



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0081925
(43) 공개일자 2020년07월08일

(51) 국제특허분류(Int. Cl.)
G10L 15/22 (2006.01) B25J 11/00 (2006.01)
B25J 13/00 (2006.01) B25J 9/16 (2006.01)
G10L 13/033 (2013.01) G10L 15/02 (2006.01)
(52) CPC특허분류
G10L 15/22 (2013.01)
B25J 11/0005 (2013.01)
(21) 출원번호 10-2018-0171954
(22) 출원일자 2018년12월28일
심사청구일자 2018년12월28일

(71) 출원인
수상에스티(주)
부산광역시 금정구 장전로 54-1, 3층(장전동)
(72) 발명자
이성중
부산광역시 금정구 수림로 85번길 14-1 (장전동)
(74) 대리인
하동엽

전체 청구항 수 : 총 7 항

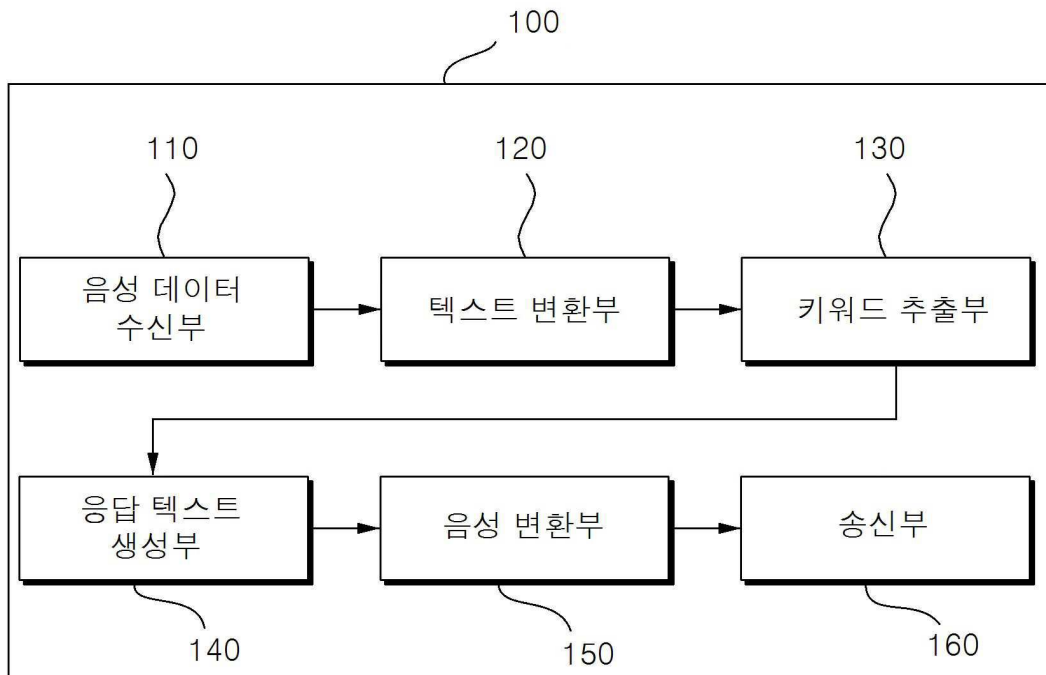
(54) 발명의 명칭 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법

(57) 요약

본 발명은 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법을 개시한다. 본 발명의 일실시예에 따른 교감형 로봇의 음성인식 시스템은, 외부 단말로부터 전송되는 음성 데이터를 수신하는 음성 데이터 수신부, 상기 음성 데이터를 텍스트로 변환하는 텍스트 변환부, 상기 변환된 텍스트로부터 키워드를 추출하는 키워드 추

(뒷면에 계속)

대표도 - 도2



출부, 상기 추출된 키워드에 대응하는 응답 텍스트를 기저장된 메타데이터로부터 추출하는 응답 텍스트 생성부, 상기 응답 텍스트를 음성 데이터로 변환하는 음성 변환부, 및 상기 변환된 음성 데이터를 상기 외부 단말로 송신하는 송신부를 포함할 수 있다.

본 발명의 일실시예에 따른 교감형 로봇은, 버튼 조작을 통해 음성 입력을 개시하기 위한 명령을 입력받는 음성 인식 버튼부, 사용자로부터 발화되는 음성을 입력받는 음성 입력부, 상기 입력된 음성의 녹음 데이터를 PCM data 형태로 외부 시스템에 전송하는 음성 전송부 및 상기 외부 시스템으로부터의 응답 데이터를 수신하고 출력하는 음성 출력부를 포함한다.

