

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2024-0003368
(43) 공개일자 2024년01월09일

(51) 국제특허분류(Int. Cl.)
G06Q 50/50 (2024.01) G10L 17/14 (2013.01)
G10L 17/22 (2013.01)
(52) CPC특허분류
G06Q 50/50 (2024.01)
G10L 15/26 (2013.01)
(21) 출원번호 10-2022-0080353
(22) 출원일자 2022년06월30일
심사청구일자 2022년06월30일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
이보원
인천광역시 미추홀구 인하로 100, 하이테크센터 1018호(용현동)
서지영
인천광역시 연수구 해송로 70, 206동 204호(송도동, 송도웹카운티2단지)
(뒷면에 계속)
(74) 대리인
양성보

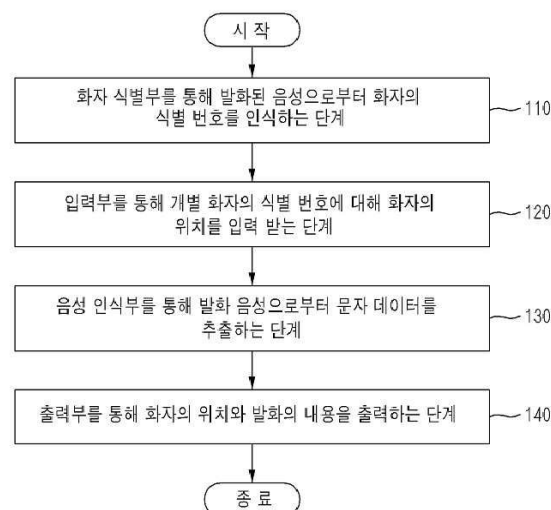
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 실시간 대화 표시 방법 및 시스템

(57) 요약

실시간 대화 표시 방법 및 시스템이 제시된다. 본 발명에서 제안하는 화자 식별부, 화자 위치 입력부, 음성 인식부 및 출력부를 포함하는 실시간 대화 표시 시스템의 실시간 대화 표시 방법은 실시간 대화 표시 시스템을 구비하고 있는 복수의 화자가 대화를 하는 상황에서 화자 식별부를 통해 발화 음성으로부터 개별 화자의 식별 번호를 인식하는 단계, 화자 위치 입력부를 통해 상기 개별 화자의 식별 번호에 대한 개별 화자의 위치를 입력 받는 단계, 음성 인식부를 통해 발화 음성으로부터 문자 데이터를 추출하는 단계 및 출력부를 통해 개별 화자의 위치와 발화의 내용을 출력하는 단계를 포함한다.

대표도 - 도1



(52) CPC특허분류

G10L 17/14 (2013.01)

G10L 17/22 (2013.01)

(72) 발명자

윤수연

인천광역시 미추홀구 인하로47번길 131, 오로라
502호(용현동)

김동현

경기도 고양시 일산서구 산현로 45, 1401동 1201
호(탄현동, 탄현마을14단지아파트)

정승훈

인천광역시 연수구 해돋이로6번길 7, 106동 402호
(송도동, 아이파크송도)

이영선

서울특별시 용산구 이촌로71길 10, 213동 502호(이
촌동, 한가람아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152759
과제번호	2020-0-01389-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능융합연구센터지원사업(국고)
연구과제명	[Ezbaro][정부] 인공지능융합연구센터지원(3차년도)
기 여 율	1/2
과제수행기관명	인하대학교 산학협력단
연구기간	2022.01.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1345345575
과제번호	2021S1A3A2A01087325
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	사회과학연구지원(R&D)
연구과제명	포용적이고 지속가능한 장애친화적 디지털 사회로의 전환
기 여 율	1/2
과제수행기관명	이화여자대학교
연구기간	2021.09.01 ~ 2022.05.31

명세서

청구범위

청구항 1

화자 식별부, 화자 위치 입력부, 음성 인식부 및 출력부를 포함하는 실시간 대화 표시 시스템의 실시간 대화 표시 방법에 있어서,

실시간 대화 표시 시스템을 구비하고 있는 복수의 화자가 대화를 하는 상황에서 화자 식별부를 통해 발화 음성으로부터 개별 화자의 식별 번호를 인식하는 단계;

화자 위치 입력부를 통해 상기 개별 화자의 식별 번호에 대한 개별 화자의 위치를 입력 받는 단계;

음성 인식부를 통해 발화 음성으로부터 문자 데이터를 추출하는 단계; 및

출력부를 통해 개별 화자의 위치와 발화의 내용을 출력하는 단계

를 포함하는 실시간 대화 표시 방법.

청구항 2

제1항에 있어서,

상기 화자 식별부를 통해 복수의 화자가 대화를 하는 상황에서 발화 음성으로부터 개별 화자의 식별 번호를 인식하는 단계는,

상기 개별 화자에 대한 고유한 식별 번호를 생성하고, 단일 또는 복수의 장치로부터 입력된 음성 신호를 통해 발화 중인 화자에 대응하는 식별 번호를 인식하는

실시간 대화 표시 방법.

