



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0039702
(43) 공개일자 2022년03월29일

- (51) 국제특허분류(Int. Cl.)
G06F 3/01 (2006.01) G06T 13/40 (2011.01)
G10L 15/26 (2006.01)
- (52) CPC특허분류
G06F 3/011 (2022.02)
G06T 13/40 (2013.01)
- (21) 출원번호 10-2022-7000276
- (22) 출원일자(국제) 2020년06월04일
심사청구일자 없음
- (85) 번역문제출일자 2022년01월05일
- (86) 국제출원번호 PCT/US2020/036068
- (87) 국제공개번호 WO 2020/247590
국제공개일자 2020년12월10일
- (30) 우선권주장
62/858,234 2019년06월06일 미국(US)

- (71) 출원인
아티, 인코포레이티드
미국, 캘리포니아 90069, 웨스트 할리우드, 피엠티비 33513, 산타 모니카 블러바드 8605
- (72) 발명자
매킨타이어-커윈, 아르만도
미국, 캘리포니아 90026, 로스 엔젤레스, 버클리셔클 3033, 아티 인코포레이티드내
호리간, 라이언
미국, 캘리포니아 90026, 로스 엔젤레스, 버클리셔클 3033, 아티 인코포레이티드내
아이젠버그, 조쉬
미국, 캘리포니아 90026, 로스 엔젤레스, 버클리셔클 3033, 아티 인코포레이티드내
- (74) 대리인
특허법인씨엔에스

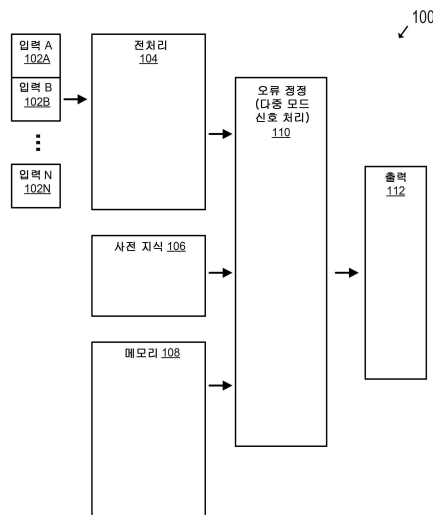
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **동적으로 반응하는 가상 캐릭터를 위한 다중 모드 모델**

(57) 요약

개시된 실시예는 다중 모드 모델을 이용하여 가상 캐릭터(또는 "아바타(avatar)")를 제어하는 방법에 관한 것이다. 다중 모드 모델은 사용자와 관련된 다양한 입력 정보를 처리하고, 다수의 내부 모델을 이용하여 입력 정보를 처리할 수 있다. 다중 모드 모델은 내부 모델을 조합하여 가상 캐릭터에 의한 믿을 수 있고 감정적으로 매력적인 반응을 만들 수 있다. 가상 캐릭터에 대한 링크는 웹 브라우저에 임베딩될 수 있고 아바타는 사용자에게 의해 가상 캐릭터와 상호 작용하기 위한 선택에 기초하여 동적으로 생성될 수 있다. 보고서가 클라이언트를 위해 생성될 수 있으며, 보고서는 클라이언트와 연관된 가상 캐릭터와 상호 작용하는 사용자의 특성에 대한 인사이트(insight)를 제공한다.

대표도 - 도1



(52) CPC특허분류

G10L 15/26 (2013.01)

G06F 2203/011 (2013.01)

G06F 2203/0381 (2013.01)

G06T 2200/24 (2013.01)

G06T 2213/08 (2013.01)

명세서

청구범위

청구항 1

가상 캐릭터를 제어하는 방법에 있어서,

장치로부터 음성 정보, 얼굴 표정 정보 및 상기 장치를 둘러싸는 환경을 나타내는 환경 정보 중 임의의 하나를 포함하는 다중 모드 입력 정보를 수신하는 단계;

상기 장치 상에 제시된 디스플레이 환경 내의 위치에 상기 가상 캐릭터를 디스플레이하는 단계;

적어도 2개의 내부 모델을 구현하여 상기 다중 모드 입력 정보의 특성을 식별하는 단계;

상기 식별된 특성 중 제1 식별된 특성이 상기 식별된 특성 중 제2 식별된 특성의 유사한 특징을 임계 개수 포함하는지 여부를 결정하기 위해 상기 적어도 2개의 내부 모델의 상기 식별된 특성을 검사하는 단계;

상기 제1 식별된 특성 및 상기 제2 식별된 특성을 가상 캐릭터 지식 모델에 포함된 상기 가상 캐릭터에 특정한 정보에 대하여 비교하여, 상기 제1 식별된 특성이 상기 식별된 특성 중 상기 제2 식별된 특성의 유사한 특징을 상기 임계 개수 포함한다고 판단한 것에 기초하여 선택 특성을 선택하는 단계;

상기 가상 캐릭터와 연관된 잠재적 동작의 라이브러리에 액세스하여 상기 선택 특성과 일치하는 동작을 결정하는 단계 - 상기 동작은 상기 가상 캐릭터에 의해 수행될 애니메이션 및 연관 오디오 모두를 포함함 -; 및

상기 장치에 제시된 상기 환경에서 상기 가상 캐릭터를 수정하고 상기 연관 오디오를 출력함으로써 상기 결정된 동작을 구현하는 단계

를 포함하는, 방법.

청구항 2

제1항에 있어서,

상기 적어도 2개의 내부 모델은, 상기 음성 정보로부터 음성 감정을 파싱할 수 있는 음성 인식 모델 및 상기 얼굴 표정 정보에 기초하여 얼굴 특징 감정을 검출할 수 있는 얼굴 특징 인식 모델을 포함하고, 상기 선택 특성은 상기 음성 감정과 상기 얼굴 특징 감정 사이에 공통적인 감정이고, 상기 결정된 동작은 상기 감정에 기초하여 결정되는, 방법.

청구항 3

제1항에 있어서,

상기 적어도 2개의 내부 모델은 사용자와 이전에 맺은 관계에 관한 정보를 포함하는 사전 지식 정보를 검색할 수 있는 사전 지식 모델을 포함하고, 상기 선택 특성은 상기 사전 지식 모델을 사용하여 처리된 상기 사전 지식 정보에 기초하여 선택되는, 방법.

청구항 4

제1항에 있어서,

상기 내부 모델은 오디오 정보로부터 컨텍스트와 의미를 도출하도록 구성된 자연어 이해 모델, 환경 정보를 식별하도록 구성된 인식 모델 및 사용자 및 다른 가상 캐릭터와 관련된 데이터를 식별하도록 구성된 소셜 시뮬레이션 모델을 포함하는, 방법.

청구항 5

제1항에 있어서,

상기 가상 캐릭터에게 상기 장치에서 사용자에게 대한 질의를 나타내는 초기 동작을 수행하도록 지시하는 단계를

더 포함하고, 상기 입력 정보는 상기 질의에 대한 상기 사용자에게 의한 응답을 나타내는, 방법.

청구항 6

제1항에 있어서,

네트워크를 통해 복수의 사용자에게 임베디드 링크를 공유하는 단계;

상기 임베디드 링크가 선택되었다는 것을 나타내는 선택을 장치 세트 중 임의의 장치로부터 수신하는 단계; 및

상기 선택을 수신하는 것에 응답하여, 상기 장치에 상기 가상 캐릭터를 디스플레이하기 위해 상기 선택을 송신한 상기 장치 세트 중의 상기 장치에 데이터 스트림을 송신하는 단계

를 포함하는, 방법.

청구항 7

제6항에 있어서,

상기 데이터 스트림의 제1 배치(batch)를 제1 시간에 전송하는 단계 - 상기 제1 배치는 상기 장치의 디스플레이 상에 상기 가상 캐릭터를 초기에 생성하기 위한 정보를 포함함 -; 및

상기 제1 시간 이후의 제2 시간에 상기 데이터 스트림의 제2 배치를 전송하는 단계 - 상기 제2 배치는 상기 가상 캐릭터에 의해 제1 동작을 출력하기 위한 정보를 포함함 -

를 더 포함하고, 상기 제1 배치는 상기 제2 시간에 폐기되는, 방법.

청구항 8

제1항에 있어서,

상기 환경의 바닥(floor)을 나타내는 상기 환경의 부분을 식별하기 위하여 환경 정보를 검사하는 단계; 및

상기 환경의 바닥을 나타내는 상기 환경의 부분 위의 제1 위치에 상기 가상 캐릭터를 위치 설정하는 단계

를 더 포함하는, 방법.

청구항 9

장치에 있어서,

상기 장치는 상기 장치에 의해 캡처된 사용자에게 관한 다중 모드 입력에 대한 응답을 제공하도록 구성되고, 상기 장치는,

적어도 하나의 메모리로서,

다중 모드 입력 정보로부터 특성을 식별하도록 구성된 적어도 2개의 내부 모델;

가상 캐릭터에 특정한 정보를 포함하는 가상 캐릭터 지식 모델; 그리고

상기 가상 캐릭터와 연관된 잠재적 동작의 라이브러리 - 각각의 동작은 상기 가상 캐릭터에 의해 수행될 애니메이션 및 연관 오디오와 연관됨 -

를 포함하는 상기 적어도 하나의 메모리;

적어도 하나의 프로세서로서,

음성 정보, 얼굴 표정 정보 및 환경을 나타내는 환경 정보 중 적어도 하나를 포함하는 다중 모드 입력 정보를 수신하고;

제1 식별된 특성이 제2 식별된 특성과 임계 유사도 내에 있는지 여부를 결정하기 위해 상기 적어도 2개의 내부 모델에 의해 식별된 특성을 검사하고;

상기 가상 캐릭터 지식 모델에 대해 제1 식별된 특성 및 제2 식별된 특성을 비교하여 선택 특성을 식별하고;

상기 가상 캐릭터와 연관된 잠재적인 동작의 라이브러리를 검사함으로써 상기 선택 특성과 일치하는 동작을 결

정하고 - 상기 동작은 상기 장치에 출력될 오디오를 포함함 -; 그리고
 상기 장치에 상기 오디오를 출력
 하도록 구성된 하나 이상의 프로세서
 를 포함하는, 장치.

청구항 10

제9항에 있어서,
 상기 적어도 하나의 프로세서는,
 상기 환경 정보로부터 도출된 환경 내의 위치에 상기 장치의 디스플레이 상에 상기 가상 캐릭터를 디스플레이하
 고; 그리고
 상기 장치에 제시되는 상기 환경 내에 상기 가상 캐릭터를 수정함으로써 상기 장치에 출력될 상기 오디오와 상
 기 가상 캐릭터에 의해 수행될 선택된 애니메이션을 모두 포함하는 상기 동작을 구현
 하도록 더 구성되는, 장치.

청구항 11

제9항에 있어서,
 상기 적어도 2개의 내부 모델은 상기 사용자와 이전에 맺은 관계에 관한 정보를 포함하는 사전 지식 정보를 검
 색할 수 있는 사전 지식 모델을 포함하고, 상기 선택 특성은 상기 사전 지식 모델을 사용하여 처리된 상기 사전
 지식 정보에 기초하여 선택되는, 장치.

청구항 12

제9항에 있어서,
 상기 적어도 2개의 내부 모델은, 상기 음성 정보로부터 음성 감정을 파싱할 수 있는 음성 인식 모델 및 상기 얼
 굴 표정 정보에 기초하여 얼굴 특징 감정을 검출할 수 있는 얼굴 특징 인식 모델을 포함하고, 상기 선택 특성은
 상기 음성 감정과 상기 얼굴 특징 감정 사이에 공통적인 감정이고, 상기 결정된 동작은 상기 감정에 기초하여
 결정되는, 장치.

청구항 13

사용자 장치의 웹 브라우저에서 가상 캐릭터를 동적으로 생성하기 위한 컴퓨터 구현 방법에 있어서,
 상기 사용자 장치의 상기 웹 브라우저에 대한 링크를 임베딩하는 단계 - 상기 링크는 상기 사용자 장치에서 실
 행되는 애플리케이션으로 상기 웹 브라우저를 연결함 -;
 상기 링크가 선택되었다는 표시를 상기 사용자 장치로부터 수신하는 단계;
 상기 가상 캐릭터를 생성하기 위하여 정보를 나타내는 데이터 스트림을 상기 애플리케이션으로부터 상기 웹 브
 라우저로 전송하는 단계;
 상기 가상 캐릭터를 상기 사용자 장치의 상기 웹 브라우저에 디스플레이하는 단계;
 음성 정보, 얼굴 표정 정보 및 환경을 나타내는 환경 정보를 포함하는 다중 모드 입력 정보를 상기 장치로부터
 수신하는 단계;
 적어도 2개의 내부 모델을 구현하여 상기 다중 모드 입력 정보의 특성을 식별하는 단계;
 상기 식별된 특성 중 제1 식별된 특성이 상기 식별된 특성 중 제2 식별된 특성의 임계 유사도 내에 있는지 여부
 를 결정하기 위해 상기 적어도 2개의 내부 모델에 의해 식별된 특성을 검사하는 단계;
 상기 제1 식별된 특성 및 상기 제2 식별된 특성을 가상 캐릭터 지식 모델에 포함된 상기 가상 캐릭터에 특정한
 정보에 대하여 비교하여, 상기 제1 식별된 특성이 상기 식별된 특성 중 상기 제2 식별된 특성의 유사한 특징을
 상기 임계 개수 포함한다고 판단한 것에 기초하여 선택 특성을 선택하는 단계;

상기 가상 캐릭터와 연관된 잠재적 동작의 라이브러리에 액세스하여 선택 특성과 일치하는 동작을 결정하는 단계 - 상기 동작은 상기 가상 캐릭터에 의해 수행될 애니메이션 및 연관 오디오 모두를 포함함 -; 및
 상기 동작을 수행하는 상기 환경에서 상기 가상 캐릭터를 디스플레이하고 상기 연관 오디오를 출력하는 단계를 포함하는, 컴퓨터 구현 방법.

청구항 14

제13항에 있어서,

상기 웹 브라우저는 상기 사용자 장치에서 실행되는 모바일 애플리케이션에 디스플레이되는 페이지를 포함하는, 컴퓨터 구현 방법.

청구항 15

제13항에 있어서,

상기 적어도 2개의 내부 모델은, 상기 음성 정보로부터 음성 감정을 파싱할 수 있는 음성 인식 모델 및 상기 얼굴 표정 정보에 기초하여 얼굴 특징 감정을 검출할 수 있는 얼굴 특징 인식 모델을 포함하고, 상기 선택 특성은 상기 음성 감정과 상기 얼굴 특징 감정 사이에 공통적인 감정이고, 상기 결정된 동작은 상기 감정에 기초하여 결정되는, 컴퓨터 구현 방법.

청구항 16

제13항에 있어서,

상기 적어도 2개의 내부 모델은 사용자와 이전에 맺은 관계에 관한 정보를 포함하는 사전 지식 정보를 검색할 수 있는 사전 지식 모델을 포함하고, 상기 선택 특성은 상기 사전 지식 모델을 사용하여 처리된 상기 사전 지식 정보에 기초하여 선택되는, 컴퓨터 구현 방법.

청구항 17

제13항에 있어서,

네트워크를 통해 복수의 사용자에게 임베디드 링크를 공유하는 단계;

상기 링크가 선택되었다는 것을 나타내는 선택을 장치 세트 중 임의의 장치로부터 수신하는 단계; 및

상기 선택을 수신하는 것에 응답하여, 상기 장치 상에 상기 가상 캐릭터를 디스플레이하기 위해 상기 선택을 송신한 상기 장치 세트의 중의 상기 사용자 장치에 데이터 스트림을 송신하는 단계

를 포함하는, 컴퓨터 구현 방법.

청구항 18

제13항에 있어서,

상기 데이터 스트림의 제1 배치(batch)를 제1 시간에 전송하는 단계 - 상기 제1 배치는 상기 장치의 디스플레이 상에 상기 가상 캐릭터를 초기에 생성하기 위한 정보를 포함함 -; 및

상기 제1 시간 이후의 제2 시간에 상기 데이터 스트림의 제2 배치를 전송하는 단계 - 상기 제2 배치는 상기 가상 캐릭터에 의해 제1 동작을 출력하기 위한 정보를 포함함 -

를 더 포함하고, 상기 제1 배치는 상기 제2 시간에 폐기되는, 컴퓨터 구현 방법.