(52) CPC특허분류

B25J 13/003 (2013.01)

B25J 9/1602 (2013.01)

G06F 21/30 (2013.01)

G10L 13/033 (2013.01)

G10L 13/08 (2013.01)

G10L 15/02 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 S2457316

부처명 중소벤처기업부

연구관리전문기관 중소기업기술정보진흥원

연구사업명 기술혁신개발사업

연구과제명 음성인식 기능을 가진 감성 교감형 봉제 지능 로봇 개발

기 여 율 1/1

주관기관 수상에스티주식회사

연구기간 2017.01.09 ~ 2019.01.08

명세서

청구범위

청구항 1

외부 단말로부터 전송되는 음성 데이터를 수신하는 음성 데이터 수신부;
 상기 음성 데이터를 텍스트로 변환하는 텍스트 변환부;
 상기 변환된 텍스트로부터 키워드를 추출하는 키워드 추출부;
 상기 추출된 키워드에 대응하는 응답 텍스트를 기저장된 메타데이터로부터 추출하는 응답 텍스트 생성부;
 상기 응답 텍스트를 음성 데이터로 변환하는 음성 변환부; 및
 상기 변환된 음성 데이터를 상기 외부 단말로 송신하는 송신부;
 를 포함하는 교감형 로봇의 음성인식 시스템.

청구항 2

제 1항에 있어서,
 상기 외부 단말의 사용자를 식별하는 고유 키(primary key)를 수신하고, 상기 고유 키에 대응하는 설정값을 독출하는 사용자 관리부
 를 더 포함하는 것을 특징으로 하는 교감형 로봇의 음성인식 시스템.

청구항 3

제 1항에 있어서,
 상기 키워드 추출부는,
 상기 변환된 텍스트에 존재하는 다수의 명사를 추출하고,
 상기 명사의 유사어 셋(set)을 생성하고, 상기 추출된 명사의 카테고리를 기설정된 카테고리에 매칭하여,
 추출된 키워드마다 유사어 셋과 카테고리 속성을 부여하는 것을 특징으로 하는 교감형 로봇의 음성인식 시스템.

청구항 4

제 3항에 있어서,
 상기 응답 텍스트 생성부는,
 상기 추출된 각각의 키워드의 유사어 셋과 카테고리 속성에 대응하여 연관된 질문 리스트 셋(set)을 각각 추출하고, 상기 질문 리스트 간의 공통 질문을 추출하여 상기 응답 텍스트를 생성하는 것을 특징으로 하는 교감형 로봇의 음성인식 시스템.

청구항 5

버튼 조작을 통해 음성 입력을 개시하기 위한 명령을 입력받는 음성인식 버튼부;
 사용자로부터 발화되는 음성을 입력받는 음성 입력부;
 상기 입력된 음성의 녹음 데이터를 PCM data 형태로 외부 시스템에 전송하되, 기설정된 MTU 단위로 데이터를 분할하여 전송하는 음성 전송부; 및
 상기 외부 시스템으로부터의 응답 데이터를 수신하고 출력하는 음성 출력부;
 를 포함하는 교감형 로봇.

청구항 6

제 5항에 있어서,

상기 출력부는,

음성코덱의 각 레지스터의 딜레이 값이 0인지 확인하고, 0이 아닌 경우에는 음성 코덱의 설정동작 대기를 위한 딜레이 함수를 콜하여 각 레지스터의 딜레이 값에 상응하는 대기시간을 부여하는 것을 특징으로 하는 교감형 로봇.

청구항 7

외부 단말로부터 전송되는 음성 데이터를 수신하는 단계;

상기 음성 데이터를 텍스트로 변환하는 단계;

상기 변환된 텍스트로부터 키워드를 추출하는 단계;

상기 추출된 키워드에 대응하는 응답 텍스트를 기저장된 메타데이터로부터 추출하는 단계;

상기 응답 텍스트를 음성 데이터로 변환하는 단계; 및

상기 변환된 음성 데이터를 상기 외부 단말로 송신하는 단계;

를 포함하는 교감형 로봇의 음성인식 방법.

발명의 설명

기술 분야

- [0001] 본 발명은 음성인식 교감형 로봇, 교감형 로봇의 음성인식 시스템 및 그 방법에 관한 것으로, 보다 상세하게는 교감형 로봇을 통해 사용자의 음성을 인식하고, 상응하는 이벤트를 생성하는 시스템 및 그 방법에 관한 것이다.

배경 기술

- [0002] 사용자와 교감이 가능한 로봇, 인형 등은 유아나 어린이가 가지고 놀면서 신체 운동 발달 및 기능을 숙달하고, 상상력이나 창의력 개발을 통해 지능 발달 등 교육적으로 중요한 역할을 하기 때문에, 관련하여 교감형 로봇 또는 인형 기술 개발이 크게 관심을 받고 있다.
- [0003] 다만, 기존의 로봇 또는 인형은 제한된 소리를 출력하거나, 동작이 없으므로 사용자로 하여금 지속적으로 새로운 관심과 흥미를 끌어내기 어려운 한계가 있었다.
- [0004] 따라서, 사용자의 소리를 인식하여 응답하되, 사용자의 음성입력으로부터 사용자의 의도를 파악하고 이에 상응하는 응답을 표현할 수 있는 로봇 및 음성인식 시스템에 관한 연구가 필요하다.
- [0005] 선행기술문헌 : 한국등록특허 제10-1791942호