청구항 3

제1항에 있어서,

상기 화자 위치 입력부를 통해 상기 개별 화자의 식별 번호에 대한 개별 화자의 위치를 입력 받는 단계는,

상기 복수의 화자 각각의 실시간 대화 표시 시스템의 출력부를 통해 문자 데이터 또는 발화자를 식별하기 위한 그래픽 객체를 출력할 위치를 입력 받는

실시간 대화 표시 방법.

청구항 4

제1항에 있어서,

상기 음성 인식부를 통해 발화 음성으로부터 문자 데이터를 추출하는 단계는,

음성인식을 위한 알고리즘을 통해 발화 음성으로부터 발화자가 의도하는 발화 내용을 포함하는 문자 데이터를 추출하는

실시간 대화 표시 방법.

청구항 5

제1항에 있어서,

상기 출력부를 통해 개별 화자의 위치와 발화의 내용을 출력하는 단계는,

상기 화자 식별부를 통해 인식된 개별 화자의 식별 번호와 상기 화자 위치 입력부를 통해 입력 받은 문자 데이터 또는 발화자를 식별하기 위한 그래픽 객체를 출력할 위치, 상기 음성 인식부를 통해 식별한 문자 데이터를

입력 받아 출력하는

실시간 대화 표시 방법.

청구항 6

실시간 대화 표시 시스템을 구비하고 있는 복수의 화자가 대화를 하는 상황에서 발화 음성으로부터 개별 화자의 식별 번호를 인식하는 화자 식별부;

상기 개별 화자의 식별 번호에 대한 개별 화자의 위치를 입력 받는 화자 위치 입력부;

발화 음성으로부터 문자 데이터를 추출하는 음성 인식부; 및

개별 화자의 위치와 발화의 내용을 출력하는 출력부

를 포함하는 실시간 대화 표시 시스템.

청구항 7

제6항에 있어서,

상기 화자 식별부는,

상기 개별 화자에 대한 고유한 식별 번호를 생성하고, 단일 또는 복수의 장치로부터 입력된 음성 신호를 통해 발화 중인 화자에 대응하는 식별 번호를 인식하는

실시간 대화 표시 시스템.

청구항 8

제6항에 있어서,

상기 화자 위치 입력부는,

상기 복수의 화자 각각의 실시간 대화 표시 시스템의 출력부를 통해 문자 데이터 또는 발화자를 식별하기 위한 그래픽 객체를 출력할 위치를 입력 받는

실시간 대화 표시 시스템.

청구항 9

제6항에 있어서,

상기 음성 인식부는,

음성인식을 위한 알고리즘을 통해 발화 음성으로부터 발화자가 의도하는 발화 내용을 포함하는 문자 데이터를 추출하는

실시간 대화 표시 시스템.

청구항 10

제6항에 있어서,

상기 출력부는,

상기 화자 식별부를 통해 인식된 개별 화자의 식별 번호와 상기 화자 위치 입력부를 통해 입력 받은 문자 데이터 또는 발화자를 식별하기 위한 그래픽 객체를 출력할 위치, 상기 음성 인식부를 통해 식별한 문자 데이터를 입력 받아 출력하는

실시간 대화 표시 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 실시간 대화 표시 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 사적인 대화 또는 업무를 위한 회의 등 복수의 참가자가 수행한 대화를 녹음하고, 녹음한 대화를 화자 별로 분리하여 텍스트형태의 스크립트로 옮길 필요가 있을 수 있다. 하지만 직접 음성 파일을 듣고 글자로 받아 적거나 음성인식 등으로 얻어진 텍스트만을 보고 화자를 구분하기 어려운 상황이 발생할 수 있다. 예를 들어 대화에 목소리가 비슷한 화자가 있거나 한 화자가 다양한 톤, 높낮이나 크기로 이야기를 할 경우에 녹음된 음성파일을 듣는 것만으로는 정확하게 화자를 나눠 스크립트를 작성할 수 없다. 따라서 현재 녹음된 음성 파일을 텍스트로 변환해주는 솔루션이 등장했지만 대화 종료 후 녹음된 음성을 텍스트로 전환만 해줄 뿐, 대화 내용의 실시간 변환 또는 텍스트 외의 부가 정보를 얻기에 부족하다. 따라서 발화 상황에서 이루어진 대화를 텍스트로 옮길 뿐 아니라, 대화의 텍스트 이외의 정보를 얻기 위해서는 실시간 화자 인식 및 음성 인식을 수행하는 기술을 고려하는 것이 바람직하다.

[0003] 이와 관련하여 한국 등록특허 제10-1396715호는 인스턴트 메시지를 포함하는 말풍선에 시각적 표현을 제공하는 방법 및 그 이동 단말에 대해 개시하고 있다. 해당 특허에서 인스턴트 메시지를 포함하는 말풍선에 시각적 표현을 제공하는 방법은 인스턴트 메시지 및 인스턴트 메시지에 대응하는 말풍선의 시각적 표현을 정의하는 외관 식별자를 수신하는 단계부터 채팅창에 식별된 시각적 표현을 갖는 말풍선을 디스플레이하는 단계를 포함한다.