청구항 19

제13항에 있어서,

상기 환경의 바닥(floor)을 나타내는 상기 환경의 부분을 식별하기 위하여 환경 정보를 검사하는 단계; 및

상기 환경의 바닥을 나타내는 상기 환경의 부분 위의 제1 위치에 상기 가상 캐릭터를 위치 설정하는 단계

를 더 포함하는, 컴퓨터 구현 방법.

청구항 20

제13항에 있어서,

상기 선택 특성 및 상기 동작에 관한 정보를 저장하는 단계;

복수의 사용자에게 대한 일련의 선택 특성 및 동작을 집계하는 단계;

상기 복수의 사용자에게 대한 상기 일련의 선택 특성 및 동작을 처리하여 상기 복수의 사용자와의 상기 가상 캐릭터의 관계(engagement)에 관한 분석 세트를 도출하는 단계; 및

상기 복수의 사용자와의 상기 가상 캐릭터의 상기 관계에 관한 상기 분석 세트를 디스플레이하기 위해 분석 대시보드를 제시하는 단계

를 포함하는, 컴퓨터 구현 방법.

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호 참조

[0002] 본 실시예는 전체 내용이 여기에 참조로 포함되는 2019년 6월 6일에 출원된 "MULTI-MODAL MODEL FOR DYNAMICALLY RESPONSIVE AVATARS(동적으로 반응하는 가상 캐릭터를 위한 다중 모드 모델)"이라는 제목의 미국 임시 특허 출원 제62/858,234호에 대한 우선권을 주장한다.

[0003] 기술분야

[0004] 개시된 교시 내용은 일반적으로 장치 상의 가상 캐릭터의 제어에 관한 것이다. 개시된 교시 내용은, 더욱 구체적으로, 다중 모드 모델을 사용하여 장치에서 동적으로 응답하는 가상 캐릭터를 제어하는 것에 관한 것이다.

배경 기술

[0005] 가상 캐릭터(또는 "아바타")는 사용자 장치(예를 들어, 스마트폰, 컴퓨터, 증강 현실 장치) 상에서 사용자와의 상호 작용을 용이하게 할 수 있다. 가상 캐릭터는 사용자 장치의 디스플레이에 표시된 환경에 묘사된 캐릭터의 가상 표현을 포함할 수 있다. 가상 캐릭터는 시각적 형태를 가질 필요가 없다. 이는 단지 사용자 장치를 통해 사용자와 의사 소통할 수 있으면 된다. 사용자로부터의 입력은 가상 캐릭터가 취할 동작을 결정하기 위해 식별되고 검사될 수 있다. 가상 캐릭터는 결정된 동작(예를 들어, 애니메이션, 대화 수행)을 취할 수 있어, 가상 캐릭터와 사용자 사이의 지속적인 상호 작용을 용이하게 할 수 있다.

[0006] 가상 캐릭터를 생성하기 위한 중요한 고려 사항은 사용자에게 의해 수신된 입력에 응답하여 가상 캐릭터에 의해 취해지는 동작의 정확성이다. 일반적으로, 입력에 대한 더욱 정확한 응답은 가상 캐릭터와 상호 작용하는 데 있어서의 사용자 경험을 증가시킬 수 있다. 반대로, 가상 캐릭터가 사용자에게 의해 제공된 입력을 잘못 해석하면, 가상 캐릭터는 입력에 대해 부정확한 응답을 제공할 수 있어, 가상 캐릭터와의 사용자 경험을 저하시킬 수 있다.

도면의 간단한 설명

[0007] 도 1은, 다양한 실시예에 따라, 가상 캐릭터에 대한 출력을 제어하기 위한 다중 모드 모델의 순서도를 도시한다.

도 2는, 다양한 실시예에 따라, 다중 모드를 사용하여 가상 캐릭터의 오류를 정정하기 위한 순서도를 도시한다.

도 3은, 다양한 실시예에 따라, 가상 캐릭터를 제어하기 위한 다중 모드 모델의 순서도를 도시한다.

도 4는, 다양한 실시예에 따라, 다중 모드 모델에 구현된 가상 캐릭터와 상호 작용하는 시스템을 도시한다.

도 5는, 다양한 실시예에 따라, 장치 상에 디스플레이되는 예시적인 가상 캐릭터를 도시한다.

도 6은, 다양한 실시예에 따라, 사용자와, 가상 캐릭터를 디스플레이하는 사용자 장치 사이의 상호 작용의 순서

도를 예시한다.

도 7은, 다양한 실시예에 따라, 가상 캐릭터 반응으로부터 인사이트를 추출하는 순서도를 도시한다.

도 8은, 다양한 실시예에 따라, 인사이트 대시보드를 도시한다.

도 9는, 다양한 실시예에 따라, 다중 모드 모델을 사용하여 가상 캐릭터를 제어하는 방법의 블록도를 도시한다.

도 10은 여기에 설명된 적어도 일부 동작들이 구현될 수 있는 프로세싱 시스템의 일례를 예시하는 블록도이다.

도면 및 표는 다양한 실시예를 단지 예시를 위해 도시한다. 본 발명에 속하는 기술 분야에서 통상의 기술을 가진 자는 기술의 원리를 벗어나지 않고 대안적인 실시예가 채용될 수 있다는 것을 인식할 것이다. 따라서, 특정 실시예가 도면에 도시되어 있지만, 기술은 다양한 수정이 가능하다.

발명을 실시하기 위한 구체적인 내용

- [0008] 아래에 설명되는 실시예는 본 발명에 속하는 기술 분야에서 통상의 기술을 가진 자가 실시예를 수행할 수 있게 하는데 필요한 정보를 나타내고, 실시예들을 실시하는 최선의 형태를 예시한다. 첨부된 도면에 비추어 다음 설명을 읽을 때, 본 발명에 속하는 기술 분야에서 통상의 기술을 가진 자는 본 개시 내용의 개념을 이해할 것이고, 여기에서 특별히 다루어 지지 않은 이러한 개념의 적용을 인식할 것이다. 이러한 개념 및 적용은 본 개시 내용 및 첨부된 청구범위의 범위에 속한다.
- [0009] 실시예는 특정 컴퓨터 프로그램, 시스템 구성, 네트워크 등을 참조하여 설명될 수 있다. 그러나, 본 발명에 속하는 기술 분야에서 통상의 기술을 가진 자는 이러한 특징이 다른 컴퓨터 프로그램 유형, 시스템 구성, 네트워크 유형 등에 동일하게 적용 가능하다는 것을 인식할 것이다. 예를 들어, "Wi-Fi 네트워크"라는 용어가 네트워크를 설명하는 데 사용될 수 있지만, 관련 실시예는 다른 유형의 네트워크에 사용될 수 있다.
- [0010] 더욱이, 개시된 기술은 특수 목적 하드웨어(예를 들어, 회로), 소프트웨어 및/또는 펌웨어로 적절하게 프로그래밍된 프로그래밍 가능한 회로 또는 특수 목적 하드웨어와 프로그래밍 가능한 회로의 조합을 사용하여 구체화될 수 있다. 따라서, 실시예는 컴퓨팅 장치(예를 들어, 기지국 또는 네트워크에 연결된 컴퓨터 서버)를 프로그래밍하여 전자 장치에 의해 생성된 비디오 콘텐츠를 검사하고, 비디오 콘텐츠에 포함된 요소를 식별하고, 분류 모델을 적용하여 적합한 동작을 결정하고, 적합한 동작을 수행하는 데 사용될 수 있는 명령어를 갖는 기계 판독 가능한 매체를 포함할 수 있다.
- [0011] 본 개시 내용에서, "아바타(avatar)"라는 용어는 "가상 캐릭터(virtual character)"와 혼용될 수 있다. 가상 캐릭터는 사용자와 의사 소통할 수 있는 사용자 장치의 가상 엔티티를 포함할 수 있다. 이러한 의사 소통은 음성(대화), 텍스트(예를 들어, 채팅 또는 메시징 애플리케이션), 이미지/비디오, 음악 및/또는 임의의 프로그램의 실행(예를 들어, 조정 권한을 갖는 사물 인터넷 임베디드 시스템에서 전등 스위치를 켜는 가상 캐릭터) 등과 같은 상이한 유형의 상호 작용을 통해 발생할 수 있다. 가상 캐릭터는 애니메이션화된 바디를 가지고 있는 것처럼 보일 수 있거나, 단지 음성 또는 텍스트 기반 에이전트일 수 있지만, 예를 들어, 애니메이션화된 바디를 가지는 것이 필요하지 않을 수 있다.
- [0012] 용어
- [0013] 여기에서 사용된 용어의 목적은 단지 실시예를 설명하기 위한 것으로, 본 개시 내용의 범위를 한정하려고 의도되지 않는다. 문맥이 허용하는 경우, 단수 또는 복수 형태를 사용하는 단어는 또한 복수 또는 단수 형태를 각각 포함할 수도 있다.
- [0014] 여기에 사용된 바와 같이, 달리 구체적으로 언급되지 않는 한, "처리하는(processing)", "컴퓨팅하는(computing)", "계산하는(calculating)", "결정하는(determining)", "디스플레이하는(displaying)", "생성하는(generating)" 등과 같은 용어는 컴퓨터의 메모리 또는 레지스터 내에서 물리적(전자적) 수량으로 표시된 데이터를 조작하여 컴퓨터의 메모리, 레지스터 또는 기타 저장 매체, 전송 또는 디스플레이 장치 내에서 물리적 수량으로 유사하게 표시되는 다른 데이터로 변환하는 컴퓨터 또는 유사한 전자 컴퓨팅 장치의 동작 및 프로세스를 의미한다.
- [0015] 여기에 사용된 바와 같이, "연결된(connected)", "결합된(coupled)" 등과 같은 용어는 2개 이상의 요소 사이의 직접적 또는 간접적인 임의의 연결 또는 결합을 의미할 수 있다. 요소 사이의 결합 또는 연결은 물리적이거나, 논리적이거나 또는 이들의 조합일 수 있다.

- [0016] "일 실시예" 또는 "하나의 실시예"에 대한 언급은 설명되는 특정 특징, 기능, 구조 또는 특성이 적어도 하나의 실시예에 포함된다는 것을 의미한다. 이러한 문구의 발생은 반드시 동일한 실시예를 지칭하는 것은 아니며, 반드시 상호 배타적인 대안적인 실시예를 지칭하는 것도 아니다.
- [0017] 문맥상 명백히 달리 요구하지 않는 한, "포함하다(comprise)" 및 "포함하는(comprising)"이라는 단어는 배타적이거나 소진적인 의미보다는 포괄적인 의미(즉, "포함하지만 이에 제한되지 않는" 의미)로 해석되어야 한다.
- [0018] 또한, "~에 기초한(based on)"이라는 용어도 배타적이거나 소진적인 의미가 아니라 포괄적인 의미로 해석되어야 한다. 따라서, 달리 언급되지 않는 한, "~에 기초한"이라는 용어는 "적어도 부분적으로 ~에 기초한"을 의미하도록 의도된다.
- [0019] "모듈(module)"이라는 용어는 소프트웨어 컴포넌트, 하드웨어 컴포넌트 및/또는 펌웨어 컴포넌트를 광범위하게 의미한다. 모듈은 통상적으로 특정된 입력(들)에 기초하여 유용한 데이터 또는 기타 출력(들)을 생성할 수 있는 기능적 컴포넌트이다. 모듈은 독립적(self-contained)일 수 있다. 컴퓨터 프로그램은 하나 이상의 모듈을 포함할 수 있다. 따라서, 컴퓨터 프로그램은 서로 다른 작업을 완료하는 책임을 갖는 다수의 모듈 또는 다수의 작업을 완료하는 책임을 갖는 단일 모듈을 포함할 수 있다.
- [0020] 다수의 항목의 목록과 관련하여 사용될 때, "또는(or)"이라는 단어는 다음과 같은 해석을 모두 포함하도록 의도된다: 목록 내의 임의의 항목, 목록 내의 모든 항목 및 목록 내의 항목의 임의의 조합.
- [0021] 여기에 설명된 임의의 프로세스에서 수행되는 단계들의 순서는 예시적이다. 그러나, 물리적 가능성에 반하지 않는 한, 단계들은 다양한 순서 및 조합으로 수행될 수 있다. 예를 들어, 단계들은 여기에 설명된 프로세스에 추가되거나 그로부터 제거될 수 있다. 마찬가지로, 단계들은 교체되거나 재정렬될 수 있다. 따라서, 임의의 프로세스에 대한 설명은 오픈 엔드(open-ended)가 되도록 의도된다.
- [0022] 시스템 개요
- [0023] 가상 캐릭터(또는 "아바타")는 사용자 장치를 통해 사용자와의 의사 소통을 용이하게 할 수 있다. 가상 캐릭터는 캐릭터(예를 들어, 책, 영화, 텔레비전으로부터의 캐릭터)의 가상 표현, 유명인의 표현, 인간의 표현, 동물의 표현, 가상의 생물의 표현, 및 현실 세계에서 통상적으로 무생물이지만 상호 작용의 목적을 위해 (말하는 책과 같은) 인간과 의사 소통할 수 있는 물체의 표현일 수 있다. 또한, 가상 캐릭터는 유행, 영혼 또는 감정과 같이 물리적 형태가 없는 엔티티 또는 활동성(animacy)을 나타낼 수 있다. 예를 들어, 사용자는 사용자와 연관된 스마트 폰을 통해 슈퍼히어로를 나타내는 가상 캐릭터와 상호 작용할 수 있다. 가상 캐릭터를 생성하는 시스템의 관점에서, 시스템은 사용자로부터 입력을 수신하고 입력에 기초하여 응답을 결정할 수 있다.
- [0024] 가상 캐릭터를 생성하기 위한 중요한 고려 사항은 사용자에게 의해 수신된 입력에 응답하여 가상 캐릭터에 의해 취해지는 동작의 정확성이다. 일반적으로, 입력에 대한 더욱 정확한 응답은 가상 캐릭터와 상호 작용하는 사용자 경험을 증가시킬 수 있다. 반대로, 가상 캐릭터가 사용자에게 의해 제공된 입력을 잘못 해석하면, 가상 캐릭터는 입력에 대해 부정확한 응답을 제공할 수 있어, 가상 캐릭터와의 사용자 경험을 저하시킬 수 있다.
- [0025] 많은 경우에, 가상 캐릭터에 제공되는 입력은 분리되어 별도로 처리된다. 예를 들어, 가상 캐릭터는 사용자의 음성 데이터 및 얼굴 특징 정보를 수신할 수 있다. 그러나, 정보는 가상 캐릭터에 대해 출력할 애니메이션/음성을 식별하는 것과 관련하여 분리되어 별도로 처리될 수 있으며, 이는 가상 캐릭터의 애니메이션/음성의 정확도를 저하시킬 수 있다.
- [0026] 가상 문자 제어 및 오류 정정을 위한 다중 모드 모델
- [0027] 본 실시예는 캐릭터의 동작을 자동으로 제어하고 가상 캐릭터에 의해 수행되는 정확한 동작을 제공하기 위해 다중 모드 모델을 사용하여 오류를 정정하는 프로그램에 적어도 부분적으로 관련될 수 있다. 다중 모드 모델은 사용자와 관련된 정보의 다수의 모드(예를 들어, 음성, 얼굴 표정, 환경 정보, 사용자와의 이전 상호 작용으로부터 수집된 정보, 시스템이 사용자에게 대해 알고 있는 정보)를 식별하고, 사용자의 입력에 응답하여 가상 캐릭터에 대한 정확한 애니메이션/동작을 동적으로 생성하기 위하여 정보를 다양한 내부 모델과 조합하는 모델을 포함할 수 있다.
- [0028] 다중 모드 모델은 다중 모드 센서의 어레이에 의해 생성된 데이터 신호로부터 추출된 정보에 대한 오류 정정을 수행할 수 있다. 다시 말하면, 현실 세계 환경으로부터 판독값을 수집하는 임의의 다중 모드 디지털 센서 어레이와 각각의 유형의 데이터 신호에 맞게 조정된 임의의 세트의 전처리 및 정보 추출 프로세스가 각각의 센서로부터 추출된 정보를 사용하여 다른 데이터 신호로부터 추출된 정보에서 오류를 정정할 수 있다. 상호 작용의 시

간에 센서에 의해 수집된 정보에 더하여, 다중 모드 오류 정정은 또한 과거로부터 사용자에게 관하여 수집된 정보를 사용할 수 있다. 예를 들어, 다중 모드 오류 정정(multi-modal error correction(MMEC))은 사용자가 과거에 캐릭터에게 말한 내용의 사본(transcript)을 사용하여 오류 정정에 연료를 공급할 수 있다. 또한, MMEC는 사용자에게 대한 정보를 포함하는 지식 베이스(온톨로지(ontology))에 저장된 정보를 이용하여 오류 정정을 안내할 수 있다. 사용자의 과거에 대한 검증된 정보는 현재 수집되고 있는 노이즈 신호를 정정하는 데 사용될 수 있다. 지식 베이스에서의 정보는 컴퓨터에 의해 자동으로 추출된 것일 수도 있고 사람에게 의해 수동으로 입력된 더욱 구조화된 정보일 수도 있다.