발명의 내용

해결하려는 과제

- [0006] 본 발명은 음성인식 교감형 로봇을 통해 사용자의 음성을 입력받고 이를 서버에 송신하여 서버에서 사용자의 음성을 분석하고 상응하는 응답 음성을 출력하도록 함으로써, 음성인식 교감형 로봇의 처리능력을 최소화하고, 비용을 절감할 수 있는 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법을 제공한다.
- [0007] 본 발명은 음성인식 교감형 로봇에 입력되는 음성을 서버로 전송하여 처리하되, MTU(Maximum Transmission Unit) 단위를 조정하여 데이터를 분할 전송함으로써, 상대적으로 낮은 사양의 하드웨어를 사용하면서도 고속의 음성인식이 가능해지는 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법을 제공한다.
- [0008] 본 발명은 무선통신을 통해 서버에 접속되어 각 사용자에게 특화된 음성을 분석하여 그에 상응하는 음성을 출력함으로써, 개별 사용자의 언어습관 등 특성에 부합하여 보다 정확한 음성인식이 가능한 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법을 제공한다.

[0009] 본 발명은 입력된 음성을 텍스트로 변환하고 상기 텍스트로부터 키워드를 추출하되, 추출된 키워드의 유사어, 카테고리 속성을 추출함으로써, 상기 유사어와 카테고리 속성에 대응하는 응답 텍스트를 보다 효과적으로 생성할 수 있는 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법을 제공한다.

과제의 해결 수단

[0010] 본 발명의 일실시예에 따른 교감형 로봇의 음성인식 시스템은, 외부 단말로부터 전송되는 음성 데이터를 수신하는 음성 데이터 수신부, 상기 음성 데이터를 텍스트로 변환하는 텍스트 변환부, 상기 변환된 텍스트로부터 키워드를 추출하는 키워드 추출부, 상기 추출된 키워드에 대응하는 응답 텍스트를 기저장된 메타데이터로부터 추출하는 응답 텍스트 생성부, 상기 응답 텍스트를 음성 데이터로 변환하는 음성 변환부, 및 상기 변환된 음성 데이터를 상기 외부 단말로 송신하는 송신부를 포함할 수 있다.

[0011] 본 발명의 일측에 따르면, 상기 외부 단말의 사용자를 식별하는 고유 키(primary key)를 수신하고, 상기 고유 키에 대응하는 설정값을 독출하는 사용자 관리부를 더 포함할 수 있다.

[0012] 본 발명의 일측에 따르면, 상기 키워드 추출부는, 상기 변환된 텍스트에 존재하는 다수의 명사를 추출하고, 상기 명사의 유사어 셋(set)을 생성하고, 상기 추출된 명사의 카테고리를 기설정된 카테고리에 매칭하여, 추출된 키워드마다 유사어 셋과 카테고리 속성을 부여할 수 있다.

[0013] 본 발명의 일측에 따르면, 상기 응답 텍스트 생성부는, 상기 추출된 각각의 키워드의 유사어 셋과 카테고리 속성에 대응하여 연관된 질문 리스트 셋(set)을 각각 추출하고, 상기 질문 리스트 간의 공통 질문을 추출하여 상기 응답 텍스트를 생성할 수 있다.

[0014] 본 발명의 일실시예에 따른 교감형 로봇은, 버튼 조작을 통해 음성 입력을 개시하기 위한 명령을 입력받는 음성 인식 버튼부, 사용자로부터 발화되는 음성을 입력받는 음성 입력부, 상기 입력된 음성의 녹음 데이터를 PCM data 형태로 외부 시스템에 전송하는 음성 전송부 및 상기 외부 시스템으로부터의 응답 데이터를 수신하고 출력하는 음성 출력부를 포함한다.

[0015] 본 발명의 일측에 따르면, 상기 출력부는, 음성코덱의 각 레지스터의 딜레이 값이 0인지 확인하고, 0이 아닌 경우에는 음성 코덱의 설정동작 대기를 위한 딜레이 함수를 콜하여 각 레지스터의 딜레이 값에 상응하는 대기시간을 부여할 수 있다.

[0016] 본 발명의 일실시예에 따른 교감형 로봇의 음성인식 방법은, 외부 단말로부터 전송되는 음성 데이터를 수신하는 단계, 상기 음성 데이터를 텍스트로 변환하는 단계, 상기 변환된 텍스트로부터 키워드를 추출하는 단계, 상기 추출된 키워드에 대응하는 응답 텍스트를 기저장된 메타데이터로부터 추출하는 단계, 상기 응답 텍스트를 음성 데이터로 변환하는 단계 및 상기 변환된 음성 데이터를 상기 외부 단말로 송신하는 단계를 포함한다.

발명의 효과

[0017] 본 발명의 일실시예에 따르면, 음성인식 교감형 로봇을 통해 사용자의 음성을 입력받고 이를 서버에 송신하여 서버에서 사용자의 음성을 분석하고 상응하는 응답 음성을 출력하도록 함으로써, 음성인식 교감형 로봇의 처리능력을 최소화하고, 비용을 절감할 수 있는 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법이 제공된다.

