에이션 케이스는 학습을 통해 생성되어 저장된 대화 시스템.

청구항 5

제 1 항에 있어서,

상기 복수 시추에이션 케이스는, 상기 사용자의 성향, 습관, 일정, 상기 사용자의 상태 정보인 표정, 옷차림, 제스처 중 적어도 하나의 비음성 정보와 음성 정보가 조합되어 이루어지고, 상기 조합에 대응하는 도메인이 설정되어 있는 대화 시스템.

청구항 6

제 1 항에 있어서, 상기 멀티 모달 입력부는,

상기 사용자의 상태 정보인 표정, 옷차림, 제스처 정보 중 적어도 하나의 비음성 정보를 영상으로 입력받는 대화 시스템.

청구항 7

제 1 항에 있어서, 상기 도메인 결정부는,

상기 미리 저장된 복수 시추에이션 케이스의 적합도 중 적합도가 가장 큰 시추에이션 케이스를 선택하여 현재 도메인으로 결정하는 대화 시스템.

청구항 8

제 7 항에 있어서,

복수 도메인이 저장된 도메인 데이터베이스부를 더 포함하고,

상기 도메인 결정부는, 상기 선택된 시추에이션 케이스의 적합도를 이용하여 각 도메인의 적합도를 산출하고, 적합도가 가장 큰 도메인을 상기 현재 도메인으로 결정하는 대화 시스템.

청구항 9

제 1 항에 있어서,

상기 멀티 모달 입력부는, 상기 사용자가 거주하는 실내의 환경 정보를 감지하는 환경 감지부를 더 포함하고,

상기 도메인 결정부는, 상기 환경 정보를 이용하여 도메인을 결정하는 대화 시스템.

청구항 10

제 1 항에 있어서,

상기 사용자의 일정 정보, 행동 이력 정보, 실내 구조 정보, 날씨 정보를 제공하는 정보제공부를 더 포함하고,

상기 도메인 결정부는, 상기 사용자의 일정, 행동 이력, 실내 구조, 날씨 정보 중 적어도 하나의 정보를 이용하여 도메인을 결정하는 대화 시스템.

청구항 11

제 1 항에 있어서,

상기 대화관리부는, 상기 도메인 결정부로 대화 문맥, 이전 대화 정보를 전송하고,

상기 도메인 결정부는, 상기 대화 문맥 및 이전 대화 정보 중 적어도 하나를 이용하여 도메인을 결정하는 대화 시스템.

청구항 12

사용자의 음성 정보 및 비음성 정보가 입력되는 멀티 모달 입력부;

상기 입력된 음성 정보에 대응하는 후보 도메인을 선정하고, 상기 입력된 비음성 정보와 미리 저장된 복수 시추에이션 케이스를 각각 비교하여 상기 복수 시추에이션 케이스 각각의 적합도를 산출하고, 상기 적합도가 가장 큰 시추에이션 케이스와 상기 후보 도메인에 기초하여 현재 도메인을 결정하는 도메인 결정부;

상기 현재 도메인에 대응하는 응답을 선택하는 대화 관리부; 및

상기 응답을 출력하는 멀티 모달 출력부를 포함하되,

상기 미리 저장된 복수 시추에이션 케이스 각각은 적어도 하나의 비음성 정보의 조합으로 이루어지고,

상기 복수 시추에이션 케이스 각각의 적합도는 상기 입력된 비음성 정보와 상기 복수 시추에이션 케이스의 비음성 정보의 일치 정도인 대화 시스템.

청구항 13

제 12항에 있어서, 상기 복수 시추에이션 케이스는,

도메인이 설정되어 있는 대화 시스템.

청구항 14

사용자의 음성 정보 및 비음성 정보를 입력받고;

상기 입력된 음성 정보 및 비음성 정보를 조합하여 시추에이션 케이스를 생성하고;

상기 생성된 시추에이션 케이스와 미리 저장된 복수 시추에이션 케이스를 각각 비교하여 상기 복수 시추에이션 케이스 각각의 적합도를 산출하고;

상기 산출된 적합도에 기초하여 현재 도메인을 결정하고;

상기 현재 도메인에 대응하는 응답을 선택하고, 상기 선택된 응답을 출력하되,

상기 미리 저장된 복수 시추에이션 케이스 각각은 적어도 하나의 음성 정보 및 비음성 정보의 조합으로 이루어지고,

상기 복수 시추에이션 케이스 각각의 적합도는 상기 생성된 시추에이션 케이스의 조합 정보와 상기 복수 시추에이션 케이스의 조합 정보 간의 일치 정도인 대화 시스템의 대화 방법.

청구항 15

제 14 항에 있어서, 상기 미리 설정된 복수 시추에이션 케이스는,

상기 사용자의 성향, 습관, 일정, 상기 사용자의 상태 정보인 표정, 옷차림, 제스처 정보, 실내 환경 정보, 날씨 정보 중 적어도 하나의 비음성 정보와 음성 정보가 조합되어 이루어지는 대화 시스템의 대화 방법.

청구항 16

제 14 항에 있어서, 상기 입력된 비음성 정보는,

상기 사용자의 상태 정보인 표정, 옷차림, 제스처 중 적어도 하나의 정보인 대화 시스템의 대화 방법.

청구항 17

삭제

청구항 18

제 14 항에 있어서, 상기 현재 도메인을 결정하는 것은,

상기 현재 도메인이 결정되면 상기 응답 출력에 대응하는 사용자의 응답을 입력받고,

상기 사용자의 응답에 대응하는 만족도를 상기 생성된 시추에이션 케이스에 반영하여 저장하는 것을 더 포함하는 대화 시스템의 대화 방법.

청구항 19

제 14 항에 있어서, 상기 현재 도메인을 결정하는 것은,

상기 선택된 시추에이션 케이스를 미리 저장된 복수 도메인에 각각 적용하여 상기 미리 저장된 복수 도메인의 적합도를 산출하고,

상기 복수 도메인의 적합도 중 적합도가 가장 높은 도메인을 현재 도메인으로 결정하는 대화 시스템의 대화 방법.

청구항 20

제 19 항에 있어서, 상기 적합도가 가장 높은 도메인을 현재 도메인으로 결정하는 것은,

상기 적합도가 가장 높은 도메인의 적합도와 기준값을 비교하고,

상기 적합도가 가장 높은 도메인의 적합도가 기준값을 초과하면 현재 도메인 결정으로 판단하고,

상기 적합도가 가장 높은 도메인의 적합도가 기준값 이하이면 현재 도메인 결정 실패로 판단하여 대화를 재요청하는 대화 시스템의 대화 방법.

청구항 21

제 14 항에 있어서, 상기 현재 도메인을 결정하는 것은,

상기 사용자가 거주하는 실내의 환경 정보를 입력받고,

상기 환경 정보를 이용하여 도메인을 결정하는 것을 더 포함하는 대화 시스템의 대화 방법.

청구항 22

제 14 항에 있어서, 상기 현재 도메인을 결정하는 것은,
 상기 사용자의 일정 정보, 행동 이력 정보, 실내 구조 정보, 날씨 정보를 제공받고,
 상기 사용자의 일정, 행동 이력, 실내 구조, 날씨 정보 중 적어도 하나의 정보를 이용하여 도메인을 결정하는 것을 더 포함하는 대화 시스템의 대화 방법.

청구항 23

제 14 항에 있어서, 상기 현재 도메인을 결정하는 것은,
 대화 문맥, 이전 대화 정보를 입력받고,
 상기 대화 문맥 및 이전 대화 정보 중 적어도 하나를 이용하여 도메인을 결정하는 것을 더 포함하는 대화 시스템의 대화 방법.

청구항 24

제 14 항에 있어서,
 상기 음성 정보의 음성인식을 수행하여 후보 도메인을 선정하고,
 상기 입력된 비음성 정보에 기초하여 미리 저장된 복수 시추에이션 케이스의 적합도를 각각 산출하고,
 상기 미리 저장된 복수 시추에이션 케이스의 적합도 중 가장 큰 적합도를 가진 시추에이션 케이스를 선택하고,
 상기 선택된 시추에이션 케이스에 대응하는 도메인과 상기 후보 도메인에 기초하여 현재 도메인을 결정하는 것을 더 포함하는 대화 시스템의 대화 방법.

발명의 설명

발명의 상세한 설명

기술 분야

[0001] 본 발명은 사용자의 다양한 자연어 질의에 대해 지능적으로 응답하거나 대화를 재요청하여 사용자와의 대화 품질을 높이기 위한 대화 시스템 및 그의 대화 방법에 관한 것이다.

배경 기술

[0002] 대화 시스템은 사용자와의 대화를 수행하고 나아가 사용자로부터 지시된 명령을 수행하는 시스템으로, 대화 시스템은 네트워크 기반의 서버 및 단말기, 그리고 로봇 등에 마련되어 있다.