- [0029] 도 1은, 다양한 실시예에 따라, 가상 캐릭터에 대한 출력을 제어하기 위한 다중 모드 모델의 순서도를 도시한다. 다중 모드 모델은 음성 인식 기술, 자연어 이해, 컴퓨터 비전(예를 들어, 사용자의 얼굴 표정 인식, 사용자의 물리적 환경 및 기타 물리적 객체에 대한 "세계 인식" 데이터), 지식 베이스(온톨로지), 특정된 가상 캐릭터의 페르소나 및 통합 다중 모드 모델로의 소셜 시뮬레이션을 조합할 수 있다. 다중 모드 모델은 정확한 가상 캐릭터를 생성하는 믿을 수 있고 감정적으로 매력적인 인공 지능 알고리즘을 만들기 위해 다수의 내부 모델을 함께 사용할 수 있다.
- [0030] 도 1에 도시된 바와 같이, 시스템은 센서 어레이로부터 원시 입력(입력(102A-N))을 수신할 수 있다. 시스템은 단일 모드 신호 처리를 수행하는 전처리(104)를 포함할 수 있다. 시스템은 지식 베이스 및 모델로 인코딩된 세계에 대한 정보를 포함하는 사전 지식(106) 및 사용자와의 이전 경험에 대해 인코딩된 정보를 저장하는 메모리(108)를 포함할 수 있다. 시스템은 오류 정정을 수행할 수 있는 다중 모드 신호 처리(110)를 포함할 수 있다. 시스템은 다중 모드 신호 처리에 기초하여 증강 결과(112)를 출력할 수 있다.
- [0031] 일부 실시예에서, 컴퓨터 시스템은 오디오 신호를 생성하는 마이크, 비디오 스트림 신호를 생성하는 카메라, 온도 신호를 생성할 수 있는 열 저항기, 어떤 냄새가 있는지에 대한 신호를 생성하는 후각 센서, GPS 센서, 가속도계 등과 같은 많은 여러 가지 센서를 포함할 수 있다. 시스템은 이러한 센서로부터 수신하여 현실 세계에서 일어나는 일을 정량화하고 이러한 신호의 디지털 표현을 사용하여 서로의 내용을 정정할 수 있다.
- [0032] 사용자와 관련하여 수신된 입력은 다중 모드 모델에 포함된 내부 모델과 결합될 수 있다. 위에서 언급된 바와 같이, 모델은 자연어 이해, 지식 모델, 페르소나, 사용자의 생활/역사에 대한 구조적 지식 및 소셜 시뮬레이션을 포함할 수 있으며, 이는 함께 가상 캐릭터와 상호 작용할 때 정확성 및 사용자 경험을 증가시킬 수 있다.
- [0033] 일례로서, 입력은 "기쁘다(I am glad)"라고 말하는 사용자의 음성을 포함할 수 있고, 여기서 내부 모델은 "기쁘다(I am glad)"라는 단어를 식별하기 위해 자연어 이해를 활용할 수 있다. 이 예에서, 제2 입력은 사용자의 표정을 식별하는 얼굴 표정 정보를 포함할 수 있다. 내부 모델은 사용자의 얼굴 표정을 결정하기 위해 컴퓨터 비전과 같은 얼굴 인식 기술을 포함할 수 있다. 시스템은 사용자가 웃고 있는지 여부를 결정하기 위해 확인할 수 있으며, 이는 사용자가 실제로 기쁘하다는 확인을 나타낼 수 있다. 얼굴 표정이 사용자가 찡그린다고 판단하거나 슬피하고 있다고 판단하면, 시스템은 모든 입력을 검사하여 "기쁘다(glad)"라는 음성 입력이 정확하지 않고 실제로는 "슬프다(sad)"이었다는 것을 식별할 수 있다. 이 예에서 예시된 바와 같이, 다중 모드 모델은 가상 캐릭터의 활동 정확도를 증가시키기 위해 다양한 하위 시스템 또는 내부 모델의 정확도를 상호 참조하고 확인할 수 있다.
- [0034] 다른 예로서, 사용자가 슈퍼히어로를 나타내는 가상 캐릭터를 말하고 음성 입력을 제공하는 경우, 시스템은 "내가 가장 좋아하는 악당은 초커이다(My favorite villain is the Choker)"라는 음성 입력을 수신할 수 있다. 음성 입력의 단어를 식별한 후, 시스템은 지식 모델을 활용하여 "초커(Choker)"라는 단어를 검사하고 그것이 내부 모델에서 슈퍼히어로와 관련된 것으로 알려진 캐릭터인 "조커(Jocker)"라는 단어와 문법적으로 가깝다는 것을 식별할 수 있다. 지식 모델은 사용자가 실제로 "조커"를 의미했다고 제안할 수 있으며, 이는 슈퍼히어로를 나타내는 가상 캐릭터에 의해 제공되는 출력의 정확도를 증가시킬 수 있다.
- [0035] 일부 실시예에서, 다중 모드 모델은 장치에 연결된 센서(예를 들어, 카메라, 마이크)로부터 캡처된 음성 데이터, 얼굴 표정 데이터 및 환경 정보를 수신할 수 있다. 음성 데이터는 사용자의 단어(예를 들어, 사용자가 말하고 있는 것으로 구성된 텍스트의 스트림)의 자동 음성 인식(automatic speech recognition(ASR))을 사용하여 검사 및 처리될 수 있다. 얼굴 표정 정보는 얼굴 표정 인식 알고리즘(예를 들어, 많은 일반적인 얼굴 표정을 인식할 수 있는 알고리즘)을 사용하여 검사 및 처리될 수 있다. 환경 정보는 많은 객체(예를 들어, 개, 식물, 커피잔, 다른 사람 등과 같은 일반적인 객체)를 인식하는 세계 인식 컴퓨터 비전 알고리즘에 의해 검사 및 처리될 수 있다. 세계 인식 알고리즘은 사용자의 물리적 세계(내부에 있는지 외부에 있는지, 밝은지 어두운지 등의 여부)에 대한 다른 데이터 포인트를 인식할 수 있다.

- [0036] 일부 실시예에서, 하나 이상의 컴퓨터 비전 알고리즘은 사용자 장치가 다수의 카메라를 갖는 장치를 갖는 장치 (예를 들어, 스마트폰)에 출력된 증강 현실을 활용하고 있는 경우에만 활성화될 수 있다. 일부 경우에, 사용자 장치는 증강 현실 디스플레이를 출력하지 않고 하나의 카메라에만 액세스할 수 있거나 카메라에 액세스하지 않을 수 있다.
- [0037] 하나의 내부 모델은 음성 입력 정보에 기초하여 ASR 데이터로부터 주어, 시점(point of view(POV)), 사건(event), 의도, 명사, 동사 등을 식별할 수 있는 자연어 이해(natural language understanding(NLU)) 모델을 포함할 수 있다. 다시 말해서, 내부 모델이 ASR을 통해 인식된 사용자로부터의 단어를 수신할 때, 내부 모델은 사용자의 음성 데이터와 관련된 인사이트(insight) 및 컨텍스트(context)의 추가 계층을 추가할 수 있다.
- [0038] 일례로서, 사용자는 "나는 하루 종일 학교에 있었고 기타 치는 법을 배웠다. 정말 재미있었다.(I was at school all day and I learned how to play the guitar. It was super fun.)"라고 말할 수 있다. 내부 모델은 사용자가 1인칭으로 말하고 있고, 그들이 주어이며, 그들의 의도가 기타를 배우는 것이었다는 것을 추출할 수 있다. 다중 모드 모델은 내부 모델을 검사하고 "나는 당신이 기타를 배우는 것을 즐기고 있어 기쁩니다!(I am glad you are having fun learning the guitar!)"와 같은 응답을 생성할 수 있다. 이것은 시스템이 "기타(guitar)"라는 단어를 이해하지만 다른 것과 관련하여 어떻게 사용되었는지, 화자가 이에 대해 어떻게 느꼈는지 또는 목표가 무엇인지 판단하지 않는 ASR 기술만 사용하는 것보다 더 정확할 수 있다. 따라서, 시스템은 사용자의 음성 입력에 긍정적 또는 부정적으로 응답할 수 없을 것이다.
- [0039] 일부 실시예에서, 내부 모델은 가상 캐릭터의 세계를 나타내는 지식 모델을 포함할 수 있다. 예를 들어, 클라이언트가 캐릭터의 가상 캐릭터를 생성하기 위한 영화 스튜디오인 경우, 지식 모델은 캐릭터를 가장 잘 묘사하기 위해 그 캐릭터와 관련된 정보를 가질 수 있다. 이 정보는 캐릭터 성격의 특정 양태를 나타내는 '페르소나(persona)'를 포함할 수 있다. 지식 모델은 다른 사용자와 가상 캐릭터의 역학을 해석하는 소셜 시뮬레이션(social simulation)을 포함할 수 있다. 일부 실시예에서, 지식 모델은 다양한 소스(예를 들어, 스크립트, 만화책, 책, 사용자 입력 등)로부터 캐릭터에 대한 정보를 수신함으로써 그 캐릭터와 관련된 추가 정보를 자동으로 생성할 수 있다.
- [0040] 시스템은 원시 데이터(예를 들어, 스트리밍 음성 데이터, 실시간 음성-텍스트)를 수신하고, 얼굴 표정 인식 데이터를 캡처하고, 사용자의 물리적 환경 및 물리적 객체에 대한 세계 인식 데이터를 캡처할 수 있다. 시스템은 원시 데이터를 서로 비교하면서 내부 모델, NLU 알고리즘, 지식 모델, 페르소나 및 소셜 시뮬레이션을 실행하여 사용자로부터 추가 컨텍스트와 의미를 도출할 수 있다. 이 정보 및 모델에 기초하여, 다중 모드 모델은 더 높은 정확도로 가상 캐릭터에 대한 출력을 생성할 수 있다.
- [0041] 컴퓨터가 세계에 대한 정보를 얻기 위해 센서를 사용할 때, 센서에 의해 변환된 데이터는 묘사하려고 시도한 세계와 관련하여 오류를 포함할 수 있다. 센서는 실제 이벤트에 의해 생성된 정보를 샘플링할 수 있다. 이러한 정보 손실 외에도, 센서에 의해 생성된 데이터 스트림을 해석하는 계산 방법도 또한 오류가 발생하기 쉽다. 이 패러다임에서의 오류에 대한 2가지 주요 위치가 있을 수 있다: 센서에서 데이터 신호로의 변환에서 그리고 데이터 스트림을 해석하는데 있어서. 다중 모드 오류 정정은 센서 출력이 데이터 신호로 변환될 때 유입되는 일부 오류를 감소시키는 데 도움이 될 수 있다. 또한, 데이터 신호가 더 높은 정확도를 가지면, 현실을 더 정확하게 나타내기 때문에, 이를 해석하기가 더 쉬울 수 있다. 이것은 다중 모드 오류 정정이 여러 측면에서 정확도를 개선하는 데 도움이 될 수 있다.
- [0042] 도 2는, 다양한 실시예에 따라, 다중 모드를 사용하여 가상 캐릭터의 오류를 정정하기 위한 순서도(200)를 도시한다. 도 2에 도시된 바와 같이, 다수의 센서(예를 들어, 이미지 센서(202A), 오디오 센서(202B), 후각 센서(202C), 센서 N(202N))는 입력 레이어(204)에서 데이터를 제공할 수 있다. 입력 레이어(204)로부터의 데이터는 디지털 이미지 신호(206A), 디지털 오디오 신호(206B), 디지털 후각 특징 신호(206C), 디지털 신호 N(206N) 등을 포함할 수 있다. 이 입력 데이터는 컴퓨터 비전 전처리(210A), 오디오 전처리(210B), 후각 전처리(210C), 전처리 유형 N(210N) 등을 포함할 수 있는 단일 모드 정보 추출(208)을 사용하여 처리될 수 있다. 다중 모드 오류 정정(216)은 처리된 정보를 수신하고 정보 스트림에서 오류를 검출할 수 있다. 예를 들어, 시각 정보 스트림(218A), 오디오 정보 스트림(218B), 후각 정보 스트림(218C), 정보 스트림 N(218N), 사전 지식(220) 및 메모리는 다중 모드 오류 정정(216)에 의해 처리되어 출력 레이어(224)에서 증강 스트림(예를 들어, 226A-N)을 도출할 수 있다.
- [0043] 각각의 장치는 다른 장치와 상호 작용할 때 고유한 버전의 다중 모드 오류 정정을 사용할 수 있다. 상식 및 세계 지식과 함께 모든 의미로부터 정보가 병합되어 현실 세계에서 실제로 일어나고 있는 일을 해독할 수 있다.

컴퓨터는 다른 모드의 정보와 세계의 이전 정보(지식 베이스 및 모델로 인코딩됨)를 합성하는 방법을 지시받을 수 있다. 다중 모드 오류 정정 스테이지에서의 각각의 연결은 최종 데이터 스트림을 개선하기 위해 서로 다른 출처로부터의 데이터 신호를 사용하는 것을 나타낼 수 있다.