[0018] 본 발명의 일실시예에 따르면, 음성인식 교감형 로봇에 입력되는 음성을 서버로 전송하여 처리하되, MTU(Maximum Transmission Unit) 단위를 조정하여 데이터를 분할 전송함으로써, 상대적으로 낮은 사양의 하드웨어를 사용하면서도 고속의 음성인식이 가능해지는 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법이 제공된다.

[0019] 본 발명의 일실시예에 따르면, 무선통신을 통해 서버에 접속되어 각 사용자에게 특화된 음성을 분석하여 그에 상응하는 음성을 출력함으로써, 개별 사용자의 언어습관 등 특성에 부합하여 보다 정확한 음성인식이 가능한 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법이 제공된다.

[0020] 본 발명의 일실시예에 따르면, 입력된 음성을 텍스트로 변환하고 상기 텍스트로부터 키워드를 추출하되, 추출된 키워드의 유사어, 카테고리 속성을 추출함으로써, 상기 유사어와 카테고리 속성에 대응하는 응답 텍스트를 보다 효과적으로 생성할 수 있는 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법이 제공된다.

도면의 간단한 설명

- [0021] 도 1은 본 발명의 실시예에 따른 음성인식 교감형 로봇을 통해 음성을 인식받고, 이를 교감형 로봇 음성인식 시스템으로 전달하여 음성인식에 대한 이벤트를 발생시키기 위한 시스템과 로봇 전체 구성을 나타낸 도면이다.
- 도 2는 본 발명의 실시예에 따른 교감형 로봇 음성인식 시스템의 세부구성을 나타낸 블록도이다.
- 도 3은 본 발명의 실시예에 따른 음성인식 교감형 로봇의 세부구성을 나타낸 블록도이다.
- 도 4는 본 발명의 실시예에 따른 교감형 로봇 음성인식 방법의 흐름을 나타낸 동작흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0022] 이하, 첨부된 도면들에 기재된 내용들을 참조하여 본 발명의 실시예들을 상세하게 설명한다. 다만, 본 발명이 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조부호는 동일한 부재를 나타낸다.
- [0024] 종래에 로봇, 인형 등을 통해 사용자의 음성을 인식하고 이에 대한 응답 메시지를 출력하는 기술은 사용자의 음성을 통해 사용자 질문의 의도를 정확하게 파악하지 못하였으며, 이에 따라 응답 메시지 또한 단순한 메시지로 이루어지는 등의 문제점이 있었다.
- [0025] 본 발명은 상기 종래 기술의 문제점을 해결하기 위해 고안된 발명으로, 본 발명의 구성을 아래에 상세하게 설명한다.
- [0026] 도 1은 본 발명의 실시예에 따른 음성인식 교감형 로봇을 통해 음성을 인식받고, 이를 교감형 로봇 음성인식 시스템으로 전달하여 음성인식에 대한 이벤트를 발생시키기 위한 시스템과 로봇 전체 구성을 나타낸 도면이다.
- [0027] 도 1을 참고하면, 연결된 스마트 기기(300)으로 음성인식 교감형 로봇(200)이 에이전트 서버와 통신을 하기 위한 절차를 마친 뒤 음성인식 교감형 로봇(200)을 통해 사용자가 인사, 질문, 감정표현 메시지 등을 음성을 통해 입력하면, 상기 입력된 음성 데이터는 로봇 음성인식 시스템(100)으로 전송할 수 있다.
- [0028] 이후, 로봇 음성인식 시스템(100)은 상기 음성 데이터를 텍스트로 변환하고, 키워드를 추출하여, 추출된 키워드에 상응하는 응답 텍스트를 생성하고, 이를 음성인식 교감형 로봇(200)에 송신하면, 음성인식 교감형 로봇(200)은 스피커 등을 통해 이를 출력하여 상기 사용자와 교감할 수 있다.
- [0029] 이때, 사용자는 음성인식 교감형 로봇(200)을 통해 음성을 입력할 수 있으며, 응답 음성을 음성인식 교감형 로봇(200)을 통하여 확인할 수 있다.
- [0030] 이하에서는 음성을 입력하고 이에 대응하는 응답 텍스트를 생성하기 위한 세부절차와 구성을 보다 상세하게 설명한다.
- [0031] 도 2는 본 발명의 실시예에 따른 교감형 로봇 음성인식 시스템의 세부구성을 나타낸 블록도이다.
- [0032] 도 2를 참고하면, 교감형 로봇 음성인식 시스템(100)은 음성 데이터 수신부(110), 텍스트 변환부(120), 키워드 추출부(130), 응답 텍스트 생성부(140), 음성 변환부(150) 및 송신부(160)를 포함한다.
- [0033] 음성 데이터 수신부(110)는 외부 단말로부터 전송되는 음성 데이터를 수신할 수 있다. 즉, 교감형 로봇(200) 또는 이와 연결된 스마트 기기(300)를 통해 사용자의 음성이 인식되면, 음성 데이터 수신부(110)는 이를 전송받아 수신할 수 있다.
- [0034] 이때, 수신하는 음성 데이터는 PCM data를 포함한 다양한 형태의 데이터일 수 있다.
- [0035] 일례로, 상기 교감형 로봇(200)은 Artik053을 사용할 수 있는데, 상기 Artik053에서 사용자의 음성이 인식되면 이는 상기 음성 데이터 수신부(110)로 전송되고, 이 과정에서 전송되는 음성 데이터가 설정된 MTU 값 이상이 되면, 여러 개의 패킷으로 분할되어 전송되도록 하여, 상기와 같이 상대적으로 낮은 사양의 하드웨어(ex. Artik053)를 사용하는 교감형 로봇(200)을 통해서도 고속의 음성인식이 지원될 수 있다.
- [0036] 또한, Artik053과 같은 저사양 하드웨어에서는 API 서비스를 사용하기 위한 SDK를 설치할 수 없으므로 상기와 같이 음성 데이터를 작은 단위의 패킷으로 나누어 전달 받으면 기존 STT API에서 마이크 입력으로 처리되던 부