[0004] 또한, 한국 등록특허 제10-2387268호는 말풍선 효과를 제공하는 메시징 서비스 제공 방법과 이를 위한 프로그램을 기록한 기록 매체 및 단말에 대해 개시하고 있다. 해당 발명은 사용자 간에 교환되는 메시지가 기록되는 말풍선을 이용하여 말풍선 효과를 제공하는 메시징 서비스 제공 방법, 이를 위한 프로그램을 기록한 기록 매체 및 사용자 단말에 관한 것으로서, 말풍선 효과를 선택하여 지정 하면, 쓰여진 말풍선이 변형되거나 상기 말풍선을 이용한 효과가 함께 출력되는 말풍선 효과를 제공한다.

[0005] 이와 관련하여 한국 공개특허 제10-2017-0018363호는 말풍선 꼬리정보로 발신인의 발신메시지에 대한 수신인의 읽음 확인 시스템에 대해 개시하고 있다. 발신자가 보낸 메시지를 수신자가 읽을 경우 발신인에게 읽음식별코드를 전송하고, 발신인의 스마트폰에서 읽음식별코드를 확인하면 말풍선의 모양을 바꿔 수신인이 읽었음을 시각적으로 간단하게 식별할 수 있는 방법을 제공한다.

[0006] 하지만, 상술한 종래기술들은 인스턴트 메시지에 새로운 정보 전달 없이 디자인에 초점을 맞추거나, 타인의 수신 확인을 위한 정보만을 담고 있다. 다른 화자라도 같은 구성 및 시각적 효과를 추가할 수 있고, 사용자가 식별자를 필요로 할 경우, 사용자의 의도에 따라 다른 사용자의 메시지에 시각적 정보를 수정할 수 없다는 한계가 있다. 따라서, 음성을 텍스트로 변환하고, 식별 효과를 필요로 하는 사용자가 발화자에 대한 정보를 직접 할당할 수 있다는 점에서 화자 인식 알고리즘에 대한 연구 및 개발이 요구된다.

선행기술문헌

특허문헌

- [0007] (특허문헌 0001) 한국 등록특허 제10-1396715호(2014.05.12)
(특허문헌 0002) 한국 등록특허 제10-2387268호(2022.04.12)
(특허문헌 0003) 한국 공개특허 제10-2017-0018363호(2017.02.17)

발명의 내용

해결하려는 과제

[0008] 본 발명이 이루고자 하는 기술적 과제는 사적인 대화 또는 업무 회의 등 복수의 참가자가 대화를 하는 상황에서 대화 내용을 실시간으로 텍스트로 변환하고, 발화한 화자를 인식하여 각 화자의 상대적 위치 및 대화 내용을 직관적으로 파악할 수 있도록 표시하기 위한 실시간 대화 표시 방법 및 시스템을 제공하는데 있다.

과제의 해결 수단

- [0009] 일 측면에 있어서, 본 발명에서 제안하는 화자 식별부, 화자 위치 입력부, 음성 인식부 및 출력부를 포함하는 실시간 대화 표시 시스템의 실시간 대화 표시 방법은 실시간 대화 표시 시스템을 구비하고 있는 복수의 화자가 대화를 하는 상황에서 화자 식별부를 통해 발화 음성으로부터 개별 화자의 식별 번호를 인식하는 단계, 화자 위치 입력부를 통해 상기 개별 화자의 식별 번호에 대한 개별 화자의 위치를 입력 받는 단계, 음성 인식부를 통해 발화 음성으로부터 문자 데이터를 추출하는 단계 및 출력부를 통해 개별 화자의 위치와 발화의 내용을 출력하는 단계를 포함한다.
- [0010] 상기 화자 식별부를 통해 복수의 화자가 대화를 하는 상황에서 발화 음성으로부터 개별 화자의 식별 번호를 인식하는 단계는 상기 개별 화자에 대한 고유한 식별 번호를 생성하고, 단일 또는 복수의 장치로부터 입력된 음성 신호를 통해 발화 중인 화자에 대응하는 식별 번호를 인식한다.
- [0011] 상기 화자 위치 입력부를 통해 상기 개별 화자의 식별 번호에 대한 개별 화자의 위치를 입력 받는 단계는 상기 복수의 화자 각각의 실시간 대화 표시 시스템의 출력부를 통해 문자 데이터 또는 발화자를 식별하기 위한 그래픽 객체를 출력할 위치를 입력 받는다.
- [0012] 상기 음성 인식부를 통해 발화 음성으로부터 문자 데이터를 추출하는 단계는 음성인식을 위한 알고리즘을 통해 발화 음성으로부터 발화자가 의도하는 발화 내용을 포함하는 문자 데이터를 추출한다.
- [0013] 상기 출력부를 통해 개별 화자의 위치와 발화의 내용을 출력하는 단계는 상기 화자 식별부를 통해 인식된 개별 화자의 식별 번호와 상기 화자 위치 입력부를 통해 입력 받은 문자 데이터 또는 발화자를 식별하기 위한 그래픽 객체를 출력할 위치, 상기 음성 인식부를 통해 식별한 문자 데이터를 입력 받아 출력한다.
- [0014] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 실시간 대화 표시 시스템은 실시간 대화 표시 시스템을 구비하고 있는 복수의 화자가 대화를 하는 상황에서 발화 음성으로부터 개별 화자의 식별 번호를 인식하는 화자 식별부, 상기 개별 화자의 식별 번호에 대한 개별 화자의 위치를 입력 받는 화자 위치 입력부, 발화 음성으로부터 문자 데이터를 추출하는 음성 인식부 및 개별 화자의 위치와 발화의 내용을 출력하는 출력부를 포함한다