- [0044] 데이터 스트림의 품질을 개선하기 위해 서로 다른 센서로부터의 정보가 어떻게 사용될 수 있는지에 대한 일부 예는 ASR 결과를 개선하기 위해 컴퓨터 비전 기반 입술 판독과 함께 오디오 기반 ASR을 사용하는 것, ASR 또는 이미지 인식을 정정하기 위해 후각 센서를 사용하는 것, ASR을 정정하기 위해 이미지/객체 인식을 사용하는 것, 이미지 인식을 정정하기 위해 ASR을 사용하는 것, ASR, CV 등을 개선하기 위해 세계에 대한 지식(실제 세계이든 가상 세계이든)을 사용하는 것을 포함할 수 있다.
- [0045] 일부 실시예에서, 가상 캐릭터가 구체적으로 언급되지 않은 이유는 이러한 오류 정정 프로세스가 상이한 유형의 시스템에서 사용될 수 있기 때문일 수 있다. 시스템은 말하는 내용, 만들어지고 있는 소리, 실제 세계에 대한 것 및 캐릭터에 대한 사용자의 물리적 반응에 대한 것의 인식을 가상 캐릭터에게 제공하는 데 사용할 수 있지만, 자율 주행차 또는 스마트 홈과 같은 다른 유형의 컴퓨터 시스템에 의해 사용될 수 있다.
- [0046] 일부 실시예에서, 다중 모드 오류 정정 페이지의 목적은 허위 부정(false negative)을 검출하고 그 값을 재분류하는 것일 수 있다. 이것은 정보 추출 파이프라인의 회수(recall)를 향상시킬 수 있다.
- [0047] 도 2에 도시된 바와 같이, 세계 지식 자산이 통합될 수 있다. 도 2에서, 세계 지식은 파이프라인의 다중 모드 오류 정정 페이지에 대한 입력을 포함할 수 있다. 세계 지식은 입력 레이어에서 센서에 의해 추출된 신호와 별도로 표시될 수 있다. 입력 레이어의 센서에 의해 추출된 정보는 현재 시간에 세계에서 발생하는 이벤트에 대한 정보를 캡처할 수 있다. 세계 정보의 예는 사용자가 말하는 내용, 카메라 프레임에서 검출될 수 있는 객체 등을 포함한다.
- [0048] 일부 실시예에서, 일부 세계 지식 자산은 사전에 수집될 수 있다. 세계 지식 자산은 실제 세계에 관한 사실 정보이거나 이야기 세계에 대한 허구적 정보(가공 인물의 세계와 같이)일 수 있는 세계에 대한 사실을 포함할 수 있다. '머그(mug)'라는 객체가 먹고 마시는 것과 의미적으로 관련된 이벤트에 대하여 사용될 수 있는 '컵(cup)'의 일종인 것처럼, 상식 지식이 세계 지식 자산에 포함될 수 있다. 또한, 가공 인물의 대포(cannon)대한 정보도 세계 지식 자산으로 코드화할 수 있다. 예를 들어, 배트맨®의 세계에서 "조커"는 영웅 "배트맨"에 대하여 심리 게임을 하는 악당이다. 배트맨 세계의 다른 악당은 "포이즌 아이비(Poison Ivy)"와 "서브 제로(Sub-Zero)"를 포함할 수 있다. 이 정보는 구조화된 지식 베이스 또는 온톨로지로 인코딩될 수 있다.
- [0049] 세계 지식/자산은 과거에 수집된 정보를 나타내기 때문에 센서 데이터 스트림과 분리될 수 있는 반면, 입력 레이어에서의 센서로부터의 신호는 컴퓨터 시스템과의 상호 작용이 활성화되어 있는 동안 지속적으로 획득되는 데이터 신호를 나타낼 수 있다. 세계 지식 베이스는 정적이거나 상호 작용/사용자 세션 동안 증가하지 않을 수 있다. 지식 베이스는 미리 선별되고 채워질 수 있으므로, 컴퓨터는 이 코드화된 정보를 효율적으로 사용하여 더 나은 결정을 내리거나 일부 상황에서는 다중 모드 오류 정정 프로세스를 지원할 수 있다.
- [0050] 이 정보는 다중 모드 오류 정정 프로세스를 향상시킬 수 있다. 일례로서, 정보는 배트맨의 세계(고담시)에 대해 컴퓨터 시스템과 대화하는 사용자를 포함할 수 있다. 사용자가 무언가를 말하면, ASR 시스템은 사용자가 "조커는 나쁜 녀석이다(The choker is a bad dude)"라고 말한 것으로 생각한다. 배트맨의 세계에서, "조커"는 악당이고, 사용자가 "조커는 나쁜 녀석이다"라고 말했을 가능성이 훨씬 더 크다. 시스템은 스마트 홈, 자동차 또는 사용자와 상호 작용하는 기타 장치에 사용될 수 있다.
- [0051] 도 3은, 다양한 실시예에 따라, 가상 캐릭터를 제어하기 위한 다중 모드 모델의 순서도(300)를 도시한다. 시스템은 사용자로부터 입력을 수신하는 것과, 다중 모드 모델을 위한 내부 모델을 구현하는 것을 포함할 수 있다.
- [0052] 가상 캐릭터의 특징을 구현하는 내부 모델은 가상 캐릭터가 실시간 또는 거의 실시간으로 동작하고 응답하도록 할 수 있다. 다양한 입력(사용자로부터의 음성 입력, 사용자의 얼굴 표정 인식)을 분리하여 별도로 처리하는 대신, 본 실시예는 가상 캐릭터에 의해 동작으로 표현되는 더욱 정확한 애니메이션/음성/출력을 이해하고 생성하기 위해 다수의 입력을 입력 및 검사할 수 있다. 수신된 입력의 예는 사용자의 음성(302), 사용자의 얼굴 표정(304) 및 사용자의 환경(306)과 관련된 정보를 포함할 수 있다(예를 들어, 비디오 피드의 배경을 스캔하고 컴퓨터 비전 알고리즘이 애완동물, 텔레비전, 가구 등과 같은 일반적인 물건을 인식할 수 있는지 보는 것에 의해).
- [0053] 또한, 입력은 자연어 이해 데이터(308), 지식 모델 데이터(310), 소셜 시뮬레이션 데이터(312) 등과 같은 내부 모델 입력을 포함할 수 있다. 다중 모드 모델(314)은 가상 캐릭터에 의해 취해지는 동작(들)을 나타내는 출력

(316)을 생성하기 위하여 입력(302 내지 310)를 처리할 수 있다. 본 실시예는 다수의 입력을 수신하고 입력을 서로 비교하여 가상 캐릭터에 대한 정확한 출력을 생성할 수 있다.

- [0054] 도 4는, 다양한 실시예에 따라, 다중 모드 모델에 구현된 가상 캐릭터와 상호 작용하기 위한 시스템(400)을 도시한다. 도 4에 도시된 바와 같이, 사용자는 사용자 장치(예를 들어, 스마트폰, 증강 현실 안경, 컴퓨터, 게임 콘솔 등)에서 하나 이상의 센서(예를 들어, 카메라, 마이크 등)로부터 입력 정보(예를 들어, 음성, 얼굴 표정, 환경 정보 등)를 제공할 수 있다. 다중 모드 모델은 사용자 장치 중 하나에서 실행되거나 인터넷과 같은 네트워크를 통해 사용자 장치에 연결된 외부 장치에서 실행될 수 있다. 다중 모드 모델은 입력 정보를 검사하고, 내부 모델을 실행하고, 센서에 의해 캡처된 디스플레이에서 가상 캐릭터를 나타내는 디스플레이를 출력할 수 있다. 예를 들어, 도 4에 도시된 바와 같이, 가상 캐릭터는 센서에 의해 캡처된 환경 내의 위치에 디스플레이된다.
- [0055] 일례로서, 사용자 장치는 객체(예를 들어, 객체(404))를 포함하는 환경(402)을 캡처할 수 있다. 디스플레이된 환경(406) 및 식별된 객체(408)는 가상 캐릭터(410)와의 인터페이스 상에 제공될 수 있다. 센서(412)는 사용자 장치를 둘러싸는 환경을 캡처할 수 있다. 예를 들어, 센서(412)는 사용자(414)에 관한 정보를 캡처할 수 있다.
- [0056] 다중 모드 모델은 애니메이션과 음성 모두를 포함하는 가상 캐릭터에 대한 출력을 생성할 수 있다. 가능한 애니메이션은 캐릭터에 대한 잠재적인 애니메이션의 라이브러리에 저장될 수 있다. 이것은 각각의 애니메이션에 대응하는 사전 녹음된 오디오 파일의 유사한 라이브러리와 연관될 수 있다. 라이브러리는 각각의 캐릭터에 대해 구축될 수 있으며 다른 캐릭터에 대해 별도의 라이브러리를 생성하기 위한 기반을 제공할 수 있다.
- [0057] 일부 실시예에서, 시스템은 모든 가상 캐릭터에 걸쳐 공통인 범용 리깅 모듈(universal rigging module)을 포함할 수 있다. 리깅은, 예를 들어, 3D 캐릭터의 뼈, 근육 및 피부를 기본적으로 생성하는 3D 애니메이션에서의 기술이다. 우리는 네트워크에서 하나 이상의 가상 캐릭터에 걸쳐 범용 리깅 모듈을 사용할 수 있다.
- [0058] 시스템은 AI에 의해 구동되는 절차 애니메이션을 생성할 수 있다. 다시 말해, AI가 "나도 기타를 사랑해요(I love the guitar too)"라고 말하고 싶다면 AI는 가상 캐릭터에게 미소를 지으라고 지시할 수 있으며 미소 짓는 그 특정 캐릭터에 대한 정적 애니메이션 파일이 없어도 절차적으로 이루어져야 한다.
- [0059] 일부 실시예에서, 다중 모드 모델은 많은 오디오 파일을 사전 녹음할 필요 없이 음성을 생성할 수 있다. 다중 모드 모델은 캐릭터의 음성에 대해 비교적 적은 양의 시간(예를 들어, 5 내지 15시간)을 수신하고, AI가 말하려고 시도하는 임의의 것에 대해 올바른 음성을 생성할 텍스트-음성/음성 합성 시스템을 생성할 수 있다.
- [0060] 본 실시예는 증강 현실 디스플레이를 출력하도록 구성된 장치 상에서 실행될 수 있지만, 본 실시예는 터치 스크린, 모니터, 스크린 등과 같은 임의의 적절한 디스플레이 상에서 실행될 수 있다. 일부 실시예에서, 본 시스템은, 특히 시스템이 FER과 세계 인식 내부 모델을 동시에 사용하는 경우, 사용자 장치에서 전면 및 후면 카메라를 모두 사용할 수 있다.
- [0061] 일부 실시예에서, 장치 상에 디스플레이되는 가상 캐릭터를 제어하는 것은 장치(예를 들어, 스마트폰, 컴퓨터, AR 가능 장치) 상에 배치된 센서(예를 들어, 카메라, 마이크)로부터 입력 정보를 수신하는 것을 포함할 수 있다. 가상 캐릭터를 제어하는 것은 입력 정보의 특성을 식별하도록 구성된 적어도 2개의 내부 모델(예를 들어, 다중 모드 모델)을 구현하고, 내부 모델에 의해 식별된 특성을 조합하여 내부 모델에 의해 식별된 특성들 사이의 유사도를 식별하는 것을 포함할 수 있다. 특성들 사이에 유사도를 찾는 것은 제1 입력 정보 소스로부터 식별된 특성(예를 들어, "나는 행복하다(I am happy)"라고 말하는 사용자)이 제1 입력 정보 소스로부터 식별된 특성(예를 들어, 눈살을 찌푸리거나 슬퍼하는 사용자의 얼굴 표정)과 일치하지 않는지를 식별할 수 있다. 특성이 임계값 아래의 유사도를 갖는 경우, 내부 모델은 가상 캐릭터에 공통된 정보를 포함하는 지식 모델을 검사하는 것에 기초하여 더 정확할 가능성이 있는 특성을 식별할 수 있다.
- [0062] 가상 캐릭터를 제어하는 것은 내부 모델에 의해 식별된 특성과 일치하는 동작을 결정하기 위해 가상 캐릭터와 연관된 잠재적 동작의 라이브러리를 검사하는 것을 포함할 수 있다. 가상 캐릭터를 제어하는 것은 가상 캐릭터로 하여금 내부 모델에 의해 식별된 특성과 일치하는 동작을 수행하게 하도록 동작을 출력하는 것을 포함할 수 있다.
- [0063] 일부 실시예에서, 가상 캐릭터를 제어하는 것은 내부 모델에 의해 식별된 제1 특성의 유사도가 임계 유사도 아래로 떨어진다고 결정하는 것을 포함할 수 있다. 가상 캐릭터를 제어하는 것은 가상 캐릭터와 일반적으로 연관된 정보를 나타내는 엔트리의 목록을 나타내는 제3 내부 모델을 검사하는 것을 포함할 수 있다. 가상 캐릭터를 제어하는 것은 가상 캐릭터와 연관된 목록에 나열된 엔트리와 제1 특성을 일치시키는 것을 포함할 수 있으며,

여기서 내부 모델에 의해 식별된 특성은 목록에 나열된 엔트리를 포함한다.

- [0064] 일부 실시예에서, 입력 정보는 사용자의 음성을 나타내는 오디오 정보, 사용자의 얼굴 표정을 나타내는 얼굴 표정 정보 및 장치의 센서에 의해 캡처된 환경을 나타내는 환경 정보를 포함한다.
- [0065] 일부 실시예에서, 내부 모델은 오디오 정보로부터 컨텍스트 및 의미를 도출하도록 구성된 자연어 이해 모델, 가상 캐릭터에 특정한 정보를 식별하도록 구성된 지식 모델 및 사용자 및 기타 가상 캐릭터와 관련하는 데이터를 식별하도록 구성된 소셜 시뮬레이션 모델을 포함한다.
- [0066] 일부 실시예에서, 가상 캐릭터를 제어하는 것은 장치의 디스플레이 상에 가상 캐릭터를 디스플레이하는 것과, 사용자에게 대한 질의를 나타내는 초기 동작을 수행하도록 가상 캐릭터에게 지시하는 것을 포함할 수 있으며, 여기서 입력 정보는 질의에 대한 사용자에게 의한 응답을 나타낸다.
- [0067] 일부 실시예에서, 센서를 통해 획득된 정보는 지식 베이스(온톨로지) 내의 정보가 수정될 필요가 있게 할 수 있다. 예를 들어, 사용자가 "내가 가장 좋아하는 피자 토핑은 anchovy이다(My favorite pizza topping is anchovy)"라고 말하고, 현재 버전의 지식 베이스에 사용자가 가장 좋아하는 피자 토핑이 버섯이라고 되어 있으면, 지식 기반이 이제 잘못되었을 수 있으며 anchovy에 대한 새로운 선호를 반영하도록 업데이트될 필요가 있다. 따라서, 센서에 의해 획득된 새로운 정보는 지식 베이스에서 오류를 정정하거나 정보를 업데이트하는 데 사용될 수 있다.
- [0068] 일부 실시예에서, 가상 캐릭터는 사용자의 과거에 대한 정보에 액세스할 수 있다. 이것은 그들의 생활에서의 사건의 타임라인을 포함할 수 있다. 예를 들어, 가상 캐릭터는 사용자가 5년 전에 빵집을 소유했다는 것을 알고 있을 수 있다. 이 정보는 다양한 방법으로 활용될 수 있다. 사용자가 제빵사라는 것을 캐릭터가 알고 있다면, 예를 들어, "knead"와 같은 제빵사에 대한 어휘 또는 전문 용어의 전체 세트가 있을 수 있다. "knead"라는 단어는 "need"라는 단어처럼 들리며, 자동 음성 인식에서 종종 잘못 표기될 수 있다. 많은 시스템은 "knead"라는 단어를 "need"로 잘못 표기할 수 있다. 사용자가 제빵사라는 것을 가상 캐릭터가 알고 있다면, 시스템은 사용자가 "knead"는 단어를 말하고 이를 더욱 정확하게 표기할 가능성이 훨씬 더 높다는 것을 이해할 수 있다. 사용자가 제빵사였다는 것을 캐릭터가 알면, 직업에 기초하여 사용할 단어 유형을 아는 것으로 더욱 정확하게 음성을 표기할 수 있다.
- [0069] 웹 브라우저에서 실행되는 완전 상호 작용형 가상 캐릭터
- [0070] 많은 경우에, 사용자는 가상 캐릭터를 생성하기 위해 사용자 장치에서 가상 캐릭터를 렌더링하는 데 필요한 정보를 다운로드한다. 사용자가 가상 캐릭터 정보를 다운로드하게 하기 위해, 클라이언트는 증강 현실 가상 캐릭터와 연관된 애플리케이션을 다운로드하도록 사용자를 설득할 수 있다. 사용자 장치에서 특정 애플리케이션을 명시적으로 다운로드하는 단계를 수행하는 것은 사용자에게 마찰을 유발할 수 있으며, 이는 사용자가 가상 캐릭터를 구현하지 못하게 할 수 있다.
- [0071] 본 실시예는 여기에 설명된 바와 같은 다중 모드 모델을 활용하는 3D 가상 캐릭터를 렌더링 및 생성하는 동적으로 렌더링된 가상 캐릭터(또는 인스턴트 가상 캐릭터)에 관한 것일 수 있다. 가상 캐릭터는 캐릭터를 닮은 3D 비주얼 "바디"를 포함할 수 있다. 인스턴트 가상 캐릭터는 하이퍼링크에 임베딩될 수 있다. 인스턴트 가상 캐릭터는 소셜 미디어 네트워크와 같은 다양한 네트워크에서 공유될 수 있다. 예를 들어, 회사는 인기 캐릭터 또는 유명인의 가상 캐릭터를 한 그룹의 팬과 공유할 수 있는 엔터테인먼트 회사일 수 있다. 회사는 하이퍼링크 또는 기타 식별자(예를 들어, YouTube® 비디오에 대한 링크), 문자 메시지 등을 공유할 수 있다. 네트워크(예를 들어, 소셜 미디어)를 통해 공유될 때, 인스턴트 가상 캐릭터는 사용자가 이를 보기 위해 애플리케이션을 다운로드할 필요 없이 소셜 미디어 피드에서 또는 웹 브라우저를 통해 즉시 실행될 수 있다.
- [0072] 인스턴트 가상 캐릭터는 웹 브라우저에서 실행될 수 있다. 이것은 사용자와 연관된 스마트폰으로의 가상 캐릭터와의 상호 작용을 가능하게 할 수 있다. 다중 모드 모듈은 짧은 로딩 시간(예를 들어, 1초)으로 모바일 웹 브라우저(예를 들어, 사파리®, 크롬®)를 통해 웹 기반 증강 현실로 사용자를 원활하게 마이그레이션할 수 있다. 따라서, 사용자는 "이 링크를 클릭하면 유명인이 당신의 집에 나타납니다."라는 웹페이지(예를 들어, 트위터®)상의 링크를 선택할 수 있다. 링크를 클릭함으로써, 가상 캐릭터가 거의 실시간으로(예를 들어, 1초 미만) 화면에 나타날 수 있다.
- [0073] 시스템은 적절한 3D 파일 포맷(예를 들어, GLTF(GL Transmission Format))으로 고도로 최적화된 3D 파일을 생성할 수 있다. 가상 캐릭터에 대한 애니메이션 및 오디오 파일은 사용자 장치에 배치(batch)로 스트리밍될 수 있어, 가상 캐릭터를 실행하는 데 필요한 이전 배치가 먼저 사용자 장치에 전송되고 우선시된다. 정보의 다른

배치를 위한 공간을 더 확보하기 위해 배치는 실행 후에 폐기될 수 있다. 사용자에게 송신된 배치는 사용자로부터의 입력과 가상 캐릭터에 의해 생성된 응답에 기초하여 수정되어 동적일 수 있다.