분을 상기와 같이 통신을 통해 전달 받은 음성 데이터를 받아오는 형식으로 변환하여 사용할 수 있고, 이를 통해 본 발명의 일실시례와 같은 저사양 하드웨어에서도 음성을 텍스트로 신속히 변환하는 스트리밍 서비스를 제공할 수 있다.

- [0037] 상기 음성 데이터를 수신한 이후 텍스트 변환부(120)는 상기 음성 데이터를 텍스트로 변환할 수 있다. 이때, 교감형 로봇 음성인식 시스템(100)은 개별 사용마다 언어습관 등이 다르므로, 개별 사용자의 특성에 맞도록 상기 외부 단말의 사용자를 식별하는 고유 키(primary key)를 수신하고, 상기 고유 키에 대응하는 설정값을 독출하기 위해 사용자 관리부를 더 포함할 수 있다.
- [0038] 즉, 사용자마다 음성인식 및 텍스트 변환을 위한 설정값을 달리하여 개별 사용자에게 최적화된 음성인식 및 텍스트 변환 과정을 진행함으로써, 사용자 맞춤형 음성인식이 이루어질 수 있다.
- [0039] 한편, 음성을 텍스트로 변환하기 위한 STT(Speech to Text) 과정에서는 클라우드를 통해 지원되는 API 등을 사용할 수 있으며, 120개 이상의 언어와 방언을 인식하고, 머신러닝 기술을 사용하여 실시간 스트리밍 또는 사전 녹음 오디오를 처리할 수 있다.
- [0040] 상기에서 음성 데이터가 텍스트로 변환되면, 키워드 추출부(130)는 상기 변환된 텍스트로부터 핵심 키워드를 추출할 수 있다.
- [0041] 이를 위해, 상기 변환된 텍스트에 존재하는 다수의 명사를 추출하고, 상기 명사의 유사어 셋(set)을 생성하고, 상기 추출된 명사의 카테고리를 기설정된 카테고리에 매칭하여, 추출된 키워드마다 유사어 셋과 카테고리 속성을 부여할 수 있다.
- [0042] 일례로, 사용자가 입력한 문장이 "내일 소풍 갈거야"인 경우, 상기 문장에 포함된 명사인 '내일', 과 '소풍'을 추출하고, '내일'의 유사어인 'tomorrow', '다음날', '이튿날' 등의 유사어 셋(set)을 추출하며, '내일'은 시간을 나타내는 단어이므로 카테고리 속성으로 '시간 단어'를 부여할 수 있다.
- [0043] 또한, '소풍'의 경우에도 유사어인 'picnic', '나들이', '야유회' 등의 유사어 셋(set)을 추출하고, '소풍'은 야외에서의 행동을 나타내는 단어이므로 카테고리 속성으로 '야외행동 단어'를 부여할 수 있다.
- [0044] 따라서, 상기 키워드에 대한 유사어 셋과 카테고리 속성을 이용하여 하기에 설명될 응답 텍스트 생성부(140)에서 사용자의 의도를 보다 정확하게 파악하여 그에 상응하는 응답 텍스트를 도출하도록 할 수 있다.
- [0045] 따라서, 이와 연관하여 응답 텍스트 생성부(140)는 상기 추출된 키워드에 대응하는 응답 텍스트를 기저장된 메타데이터로부터 추출할 수 있다.
- [0046] 이를 위해 응답 텍스트 생성부(140)는, 상기 추출된 각각의 키워드의 유사어 셋과 카테고리 속성에 대응하여 연관된 질문 리스트 셋(set)을 각각 추출하고, 상기 질문 리스트 간의 공통 질문을 추출하여 상기 응답 텍스트를 생성할 수 있다.
- [0047] 일례로, 추출된 키워드 '내일'과 관련된 질문 리스트 셋에 질문이 5개 포함되고, '소풍'과 관련된 질문 리스트 셋에 질문이 7개 포함된 경우, 상기 두 질문 리스트 셋에 내용이 최대한 중첩되는 질문을 하나 추출하고, 이를 사용자에게 응답할 텍스트로 결정할 수 있다.
- [0048] 이에 따라 음성 변환부(150)는 상기 결정된 응답 텍스트를 음성 데이터로 변환할 수 있다.
- [0049] 여기서, 텍스트를 음성 변환하기 위한 TTS(Text to Speech) 과정에서는 딥러닝 기술을 사용하여 실제 사람의 음성처럼 소리를 합성하고, 다양한 언어, 음성을 설정할 수 있으며, 상기 사용자의 설정값에 따라 사용자의 언어 습관과 유사한 형태로 음성을 생성할 수 있고, 사용자 지정어휘 또는 저장된 용어(회사 이름, 두문자어, 외래어, 신조어 등)에 따라 특정 단어의 발음을 반영하여 생성할 수 있다.
- [0050] 송신부(160)는 상기 변환된 음성 데이터를 상기 외부 단말로 송신할 수 있다. 이때, 송신되는 데이터는 MP3 형식을 포함한 다양한 형식의 음성 데이터로 송신할 수 있다.
- [0051] 상기와 같이, 교감형 로봇 음성인식 시스템을 사용하여 음성을 인식하고 이에 상응하는 응답 텍스트를 생성함으로써, 개별 사용자 맞춤형 음성인식 및 응답 텍스트 생성이 가능하며, 사용자의 의도를 보다 정확하게 파악하여 이에 부합하는 응답을 제공할 수 있는 효과가 발생할 수 있다.
- [0052] 이하에서는 음성인식 교감형 로봇을 통해 음성을 입력받고 이를 교감형 로봇 음성인식 시스템에 전송하며, 교감형 로봇 음성인식 시스템으로부터 응답 텍스트(음성 변환된 데이터)를 수신하여 출력하는 구성을 보다 상세하게