[0003] 대화 시스템은 사용자와의 대화 수행이나 사용자의 명령을 입력받기 위한 인터페이스로 키보드와 마우스를 가진다. 이러한 인터페이스는 사용자가 키보드와 마우스가 위치한 곳으로 이동한 후 키보드와 마우스와의 접촉을 통해 대화 수행이나 명령 입력을 수행해야 하기 때문에 사용자의 이용에 불편함이 있었고 또한 대화 시스템이 로봇인 경우에는 로봇의 이동성으로 인해 로봇에 장착하기 어려웠다. 이에 따라 대화 시스템은 사용자와의 인터페이스로 비접촉식 인터페이스인 음성 인식 인터페이스의 이용이 부각되고 있다.

[0004] 여기서 음성 인식은 사용자 발화 시 음성의 특징을 추출하고, 추출된 특징에 패턴인식 알고리즘을 적용시킨 후 사용자가 어떤 음소열 또는 단어열을 발화시켜 발생된 음성인가를 역추적하여 사용자의 말을 알아듣고, 사용자가 원하는 정보를 말로 전해주는 사용자와의 인터페이스이다.

[0005] 이러한 대화 시스템의 음성 인식은 사용자의 발화에 대한 음성 인식 성능이 높지 않기 때문에 대화 시스템과 사용자와의 자연스러운 대화 진행이 어렵다. 이에 따라 사용자와의 자연스러운 대화 진행을 위해 다양한 방법들이 제안되고 있는데, 그 중 하나가 도메인 기반의 음성 인식 방법이다.

[0006] 여기서 도메인 기반의 음성 인식은 날씨, 관광 등과 같은 주제(topic) 별로 다수의 도메인을 만들고 각 도메인에 대하여 특화된 언어모델을 생성한 다음, 이를 바탕으로 사용자의 발화에 대하여 1차 음성인식을 수행하여 핵심어를 인식하고, 인식된 핵심어에 해당하는 도메인에 대하여 2차 음성인식을 수행하여 사용자의 발화 의도를

과약하여 사용자와의 자연스러운 대화가 진행되도록 한다.

- [0007] 이때, 1차 음성인식과정에서 오류가 생기는 경우 오류를 복구할 별도의 기회없이 잘못 인식된 핵심어로 추출되는 도메인의 언어모델을 이용하여 2차 음성인식과정이 진행됨으로써 잘못된 인식결과를 도출해 내는 등, 음성인식의 정확도가 낮아지는 문제가 있다.
- [0008] 또한, 사용자 발화에 대응한 발화문이 적어도 두 개의 도메인에 해당하는 핵심어를 포함할 경우에는 다수의 도메인 중 하나의 도메인을 식별하는 것이 어려운 문제가 있다.
- [0009] 이와 같이 언어모델(LM: Language Model)을 이용한 도메인 기반의 음성 인식은 음성 인식된 결과만을 이용하여 도메인을 결정하기 때문에 도메인 검색 스페이스가 넓고 음성인식의 오류가 있는 경우 음성 인식의 실패 가능성이 높아 사용자의 발화 의도 파악 실패 가능성이 높으며, 음성 인식이 잘 된 경우에도 음성인식 결과가 복수 도메인에서 공통적으로 사용된 경우 도메인 결정이 쉽지 않은 문제가 있다.

발명의 내용

과제 해결수단

- [0010] 일 측면에 따르면 대화 시스템은 사용자의 음성과 비음성 정보가 입력되는 멀티 모달 입력부; 적어도 하나의 음성 및 비음성 정보가 조합되어 미리 저장된 복수 시추에이션 케이스를 가지고, 멀티 모달 입력부로부터 입력된 음성 및 비음성 정보를 기초로 미리 저장된 복수 시추에이션 케이스의 적합도를 각각 산출하여 현재 도메인을 결정하는 도메인 결정부; 현재 도메인에 대응하는 응답을 선택하는 대화 관리부; 응답을 출력하는 멀티 모달 출력부를 포함한다.
- [0011] 도메인 결정부는, 멀티 모달 입력부로부터 입력된 음성 및 비음성 정보를 조합하여 시추에이션 케이스를 생성하고, 생성된 시추에이션 케이스와 미리 저장된 복수 시추에이션 케이스를 각각 비교하여 적합도를 산출한다.
- [0012] 도메인 결정부는, 출력된 응답에 대응하는 사용자 응답을 입력받아 사용자의 만족도를 판단하고, 생성된 시추에이션 케이스에 만족도를 반영하여 업데이트하여 저장한다.
- [0013] 복수 시추에이션 케이스는 학습을 통해 생성되어 저장된다.
- [0014] 복수 시추에이션 케이스는, 사용자의 성향, 습관, 일정, 사용자의 상태 정보인 표정, 옷차림, 제스처 중 적어도 하나의 비음성 정보와 음성 정보가 조합되어 이루어지고, 조합에 대응하는 도메인이 설정되어 있다.
- [0015] 멀티 모달 입력부는, 사용자의 상태 정보인 표정, 옷차림, 제스처 정보 중 적어도 하나의 비음성 정보를 영상으로 입력받는다.
- [0016] 도메인 결정부는, 미리 저장된 복수 시추에이션 케이스의 적합도 중 적합도가 가장 큰 시추에이션 케이스를 선택하여 현재 도메인으로 결정한다.
- [0017] 복수 도메인이 저장된 도메인 데이터베이스부를 더 포함하고, 도메인 결정부는, 선택된 시추에이션 케이스의 적합도를 이용하여 각 도메인의 적합도를 산출하고, 적합도가 가장 큰 도메인을 현재 도메인으로 결정한다.
- [0018] 멀티 모달 입력부는, 사용자가 거주하는 실내의 환경 정보를 감지하는 환경 감지부를 더 포함하고, 도메인 결정부는, 환경 정보를 이용하여 도메인을 결정한다.
- [0019] 사용자의 일정 정보, 행동 이력 정보, 실내 구조 정보, 날씨 정보를 제공하는 정보제공부를 더 포함하고, 도메인 결정부는, 사용자의 일정, 행동 이력, 실내 구조, 날씨 정보 중 적어도 하나의 정보를 이용하여 도메인을 결정한다.
- [0020] 대화관리부는, 도메인 결정부로 대화 문맥, 이전 대화 정보를 전송하고, 도메인 결정부는, 대화 문맥 및 이전 대화 정보 중 적어도 하나를 이용하여 도메인을 결정한다.
- [0021] 다른 측면에 따르면 대화 시스템은 사용자의 음성과 비음성 정보가 입력되는 멀티 모달 입력부; 적어도 하나의 비음성 정보가 조합되어 미리 저장된 복수 시추에이션 케이스를 가지고, 멀티 모달 입력부로부터 입력된 음성에 대응하는 후보 도메인을 선정하고, 멀티 모달 입력부로부터 비음성 정보를 기초로 미리 저장된 복수 시추에이션 케이스의 적합도를 각각 산출하고, 적합도가 가장 큰 시추에이션 케이스와 후보 도메인에 기초하여 현재 도메인을 결정하는 도메인 결정부; 현재 도메인에 대응하는 응답을 선택하는 대화 관리부; 응답을 출력하는 멀티 모달 출력부를 포함한다.