- [0074] 일부 실시예에서, 시스템은 인스턴트 가상 캐릭터가 디스플레이 상의 적절한 위치에 위치되도록 사용자의 환경(예를 들어, 바닥)의 웹 기반 추적을 포함할 수 있다.
- [0075] 일부 실시예에서, 시스템은 실시간 또는 거의 실시간으로 사용자의 모바일 웹 브라우저에서 음성 인식, 자연어 이해, 얼굴 표정 인식 및 세계 인식 알고리즘을 수행한다.
- [0076] 일부 실시예에서, 시스템은 컴퓨터 비전 알고리즘을 사용하여 전경에 있는 객체를 식별한다. 예를 들어, 시스템은 가상 캐릭터가 소파 뒤에 서 있다고 결정할 수 있으며, 여기서 알고리즘은 소파에 의해 적절하게 가려지도록 수정하고 가상 캐릭터를 따라서 그가 그 뒤에 서 있는 것처럼 보일 수 있다.
- [0077] 일부 실시예에서, 시스템은 2명 이상의 사람들이 별도의 사용자 장치(예를 들어, 스마트 폰)를 통해 동일한 장소에서 가상 캐릭터를 볼 수 있도록 다중 사용자 세션을 지원한다. 컴퓨터 비전 알고리즘은 사용자의 환경을 추적하고 여러 장치가 약간 다른 세계 관점을 가지더라도 이를 실시간으로 다수의 장치와 동기화할 수 있다.
- [0078] 일부 실시예에서, 시스템은 다른 장소에서 인스턴트 가상 캐릭터의 동기화된 보기를 지원한다. 예를 들어, 다중 모드 모델이 유명인을 위한 가상 캐릭터(예를 들어, 유명 뮤지션)를 생성하고 그녀가 동시에 많은 사람들을 위해 노래를 공연하기를 원하는 경우, 시스템은 동시에 또는 거의 동시에 모두 음악을 공연하는 유명인 가상 캐릭터의 큰 그룹을 생성할 수 있다.
- [0079] 일부 실시예에서, 시스템은 3D 모델로 변환될 수 있는 자신 또는 다른 대상의 이미지를 촬영함으로써 사용자가 자신의 인스턴트 가상 캐릭터를 생성하는 것을 지원할 수 있다.
- [0080] 일부 실시예에서, 다중 모드 모델을 사용하여 가상 캐릭터를 동적으로 생성하는 것은 가상 캐릭터를 생성하기 위해 웹 브라우저를 애플리케이션에 링크하는 웹페이지에 대한 링크를 임베딩하는 것을 포함한다. 가상 캐릭터를 동적으로 생성하는 것은 링크가 선택되었다는 선택을 장치로부터 수신하는 것을 포함할 수 있다. 이것은 장치와 상호 작용하는 사용자가 장치의 디스플레이에 가상 캐릭터가 디스플레이되기를 원한다는 것을 나타낼 수 있다. 가상 캐릭터를 동적으로 생성하는 것은 가상 캐릭터를 생성하기 위한 정보를 나타내는 애플리케이션으로부터의 데이터 스트림을 전송하는 것 및 장치의 디스플레이에 가상 캐릭터를 디스플레이하는 것을 포함할 수 있다.
- [0081] 가상 캐릭터를 동적으로 생성하는 것은 장치에 배치된 센서로부터 입력 정보를 수신하는 것과, 다중 모드 모델에 포함된 적어도 2개의 내부 모델을 구현하는 것을 포함할 수 있다. 내부 모델은 입력 정보의 특성을 식별하고 내부 모델에 의해 식별된 특성을 조합하여 내부 모델에 의해 식별된 특성들 사이의 유사도를 식별하도록 구성될 수 있다. 가상 캐릭터를 동적으로 생성하는 것은 내부 모델에 의해 식별된 특성과 일치하는 동작을 결정하기 위해 가상 캐릭터와 연관된 잠재적 동작의 라이브러리를 검사하는 것과, 가상 캐릭터가 동작을 수행하게 하는 명령어를 나타내는 내부 모델에 의해 식별된 특성과 일치하는 동작을 출력하는 것을 포함할 수 있다.
- [0082] 일부 실시예에서, 가상 캐릭터를 동적으로 생성하는 것은 가상 캐릭터에 대한 제1 동작을 개시하는 것을 포함하고, 제1 동작은 가상 캐릭터에 의해 수행되는 음성 및 애니메이션을 포함하는 사용자에게 대한 질의를 나타내고, 입력 정보는 쿼리에 대한 사용자의 응답이다.
- [0083] 일부 실시예에서, 가상 캐릭터를 동적으로 생성하는 것은 임베디드 링크를 네트워크를 통해 복수의 사용자에게 공유하는 것, 링크가 선택되었다는 것을 나타내는 선택을 장치 세트로부터 수신하는 것 및 장치 세트에 포함된 각각의 장치에 가상 캐릭터를 생성하기 위한 정보를 나타내는 데이터 스트림을 애플리케이션으로부터 전송하는 것을 포함한다.
- [0084] 일부 실시예에서, 가상 캐릭터를 동적으로 생성하는 것은 제1 시간에 데이터 스트림의 제1 배치를 전송하는 것과, 제1 시간 후의 제2 시간에 데이터 스트림의 제2 배치를 전송하는 것을 포함하고, 제1 배치는 가상 캐릭터를 초기에 생성하기 위한 정보를 포함하고, 제2 배치는 가상 캐릭터에 의해 제1 동작을 출력하기 위한 정보를 포함하고, 제1 배치는 제2 시간에 폐기된다.
- [0085] 일부 실시예에서, 가상 캐릭터를 동적으로 생성하는 것은 환경 정보를 검사하여 환경에 포함된 바닥을 식별하는 것과, 가상 캐릭터를 환경에서 식별된 바닥 바로 위의 제1 위치에 위치 설정하는 것을 포함한다.
- [0086] 가상 캐릭터와의 다중 모드 상호 작용으로부터의 인사이트 추출

- [0087] 클라이언트(예를 들어, 기업)는 여기에 설명된 실시예 중 적어도 일부를 사용하여 가상 캐릭터를 생성할 수 있다. 클라이언트는 특정된 기간 동안 가상 캐릭터와 상호 작용하는 많은 사람들(예를 들어, 백만 명)을 가질 수 있다. 일례로서, 엔터테인먼트 회사는 많은 사람들을 위해 유명인의 가상 캐릭터를 동시에 생성할 수 있다.
- [0088] 본 실시예는 가상 캐릭터와의 사용자의 상호 작용으로부터 인사이트(insight)를 추출하는 것과 적어도 부분적으로 관련될 수 있다. 인사이트는 가상 캐릭터와의 상호 작용 동안 임의의 센서 세트에서 수집된 데이터에서 자동으로 추출될 수 있다.
- [0089] 도 5는, 다양한 실시예들에 따른, 장치(500) 상에 디스플레이되는 예시적인 가상 캐릭터(502)를 도시한다. 가상 캐릭터는 실제 인간, 가공 인간 또는 외계인 또는 말하는 개와 같은 상상의 캐릭터의 표현과 같은 움직이는 엔티티의 가상 표현을 포함할 수 있다. 가상 캐릭터는 컴퓨터 스크린 상에서, 가장 전형적으로는, 본 특허의 맥락에서, 스마트 폰 상에서 볼 수 있다. 가상 캐릭터는 화면에서 애니메이션화되는 바디의 표현을 가질 수 있다. 가상 캐릭터는 시각적으로 나타날 필요가 없으며 음성으로만 자신을 나타낼 수 있다. 또한, 단지 텍스트(채팅) 또는 이미지를 통해 의사 소통할 수도 있다.
- [0090] 가상 캐릭터는 스마트 폰에서 스피커를 통해 말함으로써 인간 사용자와 대화할 수 있다. 컴퓨터가 말하는 오디오는 텍스트-음성 합성기에 의해 합성될 수 있다. 가상 캐릭터는 말하는 행위를 통해 현실 세계에 영향을 미칠 수 있지만, 가상 캐릭터는 조명을 켜고 끄고, 음악을 켜고, 방을 통과하는 로봇의 움직임을 제어하는 것과 같이 다른 방식으로 세계에 영향을 줄 수 있다.
- [0091] 가상 캐릭터의 주요 특징은 활동성(animacy)을 포함할 수 있다는 것이다. 활동성은 현실 세계에 영향을 미치는 동작을 수행하는 캐릭터의 능력을 포함할 수 있다. 현실 세계(또는 가상 세계)에서 무언가에 반응하는 방법을 결정하는 가상 캐릭터의 능력은 이것이 활동하게 만드는 것이다.
- [0092] 도 6은, 다양한 실시예에 따라, 사용자와, 가상 캐릭터를 디스플레이하는 사용자 장치 사이의 상호 작용의 순서도(600)를 예시한다. 가상 캐릭터와의 상호 작용은 사람(602)과 가상 캐릭터(606)가 장치(604)를 통해 서로 의사 소통할 때를 포함할 수 있다. 사람 및 가상 캐릭터는 음성(608) 및/또는 바디 랭귀지(610)와 같은 다양한 형태로 의사 소통을 작동시킬 수 있다. 가상 캐릭터와의 상호 작용은 사람이 가상 캐릭터와 능동적으로 관계를 맺을 때를 포함할 수 있다. 가상 캐릭터는 사람이나 동물이 움직이는 것과 같은 방식으로 움직이지 않을 수 있지만, 가상 캐릭터는 활동성을 시뮬레이션할 수 있으며, 가상 캐릭터는 의미상 일관성 있고 심지어 흥미로운 방식으로 반응할 수 있다. 가상 캐릭터는 합성 음성(612) 및/또는 애니메이션화된 바디 랭귀지(614) 중 임의의 것을 제공할 수 있다.
- [0093] 일부 상호 작용은 대화의 형태일 수 있다. 가상 캐릭터는 휴대폰의 마이크를 사용하여 인간 사용자가 가상 캐릭터에게 말하고 있는 것을 들을 수 있다. 마이크는 음파를 음성-텍스트 모델에 의해 텍스트로 변환될 수 있는 오디오 파일로 변환하고, 그 다음 이 텍스트는 자연어 이해 모델에 의해 해석될 수 있다. 가상 캐릭터가 인간 사용자가 말한 내용을 이해하면, 음성을 합성하는 것, 물리적 제스처 응답에서의 이의 바디를 움직이게 하는 것 또는 양 유형의 반응의 조합을 통해 인간 사용자에게 다시 응답하는 방법을 결정할 수 있다.
- [0094] 또한, 가상 캐릭터는 가상 공간에서 움직여 의사 소통할 수도 있다. 가상 공간은 증강 또는 가상 현실을 나타내는 휴대폰 또는 증강 또는 가상 현실을 나타낼 수 있는 헤드셋과 같이 컴퓨터 화면에서 시각화될 수 있는 3차원 공간의 임의의 표현을 포함할 수 있다. 가상 공간에서 움직이는 가상 캐릭터의 몇 가지 예는 말하는 동안 입술이 애니메이션화되어 가상 입술을 사용하여 대화하고 있다는 것을 알리는 것, 바디 랭귀지를 변경하여 다양한 기분이나 감정을 표현하는 것, 얼굴 표정을 변경하여 다른 감정을 알리는 것, 몸의 일부의 위치 설정을 바꾸어 다른 기분을 나타내는 것(방어를 위해 팔짱을 끼고, 우위를 보여주기 위해 엉덩이에 손을 대는 등) 등을 포함할 수 있다.
- [0095] 인사이트는 원시 데이터로부터 추출된 구조화된 정보 또는 이해를 포함할 수 있으며, 이는 이러한 이해가 정확하다는 것을 우리가 얼마나 확신하는지 측정하는 일종의 통계와 결합될 수 있다. 일례로서, 농담이 재미있는지 테스트하기 위해, 시스템은 인스턴트 가상 캐릭터를 다수의 사람들(100명 또는 1000명)에게 보낼 수 있다. 가상 캐릭터는 각각의 사용자(또는 피험자)에게 농담을 말한다. 그러면, 인스턴트 가상 캐릭터는 농담이 전달된 후의 사용자의 반응을 기록한다. 시스템은 사용자의 반응을 분석하여 농담이 재미있었는지 여부를 결정하기 위하여 사용자의 반응을 분석할 수 있다. 사용자는 이를 웃는 것을 통하는 것과 같은 많은 다양한 방식으로 표현할 수 있고, 이는 사용자의 얼굴에서 웃는 것을 찾기 위해 컴퓨터 비전을 사용하고 웃음 사운드를 찾기 위하여 사용자의 오디오를 분석함으로써 컴퓨터에 의해 자동으로 분류될 수 있다. 웃음이 없거나 사용자가 부정적인 반응을

보이면, 이것도 또한 컴퓨터에 의해 자동으로 선택될 수 있으며, 이는 사용자가 농담이 재미있다고 생각하지 않았다는 것을 의미할 수 있다. 이 농담 테스트를 많은 사람들에게 보내는 실험의 규모로 인해, 시스템은 농담이 얼마나 재미있었는지에 대한 통계를 수집할 수 있다. 이 실험의 결과는 인사이트이다; (구조화된 이해의 일부인) 농담이 웃겼는지 여부를 아는 것과, (구조화된 이해에 대해 얼마나 확신하는지를 측정하는) 그것이 얼마나 재미있는지 통계적으로 아는 것.

[0096] 도 7은, 다양한 실시예에 따라, 가상 캐릭터 반응으로부터 인사이트를 추출하는 순서도(700)를 도시한다. 인간 사용자가 가상 캐릭터와 상호 작용할 때 많은 유형의 정보가 수집될 수 있다. 인간은 통상적으로 가상 캐릭터에게 말을 할 것이다. 또한, 인간은 바디 랭귀지를 통해 자신이 어떻게 느끼는지를 의사 소통할 수 있다. 사람들이 말하는 것의 텍스트와 사람들이 사용하는 신체적 반응 또는 바디 랭귀지는 우리가 중점을 두는 2가지 유형의 인간 의사 소통일 수 있고, 본 실시예는 센서에 의해 포착되어 컴퓨터에 의한 해석을 위해 디지털 신호로 변환될 수 있는 임의의 다른 유형의 의사 소통에 적용될 수 있다.

[0097] 현실 세계로부터의 자극을 가상 캐릭터가 결정을 내리기 위해 기초하는 이러한 데이터의 디지털 표현으로 바꾸기 위한 동일한 프로세스가 그 데이터에서 인사이트를 추출하기 위하여 사용될 수 있다. 시스템은 인간 사용자에게 반응하는 방법을 결정할 때 가상 캐릭터가 액세스할 수 있는 동일한 데이터로부터 인사이트를 추출할 수 있다. 본질적으로, 인간 사용자는 반응(말을 하거나 물리적 제스처를 함으로써)을 가질 수 있으며, 이 현실 세계 자극은 입력 센서 어레이에 의해 포착되어 원시 디지털 신호로 바뀐다. 신호 처리가 이 신호에 대해 수행될 수 있다(하나의 모드 또는 다수의 모드에서). 이는 깨끗하게 된 정보 신호를 제공할 수 있다. 이 신호는 적어도 2가지 용도로 사용될 수 있으며, 첫째, 이는 나중에 이 데이터로부터 인사이트가 추출될 수 있도록 장기 스토리지에 저장될 수 있고, 둘째, 캐릭터가 현실 세계의 자극에 어떻게 반응할지 결정할 수 있도록, 정보 신호가 가상 캐릭터의 의사 결정 설비로 보내어질 수 있다.

[0098] 이러한 맥락에서, 인사이트를 추출하는 것은 현실 세계 이벤트를 가상 캐릭터가 결정을 내리기 위하여 기초하는 구조화된 이해로 변환하는 것과 유사할 수 있다. 깨끗한 정보 신호를 얻기 위한 데이터 처리 파이프라인은 다수의 사용 사례에서 동일할 수 있다. 가상과의 상호 작용 동안 사용자의 반응에 대한 정보가 저장되면, 인사이트가 상호 작용에 대해 추출될 수 있다. 여기서, 인사이트는 정제되지 않은 데이터 세트로부터 추출된 일부 이해 또는 상위 레벨 정보일 수 있다.