발명의 효과

- [0015] 본 발명의 실시예들에 따른 실시간 대화 표시 방법 및 시스템을 통해 사적인 대화 또는 업무 회의 등 복수의 참가자가 대화를 하는 상황에서 대화 내용을 실시간으로 텍스트로 변환하고, 발화한 화자를 인식하여 각 화자의 상대적 위치 및 대화 내용을 직관적으로 파악할 수 있도록 표시할 수 있다. 본 발명의 실시예들에 따르면 화자의 위치 정보 및 음성인식 기술을 사용하여 실시간으로 대화 내용 및 해당 발화자의 정보가 표시되도록 하여 대화의 효율성을 높이며 나아가서는 고령자, 청각장애인 등이 대화 내용을 직접 듣지 못하는 경우에도 효과적으로 적용이 가능하다.

도면의 간단한 설명

- [0016] 도 1은 본 발명의 일 실시예에 따른 실시간 대화 표시 방법을 설명하기 위한 흐름도이다.
- 도 2는 본 발명의 일 실시예에 따른 실시간 대화 표시 시스템의 구성을 나타내는 도면이다.
- 도 3은 본 발명의 일 실시예에 따른 복수의 화자와 개별 실시간 대화 표시 시스템 간의 동작 과정을 설명하기 위한 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 복수의 화자와 하나의 실시간 대화 표시 시스템의 동작 과정을 설명하기 위한 도면이다.
- 도 5는 본 발명의 일 실시예에 따른 실시간 대화 표시를 위한 출력 화면의 예시를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0017] 한 공간에서 대화가 이루어질 때, 특히 회의와 같은 상황에서는 화자 별로 발화한 내용을 텍스트로 옮기는 것이 필요하다. 하지만 녹음한 음성파일을 듣고 직접 텍스트를 입력하는 것에는 물리적 어려움이 존재하고, 음성 인식(Automatic Speech Recognition; ASR) 기술 등으로 텍스트를 얻어내는 방법은 화자를 분리할 수 없다. 이와 관련하여 음성 파일을 텍스트로 변환해주는 솔루션이 등장하였으나, 이들은 발화자의 정보를 고려하지 않거나 사전에 수집된 발화자에 대한 데이터가 필요하다는 한계가 존재한다.
- [0018] 따라서, 본 발명은 사적인 대화 또는 업무 회의 등 복수의 참가자가 대화를 하는 상황에서 대화 내용을 실시간

으로 텍스트로 변환하고, 발화한 화자를 인식하여 각 화자의 상대적 위치 및 대화 내용을 직관적으로 파악할 수 있도록 표시하는 방법 및 시스템을 제안한다.