- [0022] 복수 시추에이션 케이스는, 도메인이 설정되어 있다.
- [0023] 또 다른 측면에 따르면 대화 시스템의 대화 방법은, 사용자의 음성 정보와 비음성 정보를 입력받고, 입력된 음성 및 비음성 정보에 기초하여 미리 저장된 복수 시추에이션 케이스의 적합도를 각각 산출하고, 산출된 결과에 기초하여 현재 도메인을 결정하고, 현재 도메인에 대응하는 응답을 선택하여 선택된 응답을 출력한다.
- [0024] 미리 설정된 복수 시추에이션 케이스는, 사용자의 성향, 습관, 일정, 사용자의 상태 정보인 표정, 옷차림, 제스처 정보, 실내 환경 정보, 날씨 정보 중 적어도 하나의 비음성 정보와 음성 정보가 조합되어 이루어진다.
- [0025] 입력된 비음성 정보는, 사용자의 상태 정보인 표정, 옷차림, 제스처 중 적어도 하나의 정보이다.
- [0026] 현재 도메인을 결정하는 것은, 입력된 음성 및 비음성 정보를 조합하여 시추에이션 케이스를 생성하고, 생성된 시추에이션 케이스와 미리 저장된 복수 시추에이션 케이스를 각각 비교하여 적합도를 산출하고, 적합도가 가장 큰 시추에이션 케이스를 선택하는 것을 더 포함한다.
- [0027] 현재 도메인을 결정하는 것은, 현재 도메인이 결정되면 응답 출력에 대응하는 사용자의 응답을 입력받고, 사용자의 응답에 대응하는 만족도를 생성된 시추에이션 케이스에 반영하여 저장하는 것을 더 포함한다.
- [0028] 현재 도메인을 결정하는 것은, 선택된 시추에이션 케이스를 미리 저장된 복수 도메인에 각각 적용하여 미리 저장된 복수 도메인의 적합도를 산출하고, 복수 도메인의 적합도 중 적합도가 가장 높은 도메인을 현재 도메인으로 결정한다.
- [0029] 적합도가 가장 높은 도메인을 현재 도메인으로 결정하는 하는 것은, 적합도가 가장 높은 도메인의 적합도와 기준값을 비교하고, 적합도가 가장 높은 도메인의 적합도가 기준값을 초과하면 현재 도메인 결정으로 판단하고, 적합도가 가장 높은 도메인의 적합도가 기준값 이하이면 현재 도메인 결정 실패로 판단하여 대화를 재요청한다.
- [0030] 현재 도메인을 결정하는 것은, 사용자가 거주하는 실내의 환경 정보를 입력받고, 환경 정보를 이용하여 도메인을 결정하는 것을 더 포함한다.
- [0031] 현재 도메인을 결정하는 것은, 사용자의 일정 정보, 행동 이력 정보, 실내 구조 정보, 날씨 정보를 제공받고, 사용자의 일정, 행동 이력, 실내 구조, 날씨 정보 중 적어도 하나의 정보를 이용하여 도메인을 결정하는 것을 더 포함한다.
- [0032] 현재 도메인을 결정하는 것은, 대화 문맥, 이전 대화 정보를 입력받고, 대화 문맥 및 이전 대화 정보 중 적어도 하나를 이용하여 도메인을 결정하는 것을 더 포함한다.
- [0033] 미리 저장된 복수 시추에이션 케이스는 적어도 하나의 비음성 정보의 조합으로 이루어지고, 음성 정보의 음성인식을 수행하여 후보 도메인을 선정하고, 입력된 비음성 정보에 기초하여 미리 저장된 복수 시추에이션 케이스의 적합도를 각각 산출하고, 미리 저장된 복수 시추에이션 케이스의 적합도 중 가장 큰 적합도를 가진 시추에이션 케이스를 선택하고, 선택된 시추에이션 케이스에 대응하는 도메인과 후보 도메인에 기초하여 현재 도메인을 결정하는 것을 더 포함한다.

효 과

- [0034] 일 측면에 따르면 도메인 결정에 반영된 복수 정보 조합으로 이루어진 시추에이션 케이스와 현재 정보, 음성 인식 결과를 이용하여 도메인을 결정함으로써 대화 검색 스페이스를 축소할 수 있고 도메인 결정 정확도를 높일 수 있어 적절한 응답을 수행할 수 있고, 이에 따라 사용자와의 자연스러운 대화를 수행할 수 있다.
- [0035] 다른 측면에 따르면 도메인 결정 정확도를 상승시켜 사용자의 발화 의도를 보다 정확하게 파악할 수 있어 사용자 발화 의도에 대응하는 적절한 서비스를 수행할 수 있고, 이에 따라 사용자의 서비스 만족도를 높일 수 있다.
- [0036] 또한 네트워크를 통해 실내 환경 정보, 사용자의 일정 정보, 사용자의 행동 이력 정보, 사용자의 집안 구조 정보 등을 제공받고, 이 정보들을 도메인 결정 시에 반영함으로써 현재 환경 조건에 따른 사용자의 발화 의도에 적절한 응답을 수행할 수 있다.

발명의 실시를 위한 구체적인 내용

- [0037] 이하에서는 첨부도면을 참조하여 본 발명에 대해 상세히 설명한다.
- [0038] 도 1은 일 실시예에 따른 대화 시스템의 구성도로서, 도메인 결정 정확도를 높여 사용자의 대화 의도에 적절한

응답을 하기 위한 대화 시스템은 멀티 모달 입력부(10), 멀티 모달 입력 처리부(20), 대화 관리부(30), 도메인 결정부(40), 데이터베이스부(50), 정보제공부(60), 멀티 모달 출력 처리부(70), 멀티 모달 출력부(80)를 포함한다.

- [0039] 멀티 모달 입력부(10)는 사용자 발화에 따른 음성 정보를 입력받는 제1입력부(11)와, 사용자로부터 비음성 정보를 입력받는 제2입력부(12)를 가지고, 입력된 정보를 멀티 모달 입력 처리부(20)로 전송한다.
- [0040] 제1입력부(11)는 사용자 음성의 음파를 받아서 그 진동에 따른 전기 신호를 발생하는 마이크로 폰과, 음성 신호를 텍스트로 변환하는 음성인식부(ASR: Automated Speech Recognizer)를 가진다.
- [0041] 음성인식부는 사용자가 발성한 음성으로부터 특징 벡터를 추출하고 추출된 특징 벡터와 음성 데이터베이스에 저장된 각 단어의 언어 모델 또는 음소 단위의 음향 모델을 비교하여 후보 단어 또는 후보 음소를 추출하고, 후보 단어 또는 후보 음소를 기초로 문장을 획득하고, 획득된 문장의 문법 구조, 문장 문맥을 판단하고, 특정 주제에 의 부합 여부를 판단하여 가장 적합한 단어나 음소 등의 핵심어를 판단하고 이에 대응하는 문장을 판단한다.
- [0042] 제2입력부(12)는 비음성인 사용자의 영상을 수집하는 카메라와 같은 영상수집부를 가지고, 영상수집부를 통해 수집된 영상을 멀티 모달 입력 처리부(20)로 전송한다.
- [0043] 아울러 멀티 모달 입력부(10)는 집안 내의 온도, 습도, 조도 등의 환경 정보를 수집하는 환경 감지부를 더 포함하는 것이 가능하고, 사용자로부터 문자를 입력받는 키보드를 더 포함하는 것도 가능하다.
- [0044] 멀티 모달 입력 처리부(20)는 멀티 모달 입력부(10)로부터 전송된 단어, 음소 등의 핵심어, 문장의 정보 중 적어도 하나를 이용하여 대화 관리를 위해 필요한 의미 정보를 분석하고, 이때 음성인식에 불필요한 요소를 제거하여 가공한다.
- [0045] 멀티 모달 입력 처리부(20)는 단어, 음소 등의 핵심어, 문장의 정보 중 적어도 하나의 의미 정보를 이용하여 언어 처리를 수행하고 언어 처리된 데이터를 대화 관리부(30) 및 도메인 결정부(40)로 전송한다.
- [0046] 멀티 모달 입력 처리부(20)는 멀티 모달 입력부(10)로부터 전송된 문장의 정보를 이용하여 대화 관리를 위해 필요한 의미 정보를 분석하는 언어이해부(SLU: Spoken Language Understanding)를 포함한다.
- [0047] 멀티 모달 입력 처리부(20)는 멀티 모달 입력부(10)로부터 전송된 영상의 필요한 요소를 제거하여 가공하고 영상 처리를 수행하여 사용자의 제스처, 옷차림, 바쁨 여부, 표정 등의 사용자 상태 정보, 사용자 주변의 물체 등을 인식하고, 인식된 영상을 대화 관리부(30) 및 도메인 결정부(40)로 전송한다.
- [0048] 대화 관리부(30)는 멀티 모달 입력 처리부(20)로부터 전송된 단어, 음소 등의 핵심어, 문장 중 적어도 하나의 의미 정보와 이전 대화에 기초하여 대화 문맥을 판단한다. 이러한 대화 관리부(30)는 이전 대화를 저장하고 있다.
- [0049] 대화 관리부(30)는 도메인 결정부(40)로 도메인 결정을 지시하고, 도메인 결정 지시 시 대화 문맥의 정보를 도메인 결정부(40)로 전송한다.
- [0050] 대화 관리부(30)는 도메인 결정부(40)에서 결정된 도메인에 대응하는 응답을 생성한다. 여기서 대화 관리부(30)는 각 도메인에 대응하는 적절한 응답 메시지를 저장 및 관리하는 응답 데이터베이스(미도시)를 가진다.
- [0051] 도메인 결정부(DR:Domain Reasoner, 40)는 멀티 모달 입력 처리부(20)로부터 전송된 사용자의 상태 정보, 실내 환경 정보, 사용자의 일정, 날씨 등의 현재 비음성 정보 중 적어도 하나의 비음성 정보와 음성인식에 대응하는 핵심어 정보를 조합하여 시추에이션 케이스를 생성하고, 생성된 시추에이션 케이스와 미리 저장된 복수 시추에이션 케이스의 적합도를 각각 산출하여 사용자의 발화 의도에 대응하는 도메인을 결정한다.
- [0052] 좀 더 구체적으로 도 2에 도시된 바와 같이 도메인 결정부(40)는 정보 입력부(41), 결정부(42), 시추에이션 케이스부(43)를 포함한다.
- [0053] 정보 입력부(41)는 멀티 모달 입력 처리부(20)로부터 사용자의 상태 정보인 제스처, 표정, 옷차림 등의 정보가 입력되고, 정보 제공부(60)로부터 날씨 정보, 일정 정보 등이 입력되며, 대화 관리부(30)로부터 대화 문맥 정보, 이전 대화 정보가 입력되고, 입력된 정보를 결정부(42)로 전송한다.
- [0054] 결정부(42)는 멀티 모달 입력부(10), 대화 관리부(30), 정보 제공부(60)로부터 전송된 적어도 하나의 현재 정보(음성 정보 및 비음성 정보)에 대응하는 코드를 데이터베이스부(50)에서 검색하고, 검색된 코드를 조합하여 시추에이션 케이스를 생성한다.