[0099] 본 실시예는 가상 캐릭터와의 상호 작용으로부터 인사이트를 추출하는 능력에 관한 것이다. 이것은, 가상 캐릭터와의 상호 작용이 더 많은 구조를 갖고 다중 모드 데이터 신호(음성, 시각적 데이터 스트림, 후각 데이터 스트림 등)를 생성하기 때문에, 텍스트 스트림으로부터 인사이트를 추출하는 것과 다를 수 있다.

[0100] 예를 들어, 장치는 사용자(702)에 의해 제공되는 자극을 나타내는 입력 센서 어레이(704)를 캡처할 수 있다. 원시 디지털 신호 어레이(706)는 다중 모드 디지털 신호 처리(708)를 사용하여 처리될 수 있다. 깨끗하게 된(또는 처리된) 정보 신호(710)는 임의의 인사이트 추출(712)을 통해 처리되고, 사용자 데이터 저장소(716)에 저장되고 그리고/또는 의사 결정 로직으로 처리되어 사용자 장치(718) 상의 가상 캐릭터(720)에 의한 반응을 출력하도록 가상 캐릭터 반응(714)을 도출할 수 있다.

[0101] 도 8은, 다양한 실시예에 따라, 인사이트 대시보드(800)를 도시한다. 도 8에 도시된 바와 같이, 클라이언트는 대시보드 또는 인사이트 포털을 보고 가상 캐릭터와 상호 작용하는 사용자를 나타내는 청중에 대한 많은 인사이트를 볼 수 있다. 사용자와 관련된 다양한 정보의 수신으로 인해, 시스템은 사용자의 다양한 특징을 식별할 수 있다. 예를 들어, 시스템은 가상 캐릭터가 취하는 동작이 어떤 것인지에 대한 반응을 나타내는 사용자의 얼굴 표정 또는 사용자의 어조를 식별할 수 있다. 또한, 수신된 정보는 사용자 주위의 환경도 포함한다.

[0102] 인사이트 대시보드는 사용자 수, 신규 사용자 수, 일일 평균 사용자 수 및 전체 상호 작용 시간과 같은 정보의 보고서를 디스플레이할 수 있다. 또한, 인사이트 대시보드는 활성 응시 시간, 음성 수다(speech verbosity), 구매 정보, 항목 선택 정보 및 감정 정보와 같은 인사이트 정보의 그래픽 표현을 포함할 수 있다.

[0103] 일부 실시예에서, 가상 캐릭터는 사용자 장치를 통해 클라이언트와 사용자 사이의 다양한 거래를 용이하게 할 수 있다. 예를 들어, 가상 캐릭터는 사용자가 아이템을 구매하기를 원한다는 것을 식별할 수 있으며, 가상 캐릭터는 사용자를 특정 웹페이지로 포워딩하거나 거래를 직접 촉진할 수 있다.

[0104] 인사이트는 가상 캐릭터와 상호 작용하는 사용자 데이터에 대한 감정 분석 및 기타 알고리즘을 실행할 수 있다. 예를 들어, 캐릭터는 가상으로 다른 캐릭터로부터 메시지를 수신하고 함께 영화를 만들기 위해 팀을 이루는 캐릭터에 대해 어떻게 느끼는지 사용자에게 물어볼 수 있다. 인사이트는 총 응답률을 식별할 수 있다(예를 들어,

인사이드는 청중의 78%가 이것이 좋은 생각이라는 것을 회사에 보여준다).

- [0105] 일부 실시예에서, 방법은 가상 캐릭터와 상호 작용하는 사용자와 연관된 장치의 그룹 각각에 가상 캐릭터를 생성하는 단계를 포함할 수 있다. 방법은 장치 그룹과 상호 작용하는 사용자에 관한 입력 정보를 수신하는 단계를 포함할 수 있다. 방법은 각각의 장치에 적어도 두 개의 내부 모델을 구현하는 단계를 포함할 수 있다. 내부 모델은 입력 정보의 특성을 식별하고 내부 모델에 의해 식별된 특성을 조합하여 내부 모델에 의해 식별된 특성들 사이의 유사도를 식별하도록 구성될 수 있다. 장치 그룹에서의 각각의 장치와 연관된 특성에 대해, 방법은 가상 캐릭터와 연관된 잠재적인 동작의 라이브러리를 검사하여 내부 모델에 의해 식별된 특성과 일치하는 동작을 결정하는 단계를 포함할 수 있다. 장치 그룹에서의 각각의 장치에 대해, 방법은 가상 캐릭터가 동작을 수행하도록 하는 명령어를 나타내는 내부 모델에 의해 식별된 특성과 일치하는 동작을 출력하는 단계를 포함할 수 있다.
- [0106] 방법은 가상 캐릭터와 장치 그룹과 연관된 사용자 사이의 상호 작용에 관한 정보를 획득하는 단계를 포함할 수 있다. 방법은 상호 작용과 관련하여 획득된 정보에 기초하여 가상 캐릭터와 사용자 사이의 상호 작용의 특징을 나타내는 보고서를 생성하는 단계를 포함할 수 있다.
- [0107] 일부 실시예에서, 보고서는 가상 캐릭터와 상호 작용하는 사용자의 수, 일일 평균 사용자의 수 및 사용자에게 의한 평균 상호 작용 시간을 포함한다.
- [0108] 일부 실시예에서, 방법은 가상 캐릭터와 상호 작용하는 사용자와 연관된 활성 응시 시간, 음성 수다, 항목 선택 정보 및 감정 정보 중 적어도 하나의 그래픽 표현을 생성하는 단계를 포함하고, 가상 표현은 보고서에 포함된다.
- [0109] 예를 들어, 도 8에 도시된 바와 같이, 인사이드 보고서는 전체 사용자 표시자(802), 지난 달에 대한 신규 사용자 표시자(804), 지난 24시간 동안의 신규 사용자 표시자(806), 일일 평균 사용자 표시자(808), 월별 평균 사용자 표시자(810), 전체 상호 작용 시간 표시자(812), 시간 대 경험 세그먼트를 나타내는 평균 응시 시간 그래프(814), 시간 대 경험 세그먼트를 나타내는 음성 수다 그래프(816), 가상 캐릭터와 상호 작용하는 것에 응답하여 사용자에게 의해 취해지는 후속 동작(예를 들어, 구매한 영화 티켓(818), 아이템 선택(820)), 가상 캐릭터 감정(822) 등 중 임의의 것을 포함할 수 있다.
- [0110] 가상 캐릭터 레이어
- [0111] 본 실시예는 가상 캐릭터가 많은 애플리케이션에서 구현되는 것을 허용할 수 있다. 예를 들어, 가상 캐릭터는 매장에서 고객이 제품을 찾는 데 도움을 주거나, 철물점 고객이 새 페인트를 선택하는 데 도움을 주거나, 터치스크린 기반의 현금 자동 인출기를 더욱 개인화되고 친근한 얼굴을 갖는 것으로 교체할 수 있다. 일부 실시예에서, 현실 세계에서의 가상 캐릭터의 위치는 매핑 애플리케이션에 디스플레이될 수 있다.
- [0112] 일부 실시예에서, 인사이드는 데이터를 기계 학습 데이터 세트로 다시 공급하고 결과를 자동으로 개선하는 개선 시스템을 포함할 수 있다. 예를 들어, 가상 캐릭터가 백만 명에게 셀카를 찍어달라고 요청하면 어떻게 될까? 시스템은 이러한 모든 사진을 사용하여 셀카 카메라 사진으로부터 사용자의 얼굴을 검출하기 위한 새로운 알고리즘을 만들 수 있다.
- [0113] 도 9는, 다양한 실시예에 따라, 다중 모드 모델을 사용하여 가상 캐릭터를 제어하는 방법의 블록도를 도시한다. 방법은 음성 정보, 얼굴 표정 정보, 환경 정보 등과 같은 입력 정보를 수신하는 단계(블록 902)를 포함할 수 있다. 수신된 정보는 다중 모드(예를 들어, 오디오, 비주얼)일 수 있다.
- [0114] 방법은 환경 내의 위치에 가상 캐릭터를 디스플레이하는 단계(블록 904)를 포함할 수 있다. 가상 캐릭터는 센서/카메라에 의해 캡처된 환경과 같은 현실 세계 환경에 오버레이되거나 추가될 수 있다. 일부 실시예에서, 가상 캐릭터는 환경에 디스플레이되고 있을 때 제1 동작을 출력할 수 있다. 제1 동작은, 예를 들어, 사용자에게 대한 질의를 포함할 수 있다.
- [0115] 방법은 입력 정보를 처리하기 위해 내부 모델을 구현하는 단계(블록 906)를 포함할 수 있다. 내부 모델은 사용자가 말한 단어, 사용자의 얼굴 표정 등과 같은 사용자에게 대한 다양한 정보를 판단하기 위해 입력 정보를 처리할 수 있다.
- [0116] 방법은 적어도 2개의 내부 모델에 의해 식별된 특성을 검사하여 제1 식별된 특성이 제2 식별된 특성에 대한 임계 유사도 내에 있는지 여부를 판단하는 단계(블록 908)를 포함할 수 있다. 특성 사이의 유사성을 식별하는 것은 가상 캐릭터와 일반적으로 연관된 특성과 일치하는 특성을 결정하기 위해 내부 모델을 참조하는 것을 포함할

수 있다.

- [0117] 일례로서, 제1 식별된 특성은 사용자가 말한 단어가 "나는 슬프다(I am sad)"라는 것을 식별하는 것을 포함할 수 있다. 제2 식별된 특성은 사용자가 웃고 있거나 행복하다는 것을 나타내는 얼굴 표정을 식별하는 것을 포함할 수 있다. 따라서, 이러한 특성들을 비교하는 것은 특성이 임계 유사도 내에 있지 않다고 결정할 수 있다. 아래에 언급되는 바와 같이, 지식 모델은 대조되는 특성이 더 정확할 가능성이 있는지 식별하는 데 사용될 수 있다. 예를 들어, 지식 모델은 가상 캐릭터에 의한 이전 동작이 사용자로부터 긍정적인 반응을 유도할 가능성이 가장 높다는 것을 식별할 수 있다. 이 예에서, 사용자가 "슬프다(I am sad)" 대신 "기쁘다(I am glad)"라고 말했다기 때문에, 사용자가 웃고 있는 것을 나타내는 특성이 더 정확할 수 있고, 제1 특성은 부정확하다.
- [0118] 방법은 가상 캐릭터 지식 모델 내에 포함된 가상 캐릭터에 특정한 정보에 대하여 제1 식별된 특성을 제2 식별된 특성과 비교하여 선택된 특징을 식별하는 단계(블록 910)를 포함할 수 있다. 지식 모델은 사용자에 의한 일반적인 응답, 가상 캐릭터와 관련된 용어, 가상 캐릭터와 종종 관련된 기타 캐릭터 등과 같은 가상 캐릭터에 특정한 정보를 포함할 수 있다.
- [0119] 방법은 가상 캐릭터에 의한 잠재적 동작의 라이브러리에 액세스하는 단계(블록 912)를 포함할 수 있다. 라이브러리는 특정 가상 캐릭터에 의한 잠재적인 애니메이션/음성을 나타내는 정보를 포함하는 데이터베이스, 레지스트리, 목록 등을 포함할 수 있다. 또한, 잠재적인 동작은 생성 모델을 사용함으로써 런타임에 생성될 수 있다.
- [0120] 방법은 처리된 입력 정보에 기초하여 가상 캐릭터에 의해 수행될 동작을 결정하는 단계를 포함할 수 있다. 다중 모드 모델은 사용자와 관련된 다양한 정보를 처리하고 다양한 정보를 조합하여 더욱 정확하게 동작을 결정할 수 있다. 동작을 결정하는 것은 처리된 입력 정보로부터 식별된 정확한 동작과 가장 근접하게 일치하는 라이브러리 내의 동작을 식별하는 것을 포함할 수 있다.
- [0121] 방법은 환경에서 동작을 수행하는 가상 캐릭터를 디스플레이하는 단계(블록 914)를 포함할 수 있다. 일부 실시예들에서, 방법은 사용자 장치에 의한 선택을 식별하는 것에 기초하여 가상 캐릭터를 동적으로 생성하는 단계(블록 916)를 포함할 수 있다. 동적으로 생성하는 단계는 짧은 지연(예를 들어, 1초 미만)으로 가상 캐릭터를 제어하는 단계를 포함할 수 있다. 사용자 장치에 의한 선택은 사용자 장치와 상호 작용하는 사용자에 의한 하이퍼링크에서의 클릭을 포함할 수 있다.
- [0122] 방법은 클라이언트에 대한 인사이트 보고서를 생성하는 단계(블록 918)를 포함할 수 있다. 인사이트 보고서는 사용자 장치에서 생성된 가상 캐릭터와 상호 작용하는 사용자에 대한 다양한 정보를 포함할 수 있다.
- [0123] 제1 예시적인 실시예에서, 가상 캐릭터를 제어하기 위한 방법이 개시된다. 방법은 장치로부터 다중 모드 입력 정보를 수신하는 단계를 포함할 수 있으며, 다중 모드 입력 정보는 음성 정보, 얼굴 표정 정보 및 장치를 둘러싸는 환경을 나타내는 환경 정보 중 임의의 하나를 포함한다. 또한, 방법은 장치 상에 제시된 디스플레이 환경 내의 위치에 가상 캐릭터를 디스플레이하는 단계를 포함할 수 있다. 또한, 방법은 적어도 2개의 내부 모델을 구현하여 다중 모드 입력 정보의 특성을 식별하는 단계를 포함할 수 있다.
- [0124] 또한, 방법은 식별된 특성 중 제1 식별된 특성이 식별된 특성 중 제2 식별된 특성의 유사한 특징을 임계 개수 포함하는지 여부를 결정하기 위해 적어도 2개의 내부 모델의 식별된 특성을 검사하는 단계를 포함할 수 있다. 또한, 방법은 제1 식별된 특성 및 제2 식별된 특성을 가상 캐릭터 지식 모델에 포함된 가상 캐릭터에 특정한 정보에 대하여 비교하여, 제1 식별된 특성이 식별된 특성 중 제2 식별된 특성의 유사한 특징을 임계 개수 포함한다고 판단한 것에 기초하여 선택 특성을 선택하는 단계를 포함할 수 있다.
- [0125] 또한, 방법은 가상 캐릭터와 연관된 잠재적 동작의 라이브러리에 액세스하여 선택 특성과 일치하는 동작을 결정하는 단계를 포함할 수 있고, 동작은 가상 캐릭터에 의해 수행될 애니메이션 및 연관 오디오 모두를 포함한다. 또한, 방법은 장치에 제시된 환경에서 가상 캐릭터를 수정하고 연관 오디오를 출력함으로써 결정된 동작을 구현하는 단계를 포함할 수 있다.
- [0126] 일부 실시예에서, 적어도 2개의 내부 모델은 음성 정보로부터 음성 감정을 파싱할 수 있는 음성 인식 모델 및 얼굴 표정 정보에 기초하여 얼굴 특징 감정을 검출할 수 있는 얼굴 특징 인식 모델을 포함하고, 선택 특성은 음성 감정과 얼굴 특징 감정 사이에서 공통적인 감정이고, 결정된 동작은 감정에 기초하여 결정된다.
- [0127] 일부 실시예에서, 적어도 2개의 내부 모델은 사용자와 이전에 맺은 관계에 관한 정보를 포함하는 사전 지식 정보를 검색할 수 있는 사전 지식 모델을 포함하고, 선택 특성은 사전 지식 모델을 사용하여 처리된 사전 지식 정보에 기초하여 선택된다.