- [0019] 본 발명의 실시예에 따르면, 화자의 위치 정보 및 음성인식 기술을 사용하여 실시간으로 대화 내용 및 해당 발화자의 정보가 표시되도록 하여 대화의 효율성을 높이며 나아가서는 고령자, 청각장애인 등이 대화 내용을 직접 듣지 못하는 경우에도 효과적으로 적용이 가능하다. 이하, 본 발명의 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0021] 도 1은 본 발명의 일 실시예에 따른 실시간 대화 표시 방법을 설명하기 위한 흐름도이다.
- [0022] 제안하는 화자 식별부, 화자 위치 입력부, 음성 인식부 및 출력부를 포함하는 실시간 대화 표시 시스템의 실시간 대화 표시 방법은 실시간 대화 표시 시스템을 구비하고 있는 복수의 화자가 대화를 하는 상황에서 화자 식별부를 통해 발화 음성으로부터 개별 화자의 식별 번호를 인식하는 단계(110), 화자 위치 입력부를 통해 상기 개별 화자의 식별 번호에 대한 개별 화자의 위치를 입력 받는 단계(120), 음성 인식부를 통해 발화 음성으로부터 문자 데이터를 추출하는 단계(130) 및 출력부를 통해 개별 화자의 위치와 발화의 내용을 출력하는 단계(140)를 포함한다.
- [0023] 단계(110)에서, 실시간 대화 표시 시스템을 구비하고 있는 복수의 화자가 대화를 하는 상황에서 화자 식별부를 통해 발화 음성으로부터 개별 화자의 식별 번호를 인식한다.
- [0024] 상기 개별 화자에 대한 고유한 식별 번호를 생성하고, 단일 또는 복수의 장치로부터 입력된 음성 신호를 통해 발화 중인 화자에 대응하는 식별 번호를 인식한다.
- [0025] 단계(120)에서, 화자 위치 입력부를 통해 상기 개별 화자의 식별 번호에 대한 개별 화자의 위치를 입력 받는다.
- [0026] 상기 복수의 화자 각각의 실시간 대화 표시 시스템의 출력부를 통해 문자 데이터 또는 발화자를 식별하기 위한 그래픽 객체를 출력할 위치를 입력 받는다.
- [0027] 단계(130)에서, 음성 인식부를 통해 발화 음성으로부터 문자 데이터를 추출한다.
- [0028] 음성인식을 위한 알고리즘을 통해 발화 음성으로부터 발화자가 의도하는 발화 내용을 포함하는 문자 데이터를 추출한다.
- [0029] 단계(140)에서, 출력부를 통해 개별 화자의 위치와 발화의 내용을 출력한다.
- [0030] 상기 화자 식별부를 통해 인식된 개별 화자의 식별 번호와 상기 화자 위치 입력부를 통해 입력 받은 문자 데이터 또는 발화자를 식별하기 위한 그래픽 객체를 출력할 위치, 상기 음성 인식부를 통해 식별한 문자 데이터를 입력 받아 화면 또는 이와 유사한 대상에 출력한다.
- [0032] 도 2는 본 발명의 일 실시예에 따른 실시간 대화 표시 시스템의 구성을 나타내는 도면이다.
- [0033] 제안하는 실시간 대화 표시 시스템은 화자 식별부(210), 화자 위치 입력부(230), 음성 인식부(220) 및 출력부(240)를 포함한다.
- [0034] 본 발명의 실시예에 따른 화자 식별부(210)는 실시간 대화 표시 시스템을 구비하고 있는 복수의 화자가 대화를 하는 상황에서 발화 음성으로부터 개별 화자의 식별 번호를 인식한다.
- [0035] 본 발명의 실시예에 따른 화자 식별부(210)는 상기 개별 화자에 대한 고유한 식별 번호를 생성하고, 단일 또는 복수의 장치로부터 입력된 음성 신호를 통해 발화 중인 화자에 대응하는 식별 번호를 인식한다.
- [0036] 본 발명의 실시예에 따른 화자 위치 입력부(230)는 상기 개별 화자의 식별 번호에 대한 개별 화자의 위치를 입력 받는다.
- [0037] 본 발명의 실시예에 따른 화자 위치 입력부(230)는 상기 복수의 화자 각각의 실시간 대화 표시 시스템의 출력부를 통해 문자 데이터 또는 발화자를 식별하기 위한 그래픽 객체를 출력할 위치를 입력 받는다.
- [0038] 본 발명의 실시예에 따른 음성 인식부(220)는 발화 음성으로부터 문자 데이터를 추출한다.
- [0039] 본 발명의 실시예에 따른 음성 인식부(220)는 음성인식을 위한 알고리즘을 통해 발화 음성으로부터 발화자가 의

도하는 발화 내용을 포함하는 문자 데이터를 추출한다.