- [0055] 그리고 결정부(42)는 생성된 시추에이션 케이스의 정보에 기초하여 시추에이션 케이스부(43)에 미리 저장된 복수 시추에이션 케이스의 적합도를 판단하고, 적합도가 가장 큰 시추에이션 케이스를 선택하여 현재 도메인으로 결정한다.
- [0056] 여기서 현재 정보는 사용자의 상태인 제스처, 표정, 옷차림, 바빔 여부 정보, 실내 환경 정보인 온도, 습도, 조도, 사용자의 일정 정보, 사용자의 행동 이력 정보, 날씨 정보, 대화 문맥 정보, 이전 대화 정보 중 적어도 두 개의 정보이고, 시추에이션 케이스의 적합도는 복수의 현재 정보를 많이 포함할수록 높아진다.
- [0057] 결정부(42)는 미리 저장된 복수 도메인에 적합도가 가장 큰 시추에이션 케이스의 적합도를 각각 반영하여 복수 도메인의 적합도를 각각 산출하고, 적합도가 가장 큰 도메인을 선택하여 최종 도메인으로 결정한다.
- [0058] 여기서 복수 시추에이션 케이스는, 사용자와 이전 대화 시 학습에 의해 생성되어 저장된 것으로, 복수 시추에이션 케이스는 당시 대화 시 사용자 응답에 따른 사용자의 만족도, 사용 빈도, 유사도가 반영되어 있다.
- [0059] 이에 따라 복수 시추에이션 케이스의 적합도 산출 시, 사용자 응답에 따른 사용자의 만족도, 사용 빈도, 유사도가 반영되어 산출되는데, 이때 사용자의 만족도 또는 사용 빈도 중 적어도 하나가 높으면 적합도는 증가하고, 사용자의 만족도 또는 사용 빈도 중 적어도 하나가 낮으면 적합도는 감소한다.
- [0060] 도메인의 적합도 산출은 베이시안 네트워크(Baysian Network)를 이용하거나, 규칙 기반(Rule-based) 추론 방식을 이용한다.
- [0061] 그리고 결정부(42)는 최종 도메인의 적합도와 기준값을 비교하여 최종 도메인의 적합도가 기준값을 초과하면 도메인 결정에 성공했다고 판단하여 결정된 도메인을 대화 관리부(30)로 전송하고, 최종 도메인의 적합도가 기준값 이하이면 도메인 결정에 실패했다고 판단하여 사용자에게 재질문을 요청한다.
- [0062] 시추에이션 케이스부(43)는 도메인 결정에 사용된 정보, 즉 사용자의 성향, 습관, 사용자의 상태인 제스처, 표정, 옷차림, 사용자의 일정, 환경 정보, 날씨 정보 중 적어도 하나의 비음성 정보와 음성 정보의 조합으로 이루어진다. 도 3에 도시된 바와 같이, 시추에이션 케이스부(43)는 제1시추에이션 케이스(SC1), 제2시추에이션 케이스(SC2) ... 제9시추에이션 케이스(SC9), 제10시추에이션 케이스(SC10) 등 복수 개의 시추에이션 케이스를 가진다. 여기서 복수 시추에이션 케이스는 사용자와 이전 대화 시 학습에 의해 생성되어 저장된 것이다.
- [0063] 여기서 사용자의 성향이나 습관은 사용자 상태 정보에 대응하는 응답에 기초하여 판단된 것이다. 예를 들어 운동을 하면서 심부름을 잘시키는지, 텔레비전 시청 중 심부름을 잘시키는지, 텔레비전 시청 중 일정관리를 하는지 등에 대한 성향을 학습에 의해 습득한 것이다.
- [0064] 그리고 시추에이션 케이스부(43)는 도메인이 결정되면 이 도메인 결정에 반영된 복수 정보를 가지는 시추에이션 케이스를 저장하고, 복수 정보를 모두 가지는 시추에이션 케이스가 있는 경우에는 해당 시추에이션 케이스의 적합도를 업데이트한다.
- [0065] 그리고, 각 시추에이션 케이스의 적합도와 미리 정해진 제거 기준이 되는 적합도를 비교하여 제거 기준이 되는 적합도 미만의 적합도를 가지는 시추에이션 케이스는 제거한다.
- [0066] 삭제
- [0067] 이와 같이 도메인 결정부(40)는 도메인 결정 시 이전에 학습을 통해 생성되어 미리 저장된 복수 시추에이션 케이스와 현재 정보를 이용함으로써 대화 검색 스페이스를 줄일 수 있고, 도메인 결정 정확도를 높일 수 있다.
- [0068] 데이터베이스부(50)는 도 4에 도시된 바와 같이, 사용자의 성향 정보에 대응하는 코드(A1, A2, A3 ...)가 저장된 성향 데이터베이스(51), 사용자의 습관 정보에 대응하는 코드(B1, B2, B3 ...)가 저장된 저장된 습관 데이터베이스(52), 사용자의 옷차림, 표정, 제스처 정보에 대응하는 코드(C1, C2, C3 ...)가 저장된 상태 데이터베이스(53), 사용자의 일정 정보에 대응한 코드(D1, D2, D3 ...)가 저장된 일정 데이터베이스(54), 실내의 온도 및 습도 정보 등의 환경 정보에 대응하는 코드(E1, E2, E3 ...)가 저장된 환경 데이터베이스(55), 날씨 정보에 대응하는 코드(F1, F2, F3 ...)가 저장된 날씨 데이터베이스(56), 사용자 반응에 대응하는 코드(G1, G2, G3 ...)가 저장된 반응 데이터베이스(57)를 포함하고, 도메인 결정부(40)의 요청이 입력되면 요청된 정보에 대응하는 코드를 검색하여 도메인 결정부(40)로 전송한다.
- [0069] 정보제공부(60)는 지역 정보 통신망(Local Area Network)과 원격 네트워크(Remote Network)를 가지고, 이 네트

워크를 통해 각종 정보를 수집하여 대화 관리부(30) 및 도메인 결정부(40)로 전송한다. 즉, 정보제공부(60)는 지역 정보 통신망(Local Area Network)을 통해 연결되어 있는 단말기(미도시)로부터 사용자의 행동 이력 정보, 실내 구조 정보, 사용자 정보 등을 제공받아 대화 관리부(30) 및 도메인 결정부(40)로 전송하고, 원격 네트워크(Remote Network)를 통해 웹사이트에 접속하여 사용자의 일정 정보, 날씨 정보 등을 제공받아 대화 관리부(30) 및 도메인 결정부(40)로 전송한다. 여기서 사용자 일정 정보는 사용자가 이전에 해당 웹사이트에 접속하여 해당 웹사이트에 기록해 둔 정보이다.