설명한다.

- [0053] 도 3은 본 발명의 실시예에 따른 음성인식 교감형 로봇의 세부구성을 나타낸 블록도이다. 이때, 음성인식 교감형 로봇(200)은 일례로 ARTIK053 보드를 내장할 수 있다.
- [0054] 도 3을 참고하면, 음성인식 교감형 로봇(200)은 음성인식 버튼부(210), 음성 입력부(220), 음성 전송부(230) 및 음성 출력부(240)를 포함할 수 있다.
- [0055] 음성인식 버튼부(210)는 버튼 조작을 통해 음성 입력을 개시하기 위한 명령을 입력받을 수 있다. 즉, 종래의 경우 스마트 스피커 등을 통해 소리를 감지하는 것으로 음성입력을 개시하나, 본 발명의 일실시예에서는 사용자가 버튼을 조작하여야만 음성 입력을 개시하므로, 사용자의 적극적인 동작을 통해 음성입력을 개시하기 전까지는 사용자들의 음성대화를 모니터링 하지 않으므로, 사용자의 의도와 무관하게 대화내용이 녹음되고 제3자에게 유출되지 않도록 관리될 수 있다.
- [0056] 상기 버튼은 교감형 로봇의 손 부분에 위치하여, 버튼의 조작시 로봇의 손을 잡는 감성을 제공함으로써, 사용자가 로봇과 보다 교감을 느끼도록 설정할 수 있다.
- [0057] 음성 입력부(220)는 사용자로부터 발화되는 음성을 입력받으며, 상기 음성인식 버튼을 통해 음성입력이 개시되면, 마이크 입력 등을 통해 사용자의 음성을 입력받을 수 있다.
- [0058] 음성 전송부(230)는 상기 입력된 음성의 녹음 데이터를 PCM data 형태로 외부 시스템에 전송할 수 있다. 즉, PCM data 형태로 데이터를 전송함으로써, 보다 효과적이고 손실없이 데이터를 전송할 수 있다.
- [0059] 한편, 음성 전송은 네트워크 인터페이스에서 세그먼트 없이 보낼 수 있는 최대 데이터그램 크기 값이며, 패킷이 한번에 보낼 수 있는 최대 크기인 MTU size를 590으로 설정 하여 전송할 수 있고, 보내는 데이터가 MTU 값 이상이 되면 여러 개의 패킷으로 분할되어 전송될 수 있어 보다 효과적으로 데이터 전송이 가능해질 수 있다.
- [0060] 음성 출력부(240)는 상기 외부 시스템으로부터 음성 데이터 형태의 응답 데이터를 수신하면 스피커 등의 출력장치를 통해 사용자가 인식할 수 있도록 응답 데이터를 출력할 수 있다.
- [0061] 한편, 상기 음성 출력부(240)는 음성 출력을 위해 사용되는 코덱 셋팅시 발생하는 지연을 최소화하기 위해 하기의 방법을 사용할 수 있다.
- [0062] 코덱 사용전의 코덱 레지스터 설정과정에서 코덱의 설정동작을 기다리기 위해 delay 함수를 콜(call)하여 script[i].delay 만큼의 대기시간을 가지는데, 실제로는 script[i].delay 값은 0인 경우가 많다. 따라서, script[i].delay 함수 자체를 콜하는 시간에 따른 지연을 방지하고자 각 레지스터의 script[i].delay 값이 0인지 확인하고, 0이 아닌 경우에만 음성 코덱의 설정동작 대기를 위한 script[i].delay 함수를 콜하여 각 레지스터의 딜레이 값에 상응하는 대기시간을 부여할 수 있다. 여기서, script[i]의 멤버는 레지스터 주소이며, script[i].delay는 각 레지스터의 딜레이 값에 해당한다.
- [0063] 상기와 같이 본 발명의 일실시예에 따른 교감형 로봇을 통해 음성인식 교감형 로봇의 처리능력을 최소화하고, 비용을 절감할 수 있는 장치가 제공될 수 있다.
- [0064] 도 4는 본 발명의 실시예에 따른 교감형 로봇 음성인식 방법의 흐름을 나타낸 동작흐름도이다.
- [0065] 한편, 하기에서는 음성인식 교감형 로봇(200)은 ARTIK053 보드를 내장하고, 소켓통신을 통해 교감형 로봇의 음성인식 시스템(100)과 통신하는 것을 일례로 설명한다.
- [0066] 이를 위해 단계(410)에서는 외부 단말로부터 전송되는 음성 데이터를 수신할 수 있다.
- [0067] 즉, ARTIK053 보드와의 소켓통신으로 Client(ARTIK053)가 교감형 로봇의 음성인식 시스템(100) Server에 접속하게 되면, 사용자 구분을 위한 User_info_check()가 실행되며, 사용자는 고유 값을 갖는 primary key를 통해 구분될 수 있다.
- [0068] Client 정보를 action_thread() 호출과 함께 넘겨주면 해당 client socket을 이용하여, google_cloud_streaming() 동작으로 사용자의 음성 데이터(PCM data)를 server로 가져올 수 있다.
- [0069] 다음으로 단계(420)에서는 상기 음성 데이터를 텍스트로 변환할 수 있다. 이때, Google cloud streaming Speech To Text API를 통해서 text로 변환할 수 있다.
- [0070] 단계(430)에서는 상기 변환된 텍스트로부터 키워드를 추출할 수 있고, 단계(440)에서는 상기 추출된 키워드에