- [0128] 일부 실시예에서, 내부 모델은 오디오 정보로부터 컨텍스트와 의미를 도출하도록 구성된 자연어 이해 모델, 환경 정보를 식별하도록 구성된 인식 모델 및 사용자 및 다른 가상 캐릭터와 관련된 데이터를 식별하도록 구성된 소셜 시뮬레이션 모델을 포함한다.
- [0129] 일부 실시예에서, 방법은 가상 캐릭터에게 장치에서 사용자에게 대한 질의를 나타내는 초기 동작을 수행하도록 지시하는 단계를 포함할 수 있고, 입력 정보는 질의에 대한 사용자에게 의한 응답을 나타낸다.
- [0130] 일부 실시예에서, 방법은 네트워크를 통해 복수의 사용자에게 임베디드 링크를 공유하는 단계를 포함할 수 있다. 또한, 방법은 임베디드 링크가 선택되었다는 것을 나타내는 선택을 장치 세트 중 임의의 장치로부터 수신하는 단계를 포함할 수 있다. 또한, 방법은 선택을 수신하는 것에 응답하여, 장치에 가상 캐릭터를 디스플레이 하기 위해 선택을 송신한 장치 세트 중의 장치에 데이터 스트림을 송신하는 단계를 포함할 수 있다.
- [0131] 일부 실시예에서, 방법은 데이터 스트림의 제1 배치(batch)를 제1 시간에 전송하는 단계를 포함할 수 있고, 제1 배치는 장치의 디스플레이 상에 가상 캐릭터를 초기에 생성하기 위한 정보를 포함한다. 또한, 방법은 제1 시간 이후의 제2 시간에 데이터 스트림의 제2 배치를 전송하는 단계를 포함할 수 있고, 제2 배치는 가상 캐릭터에 의해 제1 동작을 출력하기 위한 정보를 포함하고, 제1 배치는 제2 시간에 폐기된다.
- [0132] 일부 실시예에서, 방법은 환경의 바닥(floor)을 나타내는 환경의 부분을 식별하기 위하여 환경 정보를 검사하는 단계를 포함할 수 있다. 또한, 방법은 환경의 바닥을 나타내는 환경의 부분 위의 제1 위치에 가상 캐릭터를 위치 설정하는 단계를 포함할 수 있다.
- [0133] 제2 예시적인 실시예에서, 장치에 의해 캡처된 사용자에게 관한 다중 모드 입력에 대한 응답을 제공하도록 구성된 장치가 제공된다. 장치는 적어도 하나의 메모리를 포함할 수 있다. 적어도 하나의 메모리는 다중 모드 입력 정보로부터 특성을 식별하도록 구성된 적어도 2개의 내부 모델을 포함할 수 있다. 또한, 적어도 하나의 메모리는 가상 캐릭터에 특정한 정보를 포함하는 가상 캐릭터 지식 모델을 포함할 수 있다. 또한, 적어도 하나의 메모리는 가상 캐릭터와 연관된 잠재적인 동작의 라이브러리를 포함할 수 있으며, 각각의 동작은 가상 캐릭터에 의해 수행될 애니메이션 및 연관 오디오와 연관된다.
- [0134] 또한, 장치는 적어도 하나의 프로세서를 포함할 수 있다. 적어도 하나의 프로세서는 음성 정보, 얼굴 표정 정보 및 환경을 나타내는 환경 정보 중 적어도 하나를 포함하는 다중 모드 입력 정보를 수신하도록 구성될 수 있다. 또한, 적어도 하나의 프로세서는 제1 식별된 특성이 제2 식별된 특성과 임계 유사도 내에 있는지 여부를 결정하기 위해 적어도 2개의 내부 모델에 의해 식별된 특성을 검사하도록 구성될 수 있다. 또한, 적어도 하나의 프로세서는 가상 캐릭터 지식 모델에 대해 제1 식별된 특성 및 제2 식별된 특성을 비교하여 선택 특성을 식별하도록 구성될 수 있다.
- [0135] 또한, 적어도 하나의 프로세서는 가상 캐릭터와 관련된 잠재적인 동작의 라이브러리를 검사함으로써 선택 특성과 일치하는 동작을 결정하도록 구성될 수 있으며, 동작은 장치에 출력될 오디오를 포함한다. 또한, 적어도 하나의 프로세서는 장치에 오디오를 출력하도록 구성될 수 있다.
- [0136] 일부 실시예에서, 적어도 하나의 프로세서는 환경 정보로부터 도출된 환경 내의 위치에서 장치의 디스플레이 상에 가상 캐릭터를 디스플레이하도록 더 구성된다. 또한, 적어도 하나의 프로세서는 장치에 제시되는 환경 내에 가상 캐릭터를 수정함으로써 장치에 출력될 오디오와 가상 캐릭터에 의해 수행될 선택된 애니메이션을 모두 포함하는 동작을 구현하도록 구성될 수 있다.
- [0137] 일부 실시예에서, 적어도 2개의 내부 모델은 사용자와 이전에 맺은 관계에 관한 정보를 포함하는 사전 지식 정보를 검색할 수 있는 사전 지식 모델을 포함하고, 선택 특성은 사전 지식 모델을 사용하여 처리된 사전 지식 정보에 기초하여 선택된다.
- [0138] 일부 실시예에서, 적어도 2개의 내부 모델은, 음성 정보로부터 음성 감정을 파싱할 수 있는 음성 인식 모델 및 얼굴 표정 정보에 기초하여 얼굴 특징 감정을 검출할 수 있는 얼굴 특징 인식 모델을 포함하고, 선택 특성은 음성 감정과 얼굴 특징 감정 사이에 공통적인 감정이고, 결정된 동작은 감정에 기초하여 결정된다.
- [0139] 제3 예시적인 실시예에서, 사용자 장치의 웹 브라우저 상에서 가상 캐릭터를 동적으로 생성하기 위한 컴퓨터 구현 방법이 제공된다. 컴퓨터 구현 방법은 사용자 장치의 웹 브라우저에 대한 링크를 임베딩하는 단계를 포함할 수 있고, 링크는 사용자 장치에서 실행되는 애플리케이션으로 웹 브라우저를 연결한다. 또한, 컴퓨터 구현 방법은 링크가 선택되었다는 표시를 사용자 장치로부터 수신하는 단계를 포함할 수 있다. 또한, 컴퓨터 구현 방법은 가상 캐릭터를 생성하기 위하여 정보를 나타내는 데이터 스트림을 애플리케이션으로부터 웹 브라우저로 전송하

는 단계를 포함할 수 있다.

- [0140] 또한, 컴퓨터 구현 방법은 가상 캐릭터를 사용자 장치의 웹 브라우저에 디스플레이하는 단계를 포함할 수 있다. 또한, 컴퓨터 구현 방법은 사용자 장치로부터 다중 모드 입력 정보를 수신하는 단계를 포함할 수 있고, 다중 모드 입력 정보는 음성 정보, 얼굴 표정 정보 및 환경을 나타내는 환경 정보를 포함한다. 또한, 컴퓨터 구현 방법은 적어도 2개의 내부 모델을 구현하여 다중 모드 입력 정보의 특성을 식별하기 위한 단계를 포함할 수 있다.
- [0141] 또한, 컴퓨터 구현 방법은 제1 식별된 특성이 제2 식별된 특성에 대한 임계 유사도 내에 있는지 여부를 결정하기 위해 적어도 2개의 내부 모델에 의해 식별된 특성을 검사하는 단계를 포함할 수 있다. 또한, 컴퓨터 구현 방법은 제1 식별된 특성 및 제2 식별된 특성을 가상 캐릭터 지식 모델에 포함된 가상 캐릭터에 특정한 정보에 대하여 비교하여, 제1 식별된 특성이 제2 식별된 특성의 유사한 특징을 임계 개수 포함한다고 판단한 것에 기초하여 선택 특성을 선택하는 단계를 포함할 수 있다.
- [0142] 또한, 컴퓨터 구현 방법은 가상 캐릭터와 연관된 잠재적 동작의 라이브러리에 액세스하여 선택 특성과 일치하는 동작을 선택하는 단계를 포함할 수 있으며, 동작은 가상 캐릭터에 의해 수행될 애니메이션 및 연관 오디오 모두를 포함한다. 또한, 컴퓨터 구현 방법은 동작을 수행하는 환경에서 가상 캐릭터를 디스플레이하고 연관 오디오를 출력하는 단계를 포함할 수 있다.
- [0143] 일부 실시예에서, 웹 브라우저는 사용자 장치에서 실행되는 모바일 애플리케이션에 디스플레이되는 페이지를 포함한다.
- [0144] 일부 실시예에서, 컴퓨터 구현 방법은 선택 특성 및 동작에 관한 정보를 저장하는 단계를 포함할 수 있다. 또한, 컴퓨터 구현 방법은 복수의 사용자에게 대한 일련의 선택 특성 및 동작을 집계하는 단계를 포함할 수 있다. 또한, 컴퓨터 구현 방법은 복수의 사용자에게 대한 일련의 선택 특성 및 동작을 처리하여 복수의 사용자와의 가상 캐릭터의 관계에 관한 분석 세트를 도출하는 단계를 포함할 수 있다. 또한, 컴퓨터 구현 방법은 복수의 사용자와의 가상 캐릭터의 관계에 관한 분석 세트를 디스플레이하기 위해 분석 대시보드를 제시하는 단계를 포함할 수 있다.
- [0145] 프로세싱 시스템의 예
- [0146] 도 10은 여기에 설명된 적어도 일부 동작들이 구현될 수 있는 프로세싱 시스템의 일례를 예시하는 블록도이다. 예를 들어, 프로세싱 시스템(1000)의 일부 컴포넌트는 본 실시예에서 설명된 바와 같이 전자 장치에서 호스팅될 수 있다.
- [0147] 프로세싱 시스템(1000)은 하나 이상의 중앙 처리 유닛("프로세서")(1002), 메인 메모리(1006), 비휘발성 메모리(1010), 네트워크 어댑터(1012)(예를 들어, 네트워크 인터페이스), 비디오 디스플레이(1018), 입/출력 장치(1020), 제어 장치(1022)(예를 들어, 키보드 및 포인팅 장치), 저장 매체(1026)를 포함하는 드라이브 유닛(1024) 및 신호 생성 장치(1030)를 포함할 수 있고, 이들은 버스(1016)에 통신 가능하게 연결된다. 버스(1016)는 적절한 브리지, 어댑터 또는 컨트롤러로 연결된 물리적 버스 및/또는 포인트 투 포인트 연결을 나타내는 추상적인 것으로 도시된다. 따라서, 버스(1016)는 시스템 버스, PCI(Peripheral Component Interconnect) 버스 또는 PCI-Express 버스, HyperTransport 또는 ISA(Industry Standard Architecture) 버스, SCSI(Small Computer System Interface) 버스, USB(universal serial bus), IIC(I2C) 버스 또는 IEEE(Institute of Electrical and Electronics Engineers) 표준 1594 버스(예를 들어, "Firewire")를 포함할 수 있다.
- [0148] 프로세싱 시스템(1000)은 데스크탑 컴퓨터, 태블릿 컴퓨터, PDA(personal digital assistant), 스마트폰, 게임 콘솔, 음악 플레이어, 웨어러블 전자 장치(예를 들어, 시계 또는 피트니스 트래커), 네트워크 연결("스마트") 장치(예를 들어, 텔레비전 또는 홈 어시스턴트 장치), 가상/증강 현실 시스템(예를 들어, 헤드 마운트 디스플레이) 또는 프로세싱 시스템(1000)에 의해 취해질 동작(들)을 특정하는 명령어 세트(순차적 또는 기타)를 실행할 수 있는 다른 전자 장치와 유사한 컴퓨터 프로세서 아키텍처를 공유할 수 있다.
- [0149] 메인 메모리(1006), 비휘발성 메모리(1010) 및 저장 매체(1026)("기계 판독 가능한 매체"라고도 함)가 단일 매체인 것으로 도시되어 있지만, "기계 판독 가능한 매체" 및 "저장 매체"는 하나 이상의 명령어(1028) 세트를 저장하는 단일 매체 또는 다수의 매체(예를 들어, 중앙 집중식/분산 데이터베이스 및/또는 연관 캐시 및 서버)를 포함하는 것으로 간주되어야 한다. 또한, "기계 판독 가능한 매체" 및 "저장 매체"는 프로세싱 시스템(1000)에 의한 실행을 위한 명령어 세트를 저장, 인코딩 또는 반송할 수 있는 임의의 매체를 포함하는 것으로 간주되어야 한다.

- [0150] 일반적으로, 본 개시 내용의 실시예를 구현하기 위해 실행되는 루틴은 운영 체제 또는 특정 애플리케이션, 컴포넌트, 프로그램, 객체, 모듈 또는 명령어 시퀀스(집합적으로 "컴퓨터 프로그램"으로 지칭됨)의 일부로서 구현될 수 있다. 컴퓨터 프로그램은 통상적으로 컴퓨팅 장치에서 다양한 메모리 및 저장 장치에 다양한 시간에 설정된 하나 이상의 명령어(예를 들어, 명령(1004, 1008, 1028))를 포함한다. 하나 이상의 프로세서(1002)에 의해 판독 및 실행될 때, 명령어(들)는 프로세싱 시스템(1000)이 본 개시 내용의 다양한 양태를 포함하는 요소를 실행하기 위한 동작을 수행하게 한다.
- [0151] 더욱이, 실시예가 완전히 기능하는 컴퓨팅 장치의 맥락에서 설명되었지만, 본 발명에 속하는 기술 분야에서 통상의 기술을 가진 자는 다양한 실시예가 다양한 형태의 프로그램 제품으로서 배포될 수 있다는 것을 이해할 것이다. 본 개시 내용은 실제로 배포에 영향을 미치지 위해 사용된 특정 유형의 기계 또는 컴퓨터 판독 가능한 매체에 관계없이 적용된다.
- [0152] 기계 판독 가능한 저장 매체, 기계 판독 가능한 매체 또는 컴퓨터 판독 가능한 매체의 추가 예는 휘발성 및 비휘발성 메모리 장치(1010), 플로피 및 기타 이동식 디스크, 하드 디스크 드라이브, 광 디스크(예를 들어, CD-ROM(Compact Disk Read-Only Memory), DVD(Digital Versatile Disk))와 같은 기록 가능한 유형의 매체와, 디지털 및 아날로그 통신 링크와 같은 전송 유형의 매체를 포함한다.
- [0153] 네트워크 어댑터(1012)는 프로세싱 시스템(1000)이 프로세싱 시스템(1000) 및 외부 엔티티에 의해 지원되는 임의의 통신 프로토콜을 통해 프로세싱 시스템(1000) 외부에 있는 엔티티와 네트워크(1014)에서 데이터를 중재할 수 있게 한다. 네트워크 어댑터(1012)는 네트워크 어댑터 카드, 무선 네트워크 인터페이스 카드, 라우터, 액세스 포인트, 무선 라우터, 스위치, 다계층 스위치, 프로토콜 변환기, 게이트웨이, 브리지, 브리지 라우터, 허브, 디지털 미디어 수신기 및/또는 리피터를 포함할 수 있다.
- [0154] 네트워크 어댑터(1012)는 컴퓨터 네트워크에서 데이터를 액세스/프록시하기 위한 허가를 통제 및/또는 관리하고 서로 다른 기계 및/또는 애플리케이션 사이의 다양한 신뢰 수준을 추적하는 방화벽을 포함할 수 있다. 방화벽은 특정 세트의 기계와 애플리케이션, 기계와 기계 그리고/또는 애플리케이션과 애플리케이션 사이의 미리 정해진 액세스 권한 세트를 강화할 수 있는(예를 들어, 이러한 엔티티 사이의 트래픽 및 리소스 공유 흐름을 규제할 수 있는) 하드웨어 및/또는 소프트웨어의 임의의 조합을 갖는 임의의 개수의 모듈일 수 있다. 추가로, 방화벽은 개인, 기계 및/또는 애플리케이션에 의한 객체의 액세스 및 작동 권한을 포함하는 권한과, 허가 권한이 처해 있는 상황을 자세히 설명하는 액세스 제어 목록을 추가로 관리 및/또는 액세스할 수 있다.
- [0155] 여기에 소개된 기술은 프로그래밍 가능한 회로(예를 들어, 하나 이상의 마이크로프로세서), 소프트웨어 및/또는 펌웨어, 특수 목적의 하드와이어드(즉, 프로그래밍할 수 없는) 회로 또는 이러한 형태의 조합에 의해 구현될 수 있다. 특수 목적 회로는 하나 이상의 ASIC(Application-Specific Integrated Circuit), PLD(Programmable Logic Device), FPGA(Field-Programmable Gate Array) 등의 형태일 수 있다.
- [0156] 여기에 개시된 기술은 제품 제조 프로세스의 능력을 보완하기 위해 설계 도면으로부터 정보를 추출하는 도구 또는 시스템을 포함할 수 있다. 본 실시예는 건적/추정의 생성/수정, 제조 피드백 생성, 제조 일정/주문 프로세스 등을 보조할 수 있다.
- [0157] 비고
- [0158] 문맥상 명백히 달리 요구하지 않는 한, 설명 및 청구범위 전체에 걸쳐 "포함하는(comprise)", "포함하는(comprising)" 등의 단어는 배타적이거나 소진적인 의미가 아니라 포괄적인 의미로 해석되어야 한다; 즉, "포함하지만 이에 국한되지 않는"이라는 의미로. 여기에 사용된 바와 같이, "연결된(connected)", "결합된(coupled)" 또는 이들의 임의의 변형은 2개 이상의 요소 사이의 직접적 또는 간접적인 임의의 연결 또는 결합을 의미한다; 요소 사이의 결합 또는 연결은 물리적, 논리적 또는 이들의 조합일 수 있다. 또한, "여기에(herein)", "위에(above)", "아래에(below)"라는 단어와 이와 유사한 의미를 지닌 단어는, 본 출원에서 사용될 때, 본 출원의 임의의 특정 부분이 아닌 전체로서의 본 출원을 의미한다. 문맥이 허용하는 경우, 단수 또는 복수를 사용하는 위의 상세한 설명에서의 단어는 또한 복수 또는 단수를 각각 포함할 수 있다. 2 이상의 항목 목록과 관련하여 "또는(or)"이라는 단어는 다음의 단어에 대한 해석을 모두 포함한다: 목록 내의 임의의 항목, 목록 내의 모든 항목 및 목록 내의 항목의 임의의 조합.
- [0159] 여기에 사용된 바와 같이, "실질적으로(substantially)"라는 용어는 동작, 특성, 속성, 상태, 구조, 항목 또는 결과의 완전하거나 거의 완전한 범위 또는 정도를 지칭한다. 예를 들어, "실질적으로" 둘러싸여 있는 객체는 객체가 완전히 둘러싸여 있거나 또는 거의 완전히 둘러싸여 있다는 것을 의미한다. 절대 완전성으로부터 허용되는

정확한 허용 가능한 정도는 일부 경우에 특정 상황에 따라 달라질 수 있다. 그러나, 일반적으로 말해서, 완료에 가까워진다는 것은 절대적이고 완전한 완료를 얻은 것과 동일한 전체 결과를 얻도록 하는 것일 것이다. "실질적으로"의 사용은 동작, 특성, 속성, 상태, 구조, 항목 또는 결과의 완전하거나 거의 완전한 부족을 나타내는 부정적인 의미로 사용될 때 동일하게 적용된다.