- [0070] 멀티 모달 출력 처리부(70)는 대화 관리부(30)로부터 전송된 응답에 대응하는 텍스트의 문장 데이터 또는 행동 지시 데이터를 생성하여 멀티 모달 출력부(80)로 전송한다.
- [0071] 멀티 모달 출력부(80)는 멀티 모달 출력 처리부(70)로부터 전송된 텍스트의 문장에 대응하는 음성을 출력하거나, 행동을 수행한다. 즉 멀티 모달 출력부(80)는 응답에 대응하는 텍스트를 음성으로 변환시키는 변환부(TTS: Text To Speech)와, 변환된 음성을 적정 레벨로 증폭하여 음성으로 출력하는 스피커를 가지는 제1출력부(81)와, 응답에 대응하는 행동을 수행하는 제2출력부(82)를 포함한다.
- [0072] 제2출력부(82)는 응답에 대응하는 문자나 이모티콘을 디스플레이를 통해 표시한다.
- [0073] 여기서 제2출력부(82)는 응답을 문자로 표시하는 디스플레이 기기를 이용하는 것이 가능하고, 이때 디스플레이 기기는 사용자의 대화, 명령 입력 및 응답 표시가 모두 가능한 하는 터치 스크린(Touch screen) 이용이 가능하다.
- [0074] 대화시스템이 로봇에 마련된 경우에는 제2출력부(82)는 응답에 대응하는 행동을 수행하도록 로봇의 각 관절을 구동시킨다.
- [0075] 도 5는 일 실시예에 따른 대화 시스템의 대화 방법의 순서도로, 도 1 내지 도 4를 참조하여 설명하도록 한다.
- [0076] 사용자 발화에 따른 음성 정보 및 비음성 정보인 사용자의 영상 정보, 실내 환경 정보 등의 멀티 모달 정보를 입력(101) 받는다.
- [0077] 사용자가 발성한 음성으로부터 특징 벡터를 추출하고 추출된 특징 벡터와 음성 데이터베이스에 저장된 각 단어의 언어 모델 또는 음소 단위의 음향 모델을 비교하여 후보 단어 또는 후보 음소를 추출하고, 후보 단어 또는 후보 음소를 기초로 문장을 획득하고, 획득된 문장의 문법 구조, 문장 문맥을 판단하여 가장 적합한 단어나 음소를 판단하고 이에 대응하는 문장을 판단함으로써 음성 인식을 수행(102)한다.
- [0078] 그리고 단어, 음소, 문장의 정보 중 적어도 하나를 이용하여 대화 관리를 위해 필요한 의미 정보를 분석하는 음성 이해를 수행(103)하고 언어 처리를 수행하며, 문장, 단어, 음소의 의미 정보와 이전 대화에 기초하여 대화 문맥을 판단한다.
- [0079] 다음 음성인식 결과에 대응하는 상위 복수 개의 문장, 단어, 음소를 입력으로 하여 신뢰도를 산출하고, 기준 이상의 신뢰도를 갖는 단어들 중에서 핵심어를 인식(104)한다.
- [0080] 다음 영상 처리를 수행하여 사용자 상태 정보인 제스처, 옷차림, 표정 정보를 인식하고, 사용자 주변의 물체 등을 인식한다. 이때 사용자 상태 정보에 기초하여 사용자의 성향이나 습관 등도 판단하는 것도 가능하다.
- [0081] 그리고 지역 정보 통신망(Local Area Network)을 통해 연결되어 있는 단말기(미도시)로부터 사용자의 행동 이력 정보, 실내 구조 정보, 사용자 개인 정보 등을 제공받고, 원격 네트워크(Remote Network)를 통해 웹사이트에 접속하여 사용자의 일정 정보, 날씨 정보 등을 제공받는다. 여기서 사용자 일정 정보는 사용자가 이전에 해당 웹사이트에 접속하여 해당 웹사이트에 기록해 둔 정보이다.
- [0082] 다음 적어도 하나의 현재 정보가 입력되면, 즉 사용자의 상태 정보인 제스처, 표정, 옷차림 정보, 사용자의 성향, 습관 정보, 실내 환경 정보, 사용자의 행동 이력 정보, 날씨 정보 중 적어도 하나의 비음성 정보가 입력되면, 입력된 적어도 하나의 비음성 정보와 음성 정보인 핵심어를 조합하여 시추에이션 케이스를 생성(105)한다. 이때 현재 정보, 대화 문맥 정보, 이전 대화 정보를 더 반영하는 것도 가능하다.
- [0083] 그리고 생성된 시추에이션 케이스와 미리 저장된 복수 시추에이션 케이스의 적합도를 산출하여 현재 도메인을 결정(106)하고, 현재 도메인의 성공 여부를 판단하며, 이때 현재 도메인의 결정이 성공이면 현재 도메인에 대응하는 응답을 생성하고 생성된 응답에 대응하는 텍스트의 문장 데이터를 음성으로 변환하고, 변환된 음성을 적정 레벨로 증폭하여 음성으로 출력(108)하고, 또한 생성된 응답에 대응하는 행동을 수행한 후 사용자 응답을 입력(109)받는다. 아울러 대화시스템이 로봇에 마련된 경우에는 응답에 대응하는 행동을 수행하도록 로봇의 각 관절

을 구동시킨다.

- [0084] 반면, 현재 도메인의 결정이 실패이면 사용자에게 대화를 다시 요청(110)한다.
- [0085] 도 6을 참조하여 도메인 결정(106)을 좀 더 구체적으로 설명하면, 생성된 시추에이션 케이스와 미리 저장된 복수 시추에이션 케이스의 적합도를 각각 산출(106a)한다.
- [0086] 여기서 복수 시추에이션 케이스 적합도 산출은, 생성된 시추에이션 케이스의 조합 정보와 미리 저장된 복수 시추에이션 케이스의 조합 정보를 비교하여 정보의 일치 정도를 산출하는 것이다.
- [0087] 아울러 복수 시추에이션 케이스 적합도 산출은, 생성된 시추에이션 케이스의 조합 정보와 미리 저장된 복수 시추에이션 케이스의 조합 정보를 비교하여 정보의 유사 정도를 반영하는 것도 가능하다.
- [0088] 이때 복수 시추에이션 케이스 적합도 산출은 대화 당시 사용자 응답에 따른 사용자의 만족도, 사용 빈도, 유사도를 반영한다.
- [0089] 즉, 시추에이션 케이스의 적합도는 사용자의 만족도 또는 사용 빈도 중 적어도 하나가 높으면 증가하고, 사용자의 만족도 또는 사용 빈도 중 적어도 하나가 낮으면 감소한다.
- [0090] 다음 적합도가 가장 큰 시추에이션 케이스를 선택(106b)하고, 적합도가 가장 큰 시추에이션 케이스의 적합도를 반영하여 복수 도메인의 적합도를 각각 산출(106c)하고, 적합도가 가장 큰 도메인을 선택(106d)하여 최종 도메인으로 결정한다.
- [0091] 여기서 도메인의 적합도 산출은 베이시안 네트워크(Baysian Network)를 이용하거나, 규칙 기반(Rule-based) 추론 방식을 이용한다.
- [0092] 도 7에 도시된 바와 같이, 베이시안 네트워크(Baysian Network)를 참조하여 설명한다. 도 7과 같이 사용자가 의도한 도메인이 심부름1 도메인일 경우의 확률을 알기 위해 운동, 외출, 시추에이션 케이스 적합도, 음성인식에 따른 후보 도메인을 증거 값으로 사용한다. 이때 운동 여부를 알기 위한 증거 값으로는 사용자의 제스처가 역동적 제스처인지 여부, 사용자의 옷차림 중 외출복장인지의 여부, 일정 여부를 사용하였고, 외출 여부를 알기 위한 증거 값으로 사용자의 옷차림 중 외출복장인지의 여부, 일정 여부를 사용하였다.
- [0093] 이와 같이 일정관리 도메인, 집안관리 도메인 별로 유사한 베이시안 네트워크를 구성하여 각 모델 중에서 가장 높은 확률 값을 갖는 도메인을 현재의 도메인으로 결정하게 된다.
- [0094] 다음 도메인 결정 성공 여부를 판단(107)한다. 즉 최종 도메인의 적합도와 기준값을 비교(107a)하여 최종 도메인의 적합도가 기준값을 초과하면 도메인 결정에 성공했다고 판단(107b)하고 대화 응답(108)을 수행한다.
- [0095] 이후 사용자의 응답을 입력(109)받아 응답에 대응하는 사용자 만족도를 해당 시추에이션 케이스에 반영한다.
- [0096] 즉, 도메인 결정이 성공했다고 하더라도 잘못된 결정일 수 있기 때문에 사용자의 응답을 입력받아 도메인 결정 성공 여부를 재확인하는 것이다. 이때, 사용자의 응답을 통해 최종 도메인 결정의 성공 여부를 결정하게 된다.
- [0097] 그리고 현재 도메인 결정 시 반영된 복수 정보를 모두 가지는 시추에이션 케이스를 저장한다. 이때 사용자의 응답에 따른 만족도도 반영하여 저장한다.
- [0098] 반면 최종 도메인의 적합도가 기준값 이하이면 도메인 결정에 실패(107c)했다고 판단하여 사용자에게 대화를 재요청(110)한다.
- [0099] 도메인 결정 실패 시, 사용자에게 대화를 재요청하고 사용자에게 의해 대화 재요청에 따른 음성이 인식되면 음성 인식된 음성 정보 및 현재의 비음성 정보를 이용하여 도메인 결정을 재수행한다. 이때 도메인 결정이 성공되면 대화응답을 생성 및 출력한 후 사용자의 응답을 기다린다.
- [0100] 그리고, 복수 시추에이션 케이스의 적합도와 미리 정해진 제거 적합도를 비교하여 제거 적합도 미만의 적합도를 가지는 시추에이션 케이스는 제거한다. 여기서 복수 시추에이션 케이스는 가장 일어날 법한 케이스를 학습한 결과이다.
- [0101] 이와 같이 도메인 결정에 반영된 복수 정보 조합으로 이루어진 시추에이션 케이스와 현재 정보, 음성 인식 결과를 이용하여 도메인을 결정함으로써 대화 검색 스페이스를 축소할 수 있고 도메인 결정 정확도를 높일 수 있어 적절한 응답을 수행할 수 있고, 이에 따라 사용자와의 자연스러운 대화를 수행할 수 있다.
- [0102] 또한 도메인 결정 정확도를 상승시켜 사용자의 발화 의도를 보다 정확하게 파악할 수 있어 사용자 발화 의도에

대응하는 적절한 서비스를 수행할 수 있고, 이에 따라 사용자의 서비스 만족도를 높일 수 있다.