대응하는 응답 텍스트를 기저장된 메타데이터로부터 추출할 수 있다.

[0071] 단계(450)에서는 상기 응답 텍스트를 음성 데이터로 변환할 수 있고, 단계(460)에서는 상기 변환된 음성 데이터를 상기 외부 단말로 송신할 수 있다.

[0072] 이를 위해 응답 text는 AWS Polly Text To Speech API를 이용하여 1-Chenal, Mono, 22050HZ의 mp3 file로 생성되고 해당 mp3 file을 FFmpeg module를 사용하여 2-Channel Stereo 44000HZ로 변환 및 ARTIK053 보드에 전달하는 과정이 진행될 수 있다.

[0074] 상기와 같이 본 발명의 일실시례에 따르면, 음성인식 교감형 로봇을 통해 사용자의 음성을 입력받고 이를 서버에 송신하여 서버에서 사용자의 음성을 분석하고 상응하는 응답 음성을 출력하도록 함으로써, 음성인식 교감형 로봇의 처리능력을 최소화하고, 비용을 절감할 수 있는 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법이 제공된다.

[0075] 본 발명의 일실시례에 따르면, 무선통신을 통해 서버에 접속되어 각 사용자에게 특화된 음성을 분석하여 그에 상응하는 음성을 출력함으로써, 개별 사용자의 언어습관 등 특성에 부합하여 보다 정확한 음성인식이 가능한 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법이 제공된다.

[0076] 본 발명의 일실시례에 따르면, 입력된 음성을 텍스트로 변환하고 상기 텍스트로부터 키워드를 추출하되, 추출된 키워드의 유사어, 카테고리 속성을 추출함으로써, 상기 유사어와 카테고리 속성에 대응하는 응답 텍스트를 보다 효과적으로 생성할 수 있는 음성인식 교감형 로봇, 교감형 로봇 음성인식 시스템 및 그 방법이 제공된다.

[0077] 또한, 본 발명의 일실시례에 따르면, 저사양 하드웨어를 이용하여 전력소모가 적고 가벼워 휴대가 용이해지고, 초기비용을 현저히 낮출 수 있으며, 사용자가 이동중에도 고속 음성인식 서비스를 제공받을 수 있는 효과가 발생될 수 있다.

[0079] 또한 본 발명의 일실시례에 따른, 교감형 로봇 음성인식 방법은 다양한 컴퓨터로 구현되는 동작을 수행하기 위한 프로그램 명령을 포함하는 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0080] 이상과 같이 본 발명의 일실시례는 비록 한정된 실시예와 도면에 의해 설명되었으나, 본 발명의 일실시례는 상기 설명된 실시예에 한정되는 것은 아니며, 이는 본 발명이 속하는 분야에서 통상의 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명의 일실시례는 아래에 기재된 특허청구범위에 의해서만 파악되어야 하고, 이의 균등 또는 등가적 변형 모두는 본 발명 사상의 범주에 속한다고 할 것이다.