- [0040] 본 발명의 실시예에 따른 출력부(240)는 개별 화자의 위치와 발화의 내용을 출력한다.
- [0041] 본 발명의 실시예에 따른 출력부(240)는 상기 화자 식별부를 통해 인식된 개별 화자의 식별 번호와 상기 화자 위치 입력부를 통해 입력 받은 문자 데이터 또는 발화자를 식별하기 위한 그래픽 객체를 출력할 위치, 상기 음성 인식부를 통해 식별한 문자 데이터를 입력 받아 출력한다.
- [0043] 도 3은 본 발명의 일 실시예에 따른 복수의 화자와 개별 실시간 대화 표시 시스템 간의 동작 과정을 설명하기 위한 도면이다.
- [0044] 본 발명의 일 실시예에 따른 개별 화자가 각각의 실시간 대화 표시 시스템을 가지고 있고 각각의 실시간 대화 표시 시스템이 화자의 인근에 위치한 경우의 개별 실시간 대화 표시 시스템 간의 동작 과정을 설명한다.
- [0045] 본 발명의 일 실시예에 따르면, 모든 참여자의 개별 실시간 대화 표시 시스템은 개별 기기 내에서 화자 식별부(311, 321, 331) 및 클라우드 기반의 음성인식 인터페이스 (Speech-to-text Application Programming Interface; STT API) 등을 이용하는 음성 인식부(312, 322, 332)가 존재하고, 사용자 및 타 참여자의 음성 인식부(312, 322, 332)의 결과는 블루투스, 네트워크, NFC 등의 기기간 통신을 통해 사용자 기기(310)의 출력부(314)에 반영된다. 사용자 기기(310)의 위치 입력부(313)는 사용자 기기(310)로부터 개별 화자에 대한 위치를 입력 받아 말풍선의 꼬리를 출력할 방향을 얻고, 사용자 기기(310)의 출력부(314)를 통해 반영된다.
- [0046] 더욱 상세하게는, 실시간 대화 표시 시스템을 구비하고 있는 사용자 기기(310), 참여자 기기1(320) 및 사용자 기기N(330)이 대화에 참여하는 상황에서, 각각의 실시간 대화 표시 시스템의 화자 식별부(311, 321, 331)를 통해 발화된 음성으로부터 화자의 식별 번호를 인식한다. 이때, 각각의 실시간 대화 표시 시스템의 MAC 주소를 이용하여 개별 화자에 대해 식별 번호를 생성하고 각각의 실시간 대화 표시 시스템에 일정 크기 이상으로 입력된 음성 신호에 대해 화자의 식별 번호를 기 생성한 식별 번호로 설정한다.
- [0047] 사용자 기기(310)의 화자 위치 입력부(313)는 개별 화자의 식별 번호에 대해 개별 화자의 위치를 입력 받는다. 이때, 사용자 기기(310)의 출력부(314)를 통해 발화자를 식별하기 위한 말풍선의 꼬리를 출력할 위치를 입력 받는다.
- [0048] 각각의 실시간 대화 표시 시스템의 음성 인식부(312, 322, 332)를 통해 발화 음성으로부터 문자 데이터를 추출한다. 각각의 실시간 대화 표시 시스템에 일정 크기 이상으로 입력된 음성 신호에 한해 ASR 또는 이와 유사한 알고리즘을 통해 발화 음성으로부터 텍스트를 추출한다.
- [0049] 이후, 사용자 기기(310)의 출력부(314)를 통해 화자의 위치와 발화의 내용을 출력한다. 상기 화자의 식별부(311, 321, 331)를 통해 식별한 화자의 식별 번호와 상기 화자 위치 입력부(313)를 통해 입력 받은 말풍선 꼬리의 위치, 상기 음성 인식부(312, 322, 332)를 통해 식별한 텍스트를 입력 받아 화면에 텍스트와 말풍선 꼬리를 포함하는 말풍선을 출력한다.
- [0051] 도 4는 본 발명의 일 실시예에 따른 복수의 화자와 하나의 실시간 대화 표시 시스템의 동작 과정을 설명하기 위한 도면이다.
- [0052] 본 발명의 일 실시예에 따른 개별 화자가 실시간 대화 표시 시스템을 가지고 있고 실시간 대화 표시 시스템이 화자의 인근에 위치한 경우에 대하여 설명한다. 참여자의 음성은 화자 식별부(410)을 통해 분리되고, 클라우드 기반의 음성인식 인터페이스(Speech-to-text Application Programming Interface; STT API) 등을 이용하는 음성 인식부(420)에 각각 입력되며, 그 결과물은 스마트폰 화면(다시 말해, 출력부)(440)을 통해 출력된다. 화자 위치 입력부(430)는 사용자 기기로부터 개별 화자에 대한 위치를 입력 받아 말풍선의 꼬리를 출력할 방향을 얻고, 이는 출력부(440)를 통해 반영된다.
- [0053] 본 발명은 다양한 화자 식별부, 음성인식부, 화면 위치 입력부, 및 출력부를 적용하여 범용적 사용이 가능하며, 예시에 국한되지 않는다.
- [0054] 상기 두 가지 실시 예는 공통적으로, 개별 화자의 스마트폰을 사용하며, MAC 주소를 이용하여 개별 화자에 대해 식별 번호를 생성하고 장치에 일정 크기 이상으로 입력된 음성 신호에 대해 화자의 식별 번호를 기 생성한 식별

번호로 설정한다. 화자 식별부를 통해 얻어진 화자 식별 번호에 따라 말풍선의 색으로 화자를 구별하고, 화자 위치 입력부를 통해 입력 받은 화자의 위치에 기반하여 말풍선 꼬리의 방향으로 화자의 위치를 표시한다. 음성 입력은 스마트폰의 마이크를 통해 전해지며, 음성 인식부에서 폰에서 해당 화자의 음성을 인식한다. 스마트폰 화면에 텍스트를 보이기 위해 채팅방을 형식의 화면을 사용하며, 화자 식별부, 음성인식부, 화자 위치 입력부의 결과가 통합되어 나타난다.