- [0103] 또한 네트워크를 통해 실내 환경 정보, 사용자의 일정 정보, 사용자의 행동 이력 정보, 사용자의 집안 구조 정보 등을 제공받고, 이 정보들을 도메인 결정 시에 반영함으로써 현재 환경 조건에 따른 사용자의 발화 의도에 적절한 응답을 수행할 수 있다.
- [0104] 도 8은 다른 실시예에 따른 대화 시스템의 도메인 결정부의 상세 구성도이다.
- [0105] 도메인 결정부(DR:Domain Reasoner, 40)는 음성인식에 대응하는 도메인을 결정하여 도메인 후보를 선정하고, 사용자의 상태, 실내의 온도, 습도, 조도 등의 환경, 사용자의 일정, 날씨 등의 현재 정보, 음성 인식 결과, 시추에이션 케이스의 적합도를 이용하여 사용자의 발화 의도에 대응하는 도메인을 결정한다.
- [0106] 좀 더 구체적으로 도 8에 도시된 바와 같이 도메인 결정부(40)는 정보 입력부(41), 결정부(42), 시추에이션 케이스부(43), 도메인 후보선정부(44), 도메인 데이터베이스부(45)를 포함한다.
- [0107] 정보 입력부(41)는 멀티 모달 입력 처리부(20)로부터 사용자의 상태 정보인 제스처, 표정, 옷차림 등의 정보가 입력되고, 정보 제공부(60)로부터 날씨 정보, 일정 정보 등이 입력되며, 대화 관리부(30)로부터 대화 문맥 정보, 이전 대화 정보가 입력되고, 입력된 정보를 결정부(42)로 전송한다.
- [0108] 결정부(42)는 멀티 모달 입력부(10), 대화 관리부(30), 정보 제공부(60)로부터 전송된 복수의 현재 정보에 대응하는 코드를 데이터베이스부(50)에서 검색하고, 검색된 코드를 가지는 시추에이션 케이스를 검색하고, 이때 검색된 시추에이션 케이스의 도메인과 후보 도메인에 기초하여 도메인을 결정한다.
- [0109] 여기서 현재 정보는 사용자의 상태인 제스처, 표정, 옷차림 정보, 실내 환경 정보, 사용자의 일정 정보, 사용자의 행동 이력 정보, 날씨 정보, 대화 문맥 정보, 이전 대화 정보 중 적어도 하나의 비음성 정보이다.
- [0110] 도메인 결정을 좀 더 구체적으로 설명하면, 결정부(42)는 입력된 비음성 정보에 기초하여 복수 시추에이션 케이스의 적합도를 산출하고, 적합도가 가장 큰 시추에이션 케이스를 선택하고, 이 시추에이션 케이스의 적합도를 반영하여 각 후보 도메인의 적합도를 산출하고, 적합도가 가장 큰 도메인을 선택하여 최종 도메인으로 결정한다.
- [0111] 여기서 시추에이션 케이스의 적합도 산출은, 현재 입력된 비음성 정보와 복수 시추에이션 케이스의 비음성 정보의 일치 정도이다. 이 시추에이션 케이스의 적합도 산출은 사용자 응답에 따른 사용자의 만족도, 사용 빈도, 유사도를 반영한다.
- [0112] 그리고 도메인의 적합도 산출은 베이시안 네트워크(Baysian Network)를 이용하거나, 규칙 기반(Rule-based) 추론 방식을 이용한다.
- [0113] 그리고 결정부(42)는 최종 도메인의 적합도와 기준값을 비교하여 최종 도메인의 적합도가 기준값을 초과하면 도메인 결정에 성공했다고 판단하여 결정된 도메인을 대화 관리부(30)로 전송하고, 최종 도메인의 적합도가 기준값 이하이면 도메인 결정에 실패했다고 판단하여 사용자에게 재질문을 요청한다.
- [0114] 시추에이션 케이스부(43)는 도메인 결정에 사용된 정보, 즉 사용자의 성향, 습관, 사용자의 상태인 제스처, 표정, 옷차림, 사용자의 일정, 환경 정보, 날씨 정보 중 적어도 하나의 비음성 정보의 조합으로 이루어진다.
- [0115] 여기서 사용자의 성향이나 습관은 사용자 상태 정보에 대응하는 응답에 기초하여 판단된 것이다. 예를 들어 운동을 하면서 심부름을 잘시키는지, 텔레비전 시청 중 심부름을 잘시키는지, 텔레비전 시청 중 일정관리를 하는지 등에 대한 성향을 학습에 의해 습득한 것이다.
- [0116] 그리고 시추에이션 케이스부(43)는 현재 도메인 결정에 반영된 정보를 가지는 시추에이션 케이스를 저장하거나, 이 정보를 모두 가지는 시추에이션 케이스가 있는 경우에는 해당 시추에이션 케이스의 적합도를 업데이트한다.
- [0117] 그리고, 각 시추에이션 케이스의 적합도와 미리 정해진 제거 적합도를 비교하여 제거 적합도 미만의 적합도를 가지는 시추에이션 케이스는 제거한다.
- [0118] 도메인 후보 선정부(44)는 멀티 모달 입력부(10)에서 인식된 음성인식결과 중 상위 복수 개의 문장과 단어 및 음소를 입력으로 하여 신뢰도를 산출하고, 기준 이상의 신뢰도를 갖는 단어들 중에서 핵심어를 선택하고, 선택된 핵심어를 근거로 후보 도메인들을 추출하고 추출된 후보 도메인 중 기준 스코어 이상의 스코어를 갖는 도메인을 후보 도메인으로 선정한다. 이 도메인 후보 선정부(44)는 도메인 데이터베이스부(45)에 저장된 언어 모델에 기초하여 일정 개수의 후보 도메인을 선정한다.