부호의 설명

[0081] 100: 교감형 로봇 음성인식 시스템

110: 음성 데이터 수신부

120: 텍스트 변환부

130: 키워드 추출부

140: 응답 텍스트 생성부

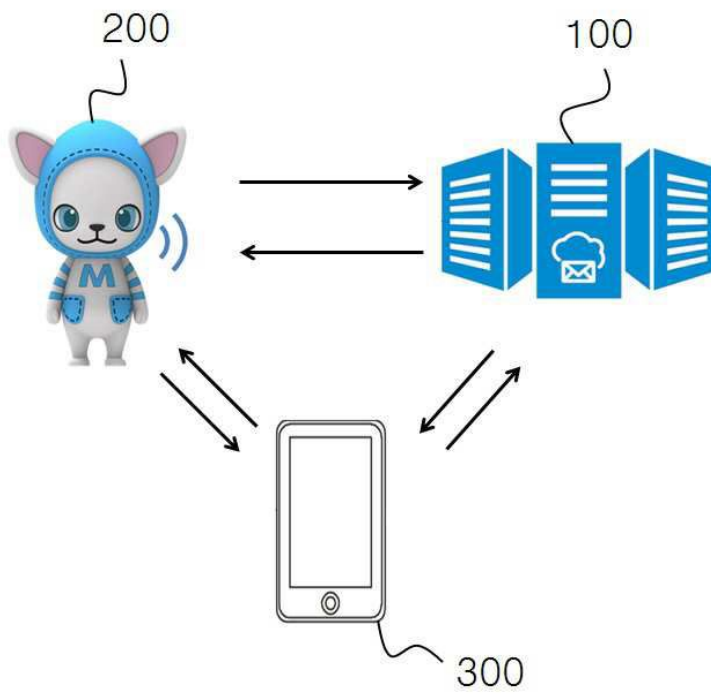
150: 음성 변환부

160: 송신부

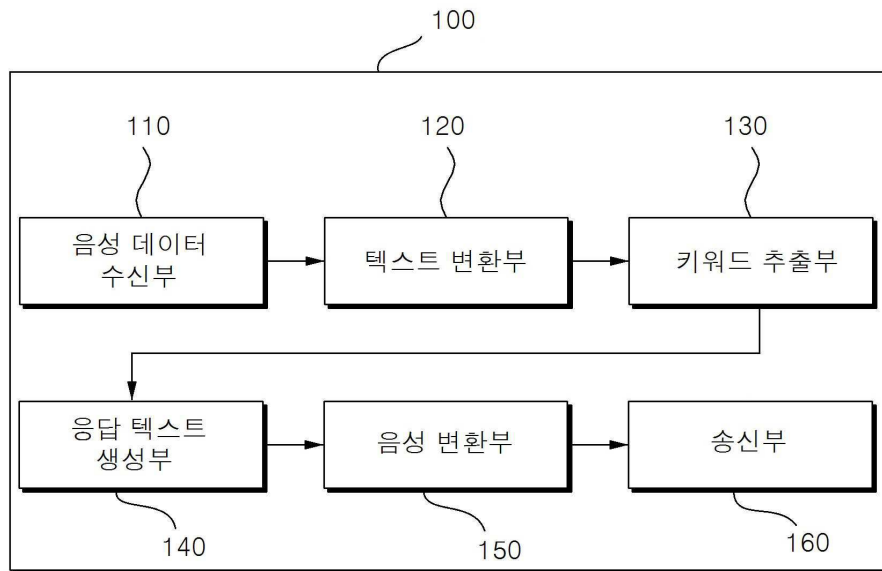
- 200: 교감형 로봇
- 210: 음성인식 버튼부
- 220: 음성 입력부
- 230: 음성 전송부
- 240: 음성 출력부

도면

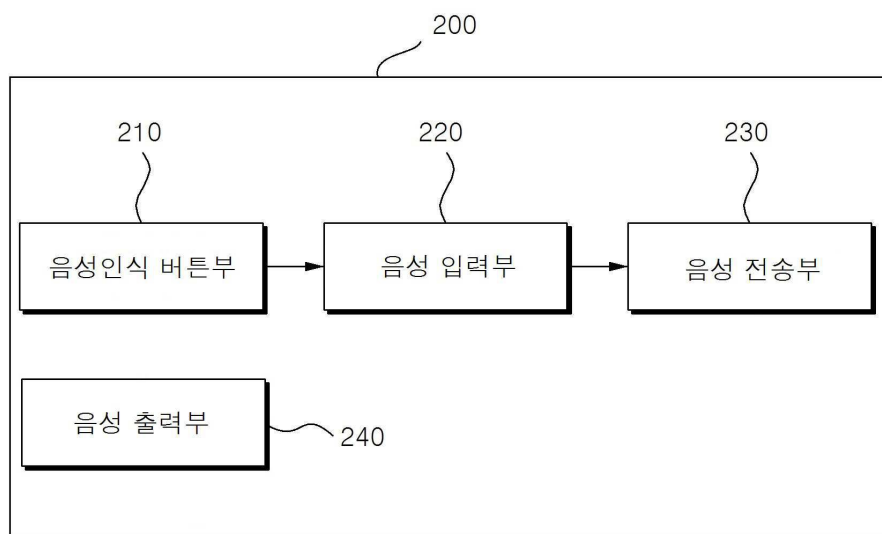
도면1



도면2



도면3



도면4