- [0055] 더욱 상세하게는, 실시간 대화 표시 시스템의 화자 식별부(410)를 통해 발화된 음성으로부터 화자의 식별 번호를 인식한다.
- [0056] 실시간 대화 표시 시스템의 난수 생성기를 이용하여 개별 화자에 대해 식별 번호를 생성하고 입력된 음성 신호에 대해 화자 인식 알고리즘 또는 이와 유사한 알고리즘을 통해 각 화자의 식별 번호에 대해 음성 신호를 화자 음성1(411) 및 화자 음성2(412)로 분리한다.
- [0057] 실시간 대화 표시 시스템의 화자 위치 입력부(430)를 통해 개별 화자의 식별 번호에 대해 화자의 위치를 입력 받는다. 개별 화자에 대해 실시간 대화 표시 시스템의 출력부(440)를 통해 발화자를 식별하기 위한 말풍선의 꼬리를 출력할 위치를 입력 받는다.
- [0058] 실시간 대화 표시 시스템의 음성 인식부(420)를 통해 발화 음성으로부터 문자 데이터를 추출한다. 화자 상기 화자 식별부(410)를 통해 분리된 음성 신호에 대해 ASR 또는 이와 유사한 알고리즘을 통해 발화 음성으로부터 텍스트를 추출한다.
- [0059] 이후, 실시간 대화 표시 시스템의 출력부(440)를 통해 화자의 위치와 발화의 내용을 출력한다. 상기 화자 식별부(410)를 통해 식별한 화자의 식별 번호와 상기 입력부를 통해 입력 받은 말풍선 꼬리의 위치, 상기 음성 인식부를 통해 식별한 텍스트를 입력 받아 화면에 텍스트와 말풍선 꼬리를 포함하는 말풍선을 출력한다.
- [0061] 도 5는 본 발명의 일 실시예에 따른 실시간 대화 표시를 위한 출력 화면의 예시를 나타내는 도면이다.
- [0062] 도 5(a) 내지 도 5(g)는 음성입력과 사용자 입력에 따라 스마트폰 화면이 바뀌는 흐름을 나타낸다. 화자 구별을 더욱 확실하게 하기 위해 화자 식별 ID에 대응하는 새로운 ID가 사용될 수 있으며, 발화한 시간 또한 표시될 수 있다.
- [0063] 도 5(a) 및 도 5(d)는 새로운 화자의 음성인식 결과가 전달 되면, 사용자의 입력이 있기 전까지 화자의 위치가 표기 되지 않음을 나타낸다.
- [0064] 도 5(b) 및 도 5(e)는 화자의 입력이 스마트폰에 전달되어 화자의 위치가 지정됨을 보인다.
- [0065] 도 5(c) 및 도 5(f)는 각각 도 5(b) 및 도 5(e)에서 지정된 화자의 위치가 표시된 결과를 나타낸다.
- [0066] 도 5(g)는 지정된 화자의 위치가 지속되는 것을 보인다.
- [0068] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.
- [0069] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로

(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

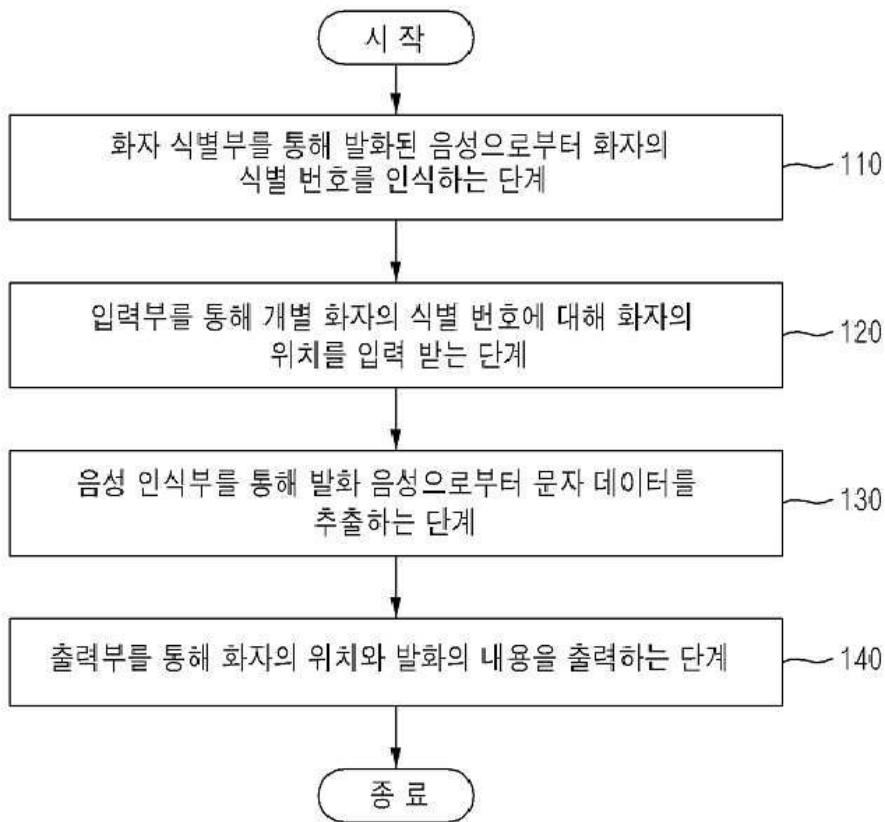
[0070] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic 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.

[0071] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

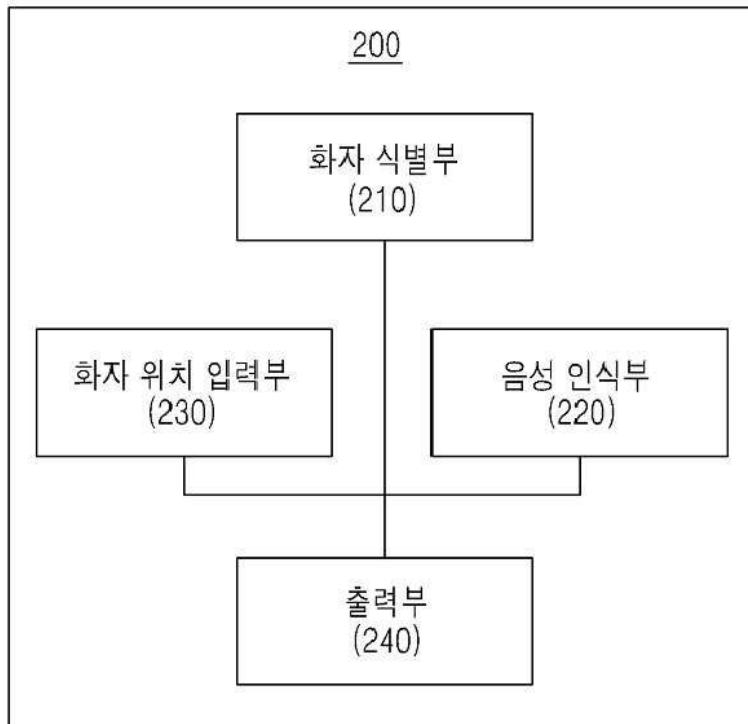
[0072] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