- [0119] 도메인 데이터베이스부(45)는 복수 토픽에 각각 대응하는 도메인이 저장되어 있다. 이 도메인 데이터베이스부(45)는 도메인 별로 학습하여 미리 구해진 도메인별 언어모델이 저장되어 있다.
- [0120] 도 9는 다른 실시예에 따른 대화 시스템의 대화 방법의 순서도로, 도 8을 참조하여 설명하도록 한다.
- [0121] 사용자 발화에 따른 음성 정보 및 비음성 정보인 사용자의 영상 정보, 실내 환경 정보 등의 멀티 모달 정보를 입력(201) 받는다.
- [0122] 사용자가 발성한 음성으로부터 특징 벡터를 추출하고 추출된 특징 벡터와 음성 데이터베이스에 저장된 각 단어의 언어 모델 또는 음소 단위의 음향 모델을 비교하여 후보 단어 또는 후보 음소를 추출하고, 후보 단어 또는 후보 음소를 기초로 문장을 획득하고, 획득된 문장의 문법 구조, 문장 문맥을 판단하여 가장 적합한 단어나 음소를 판단하고 이에 대응하는 문장을 판단함으로써 음성 인식을 수행(202)한다.
- [0123] 그리고 단어, 음소, 문장의 정보 중 적어도 하나를 이용하여 대화 관리를 위해 필요한 의미 정보를 분석하는 음성 이해를 수행(203)하고 언어 처리를 수행하며, 문장, 단어, 음소 중 적어도 하나의 의미 정보와 이전 대화에 기초하여 대화 문맥을 판단한다.
- [0124] 다음 음성인식 결과에 대응하는 상위 복수 개의 문장, 단어, 음소를 입력으로 하여 신뢰도를 산출하고, 기준 이상의 신뢰도를 갖는 단어들 중에서 핵심어를 선택하고, 선택된 핵심어를 근거로 적어도 하나의 후보 도메인을 추출하고 추출된 후보 도메인 중 기준 스코어 이상의 스코어를 갖는 도메인을 후보 도메인으로 선정한다. 이 후보 도메인 선정은, 도메인 데이터베이스부(45)에 저장된 도메인별 언어 모델에 기초하여 적어도 하나의 후보 도메인을 선정(204)한다.
- [0125] 다음 영상 처리를 수행하여 사용자 상태 정보인 제스처, 옷차림, 표정 정보를 인식하고, 사용자 주변의 물체 등을 인식한다. 이때 사용자 상태 정보에 기초하여 사용자의 성향이나 습관 등도 판단하는 것도 가능하다.
- [0126] 그리고 지역 정보 통신망(Local Area Network)을 통해 연결되어 있는 단말기(미도시)로부터 사용자의 행동 이력 정보, 실내 구조 정보, 사용자 개인 정보 등을 제공받고, 원격 네트워크(Remote Network)를 통해 웹사이트에 접속하여 사용자의 일정 정보, 날씨 정보 등을 제공받는다. 여기서 사용자 일정 정보는 사용자가 이전에 해당 웹사이트에 접속하여 해당 웹사이트에 기록해 둔 정보이다.
- [0127] 그리고 입력된 비음성 정보에 기초하여 복수 시추에이션 케이스의 적합도를 각각 산출하고, 적합도가 가장 높은 시추에이션 케이스를 선택(205)한다. 여기서 각 시추에이션 케이스는 조합된 복수 정보에 대응하는 도메인이 설정되어 있다. 이때 현재 정보, 대화 문맥 정보, 이전 대화 정보를 더 반영하는 것도 가능하다.
- [0128] 다음 선택된 시추에이션 케이스의 적합도를 반영하여 각 후보 도메인의 적합도를 산출하고, 적합도가 가장 큰 도메인을 선택하여 최종 도메인으로 결정(206)한다. 여기서 도메인의 적합도 산출은 베이시안 네트워크(Baysian Network)를 이용하거나, 규칙 기반(Rule-based) 추론 방식을 이용한다.
- [0129] 다음 도메인 결정 성공 여부를 판단(207)한다. 즉 최종 도메인의 적합도와 기준값을 비교하여 최종 도메인의 적합도가 기준값을 초과하면 도메인 결정에 성공했다고 판단하여 대화 응답(208)을 수행하고, 이후 사용자의 응답을 입력(209)받아 응답에 대응하는 사용자 만족도를 해당 시추에이션 케이스에 반영한다.
- [0130] 대화 응답 수행은, 결정된 도메인에 대응하는 응답을 생성하고 생성된 응답에 대응하는 텍스트의 문장 데이터 또는 행동 지시 데이터를 생성한다. 이때 응답에 대응하는 텍스트를 음성으로 변환하고, 변환된 음성을 적정 레벨로 증폭하여 음성으로 출력하고, 또한 응답에 대응하는 행동을 수행한다. 아울러 대화시스템이 로봇에 마련된 경우에는 응답에 대응하는 행동을 수행하도록 로봇의 각 관절을 구동시킨다.
- [0131] 그리고 도메인 결정이 성공했다고 하더라도 잘못된 결정일 수 있기 때문에 사용자의 응답을 입력받아 도메인 결정을 재확인한다. 이때 사용자의 응답은 현재 도메인 결정 시 정보들의 조합으로 생성된 시추에이션 케이스에 반영되어 저장된다.
- [0132] 반면 최종 도메인의 적합도가 기준값 이하이면 도메인 결정에 실패했다고 판단하여 사용자에게 대화를 재요청(210)한다.
- [0133] 도메인 결정 실패 시, 사용자에게 대화를 재요청하고 사용자에게 의해 대화 재요청에 따른 음성이 인식되면 이 음성인식 및 현재의 비음성 정보를 이용하여 도메인 결정을 재수행한다. 이때 도메인 결정이 성공되면 대화응답을 생성 및 출력한 후 사용자의 응답을 기다린다.

- [0134] 이와 같이 도메인 결정에 반영된 복수 정보 조합으로 이루어진 시추에이션 케이스와 현재 정보, 음성 인식 결과를 이용하여 도메인을 결정함으로써 대화 검색 공간을 축소할 수 있고 도메인 결정 정확도를 높일 수 있어 적절한 응답을 수행할 수 있고, 이에 따라 사용자와의 자연스러운 대화를 수행할 수 있다.
- [0135] 또한 도메인 결정 정확도를 상승시켜 사용자의 발화 의도를 보다 정확하게 파악할 수 있어 사용자 발화 의도에 대응하는 적절한 서비스를 수행할 수 있고, 이에 따라 사용자의 서비스 만족도를 높일 수 있다.
- [0136] 또한 네트워크를 통해 실내 환경 정보, 사용자의 일정 정보, 사용자의 행동이력 정보, 사용자의 집안 구조 정도 등을 제공받고, 이 정보들을 도메인 결정 시에 반영함으로써 현재 환경 조건에 따른 사용자의 발화 의도에 적절한 응답을 수행할 수 있다.

도면의 간단한 설명

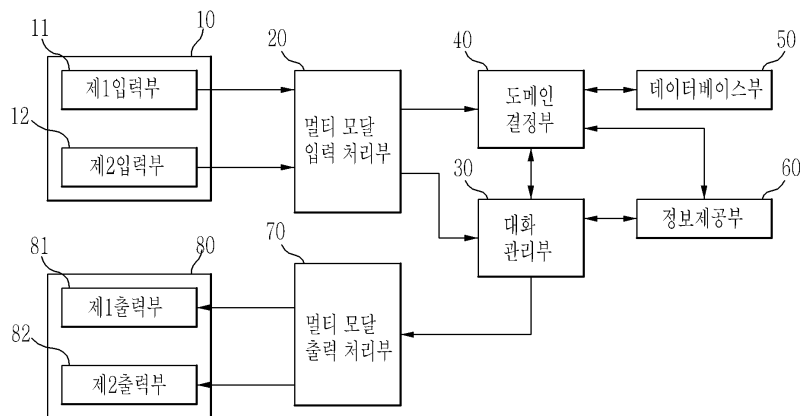
- [0137] 도 1은 일 실시예에 따른 대화 시스템의 구성도이다.
- [0138] 도 2는 일 실시예에 따른 대화 시스템에 마련된 도메인 결정부의 상세 구성도이다.
- [0139] 도 3은 일 실시예에 따른 대화 시스템에 마련된 도메인 결정부의 시추에이션 케이스부의 상세 구성도이다.
- [0140] 도 4는 일 실시예에 따른 대화 시스템에 마련된 데이터베이스부의 상세 구성도이다.
- [0141] 도 5는 일 실시예에 따른 대화 시스템의 대화 방법 순서도이다.
- [0142] 도 6은 일 실시예에 따른 대화 시스템의 대화 방법 중 도메인 결정 방법의 상세 순서도이다.
- [0143] 도 7은 일 실시예에 따른 대화 시스템의 대화 방법 중 도메인 결정 방법의 예시도이다.
- [0144] 도 8은 다른 실시예에 따른 대화 시스템에 마련된 도메인 결정부의 상세 구성도이다.
- [0145] 도 9는 다른 실시예에 따른 대화 시스템의 대화 방법 중 도메인 결정 방법의 상세 순서도이다.

[0146] *도면의 주요부분에 대한 부호 설명*

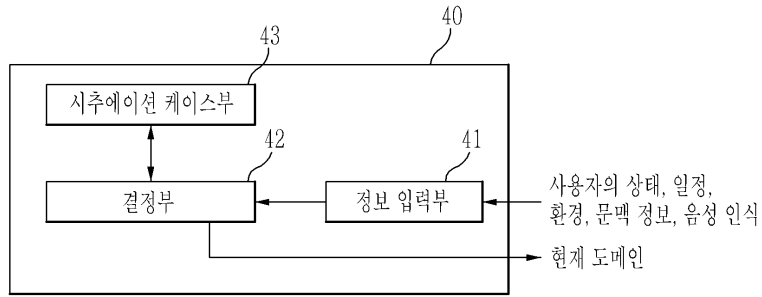
- [0147] 10: 멀티 모달 입력부
- 20: 멀티 모달 입력 처리부
- [0148] 30: 대화 관리부
- 40: 도메인 결정부
- [0149] 50: 데이터베이스부
- 60: 정보제공부
- [0150] 70: 멀티 모달 출력 처리부
- 80: 멀티 모달 출력부