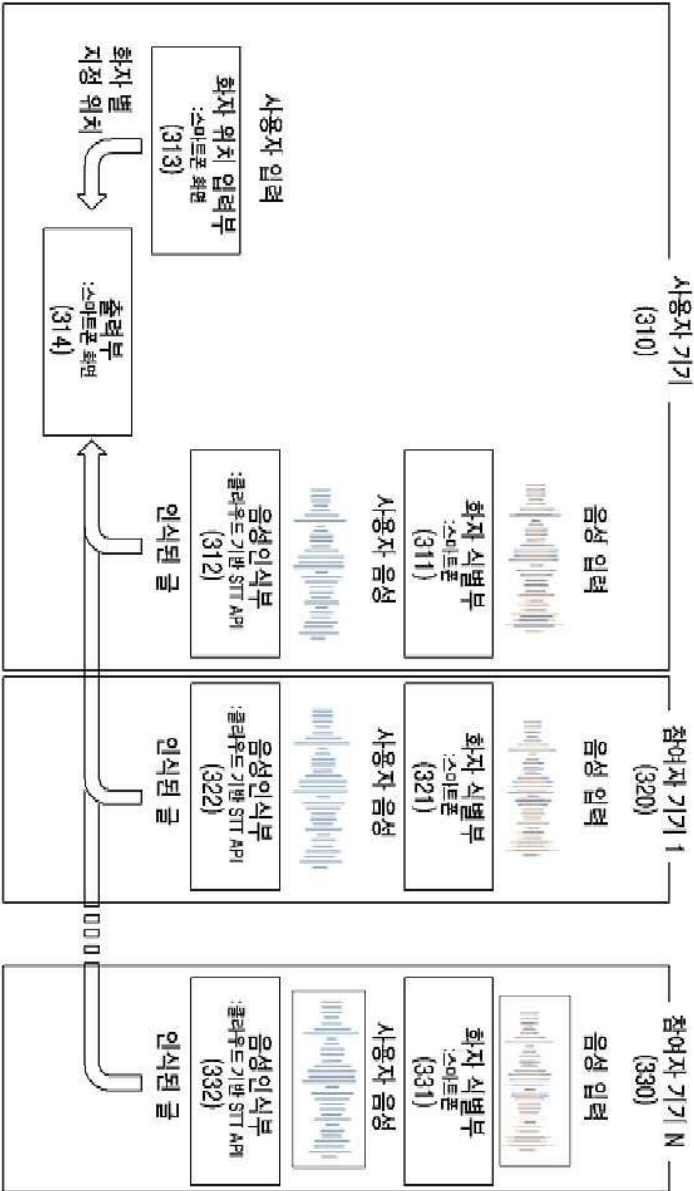
도면1



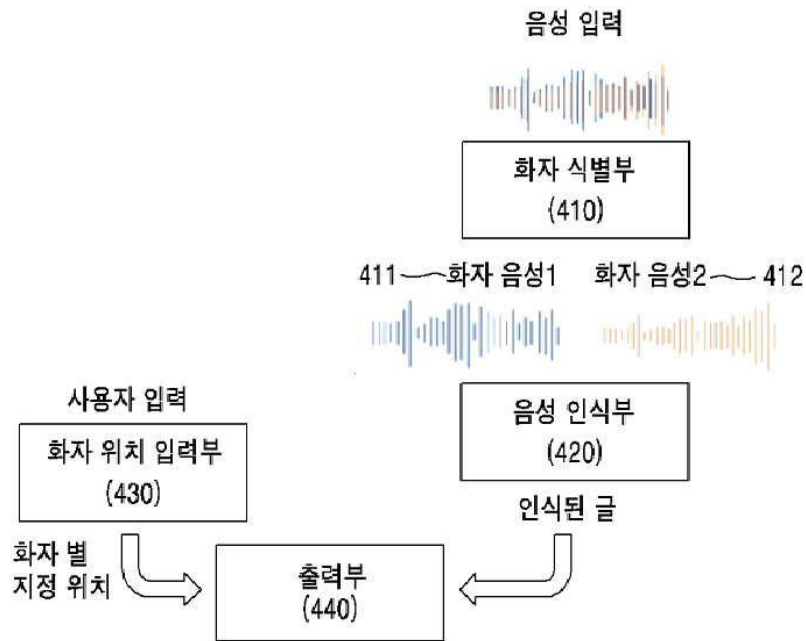
도면2



도면3



도면4



도면5