도면

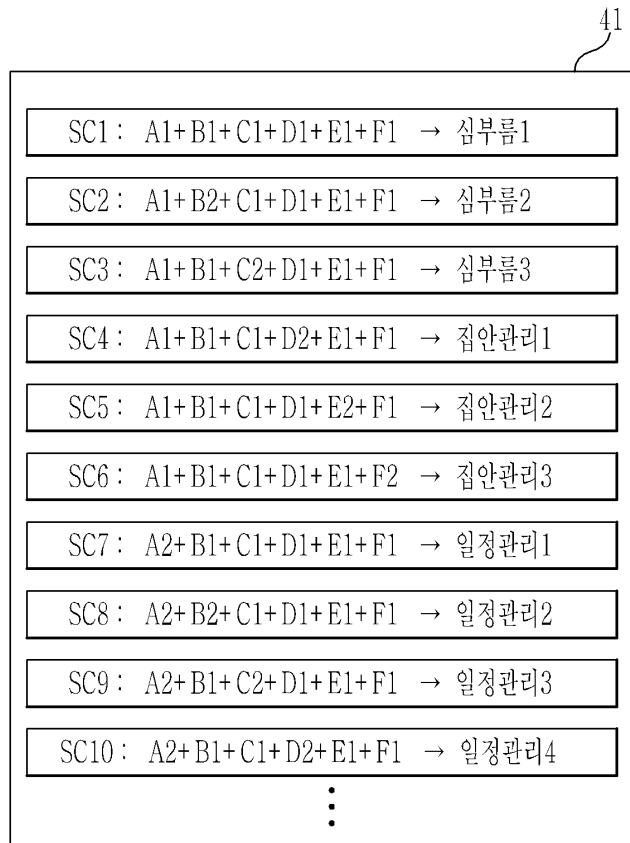
도면1



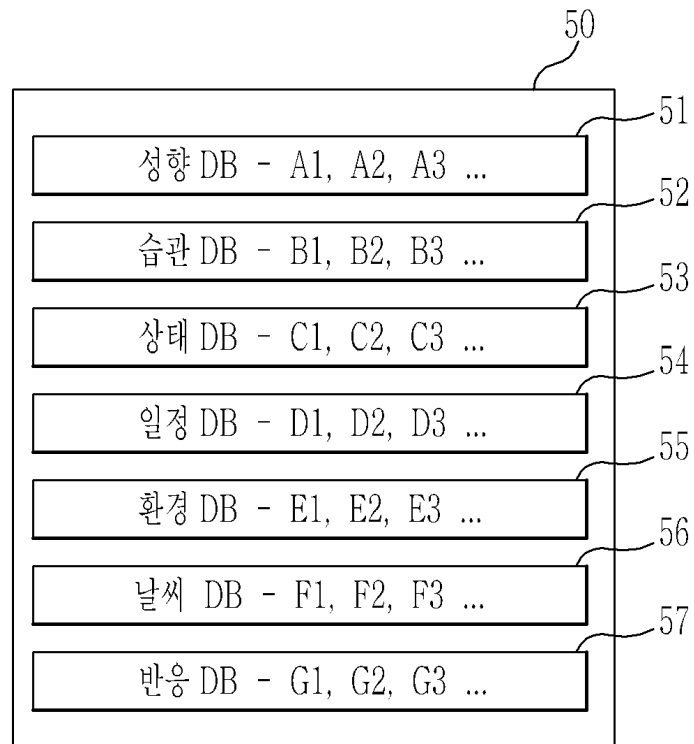
도면2



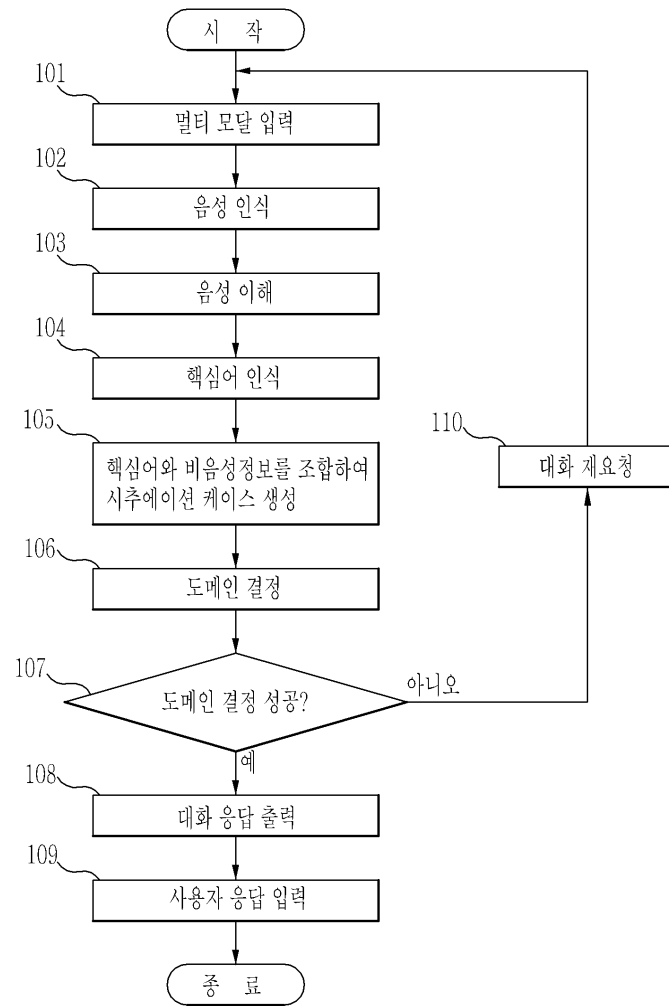
도면3



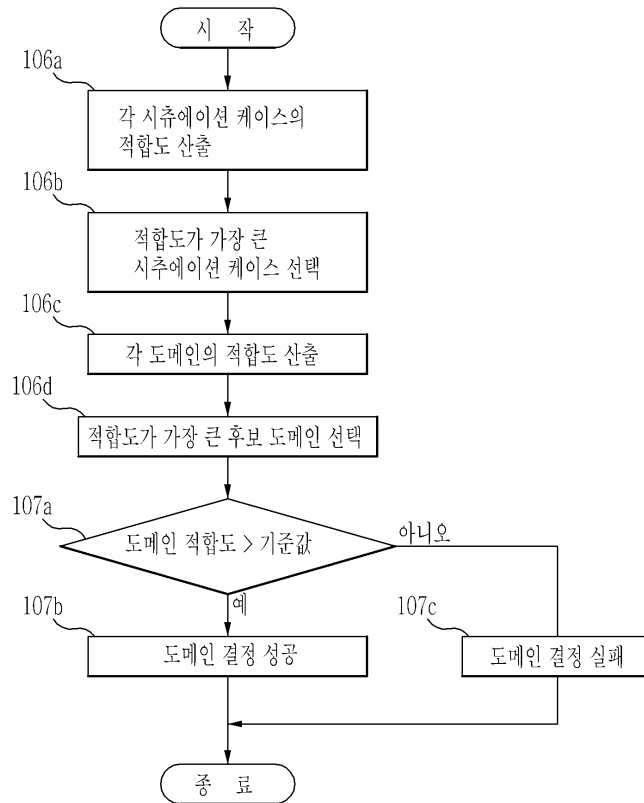
도면4



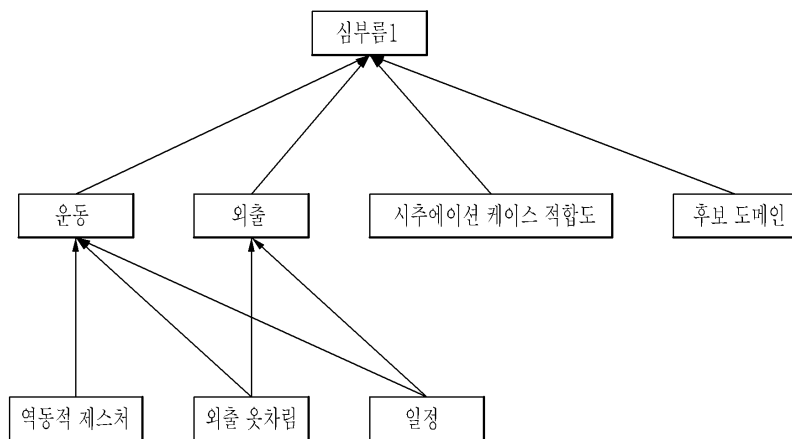
도면5



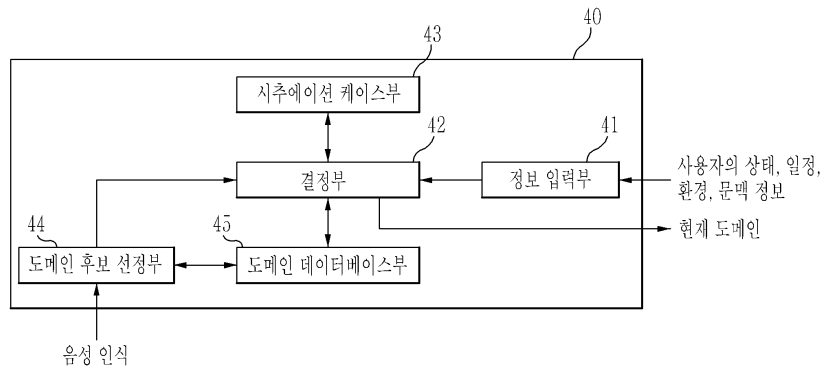
도면6



도면7



도면8



도면9