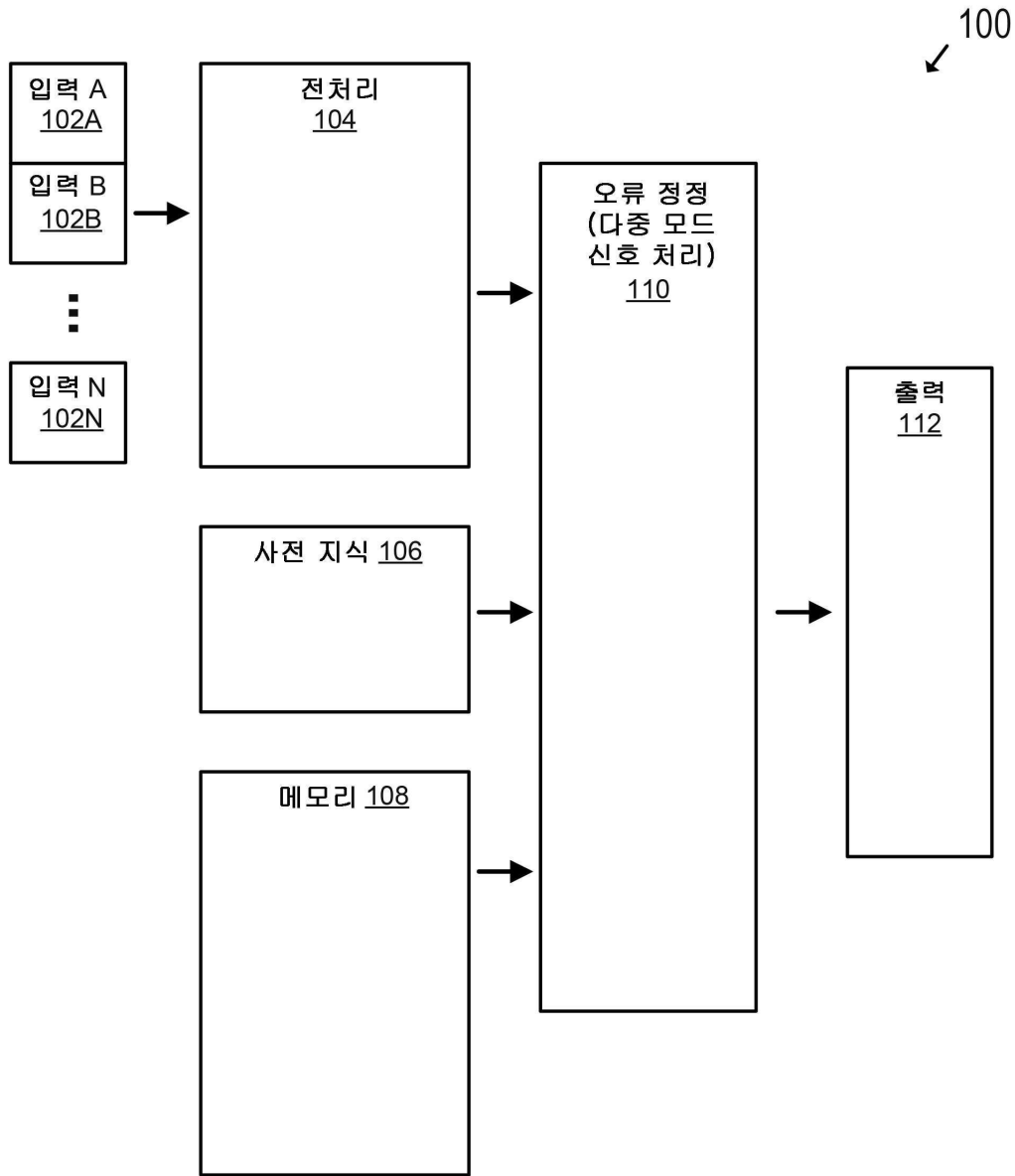
[0160] 본 발명의 실시예에 대한 상기 상세한 설명은 소진적인 것이거나 본 발명을 위에 개시된 정확한 형태로 제한하려고 의도되지 않는다. 본 발명에 대한 특정 예가 예시의 목적으로 위에서 설명되었지만, 관련 기술 분야에서의 통상의 기술자가 인식하는 바와 같이, 본 발명의 범위 내에서 다양한 균등한 변형이 가능하다. 예를 들어, 프로세스 또는 블록이 주어진 순서로 제시되지만, 대안적인 구현은 다른 순서로 단계를 갖는 루틴을 수행하거나 블록을 갖는 시스템 사용할 수 있고, 일부 프로세스 또는 블록은 삭제, 이동, 추가, 세분화, 조합 및/또는 수정되어 대안 또는 하위 조합을 제공할 수 있다. 이러한 프로세스 또는 블록의 각각은 다양한 방식으로 구현될 수 있다. 또한, 프로세스 또는 블록이 직렬로 수행되는 것으로 도시되는 경우가 있지만, 이러한 프로세스 또는 블록은 대신에 병렬로 수행 또는 구현되거나 다른 시간에 수행될 수 있다. 또한, 여기에 언급된 임의의 특정 숫자는 단지 예일뿐이다: 대안적인 구현은 다른 값 또는 범위를 사용할 수 있다.

[0161] 여기에 제공된 본 발명의 교시는 반드시 위에서 설명된 시스템이 아닌 다른 시스템에 적용될 수 있다. 위에서 설명된 다양한 예의 요소 및 동작은 본 발명의 추가 구현을 제공하기 위해 조합될 수 있다. 본 발명의 일부 대안적인 구현은 위에서 언급한 구현에 대한 추가 요소를 포함할 수 있을 뿐만 아니라 더 적은 요소를 포함할 수도 있다.

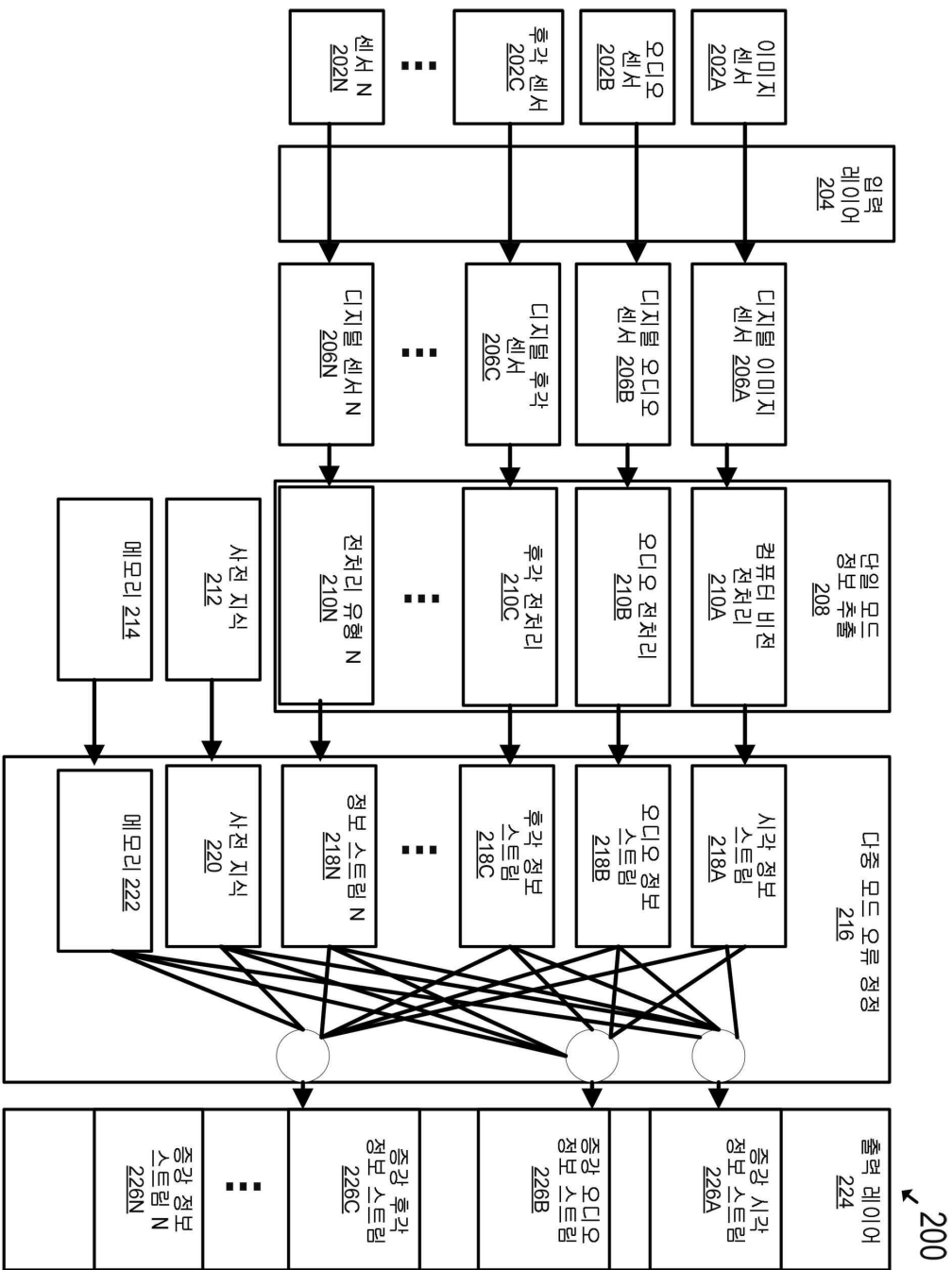
[0162] 상기 상세한 설명에 비추어 본 발명에 이러한 변경 및 기타 변경이 이루어질 수 있다. 위의 설명은 본 발명의 특정 예를 설명하고 고려되는 최상의 모드를 설명하지만, 위의 내용이 텍스트에 얼마나 상세하게 나타나 있는지에 관계없이, 본 발명은 다양한 방식으로 실시될 수 있다. 시스템의 세부사항은 그 특정 구현에서 상당히 다를 수 있지만 여전히 여기에 개시된 발명에 포함된다. 위에서 언급된 바와 같이, 본 발명의 특정 특징 또는 양태를 설명할 때 사용되는 특정 용어는 그 용어가 연관된 본 발명의 임의의 특정 특성, 특징 또는 양태로 제한되도록 용어가 여기에서 재정의되고 있다는 것을 의미하는 것으로 간주되어서는 안 된다. 일반적으로, 위의 상세한 설명 섹션이 그러한 용어를 명시적으로 정의하지 않는 한, 이어지는 청구범위에 사용된 용어는 본 발명을 명세서에 개시된 특정 예로 제한하는 것으로 해석되어서는 안 된다. 따라서, 본 발명의 실제 범위는 개시된 예뿐만 아니라 청구범위 하에 본 발명을 실시하거나 구현하는 모든 균등한 방식을 포함한다.

도면

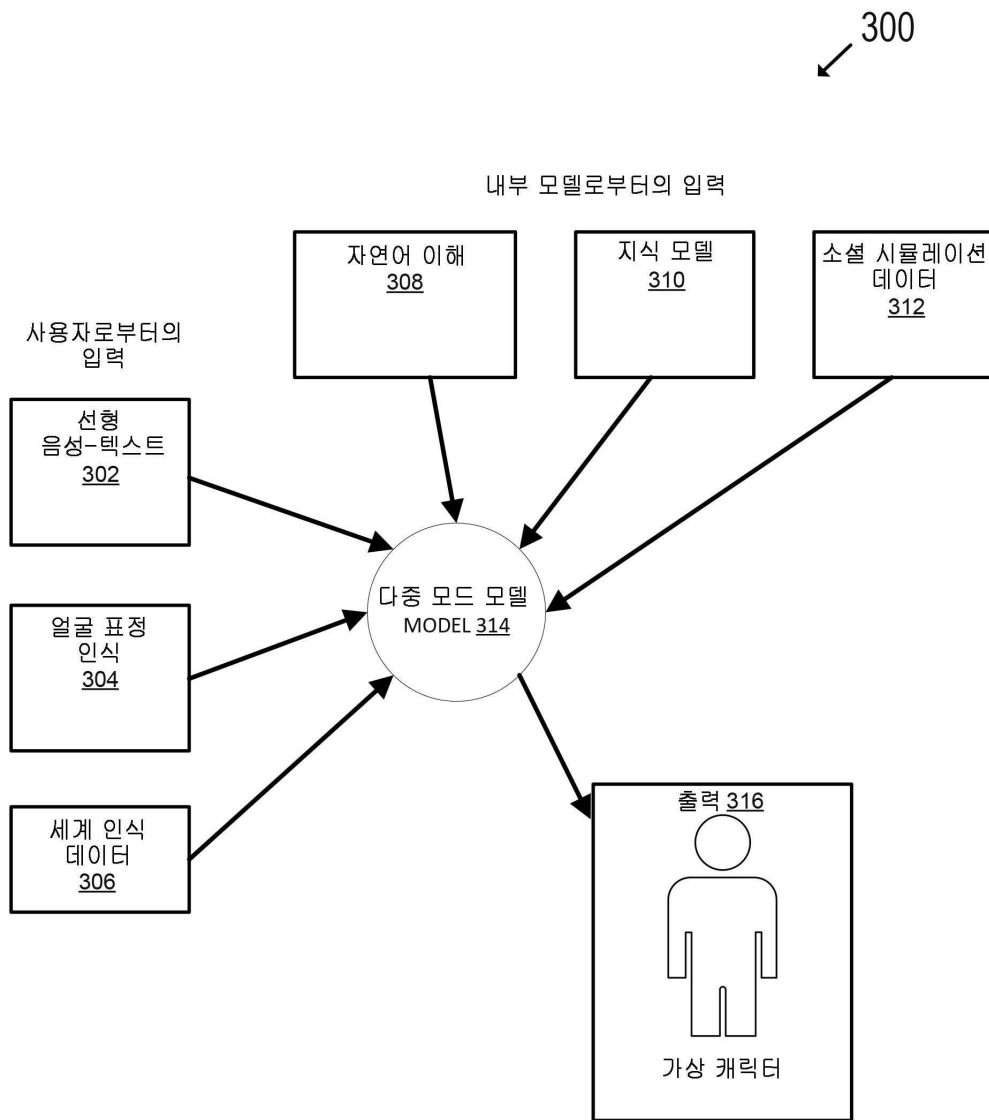
도면1



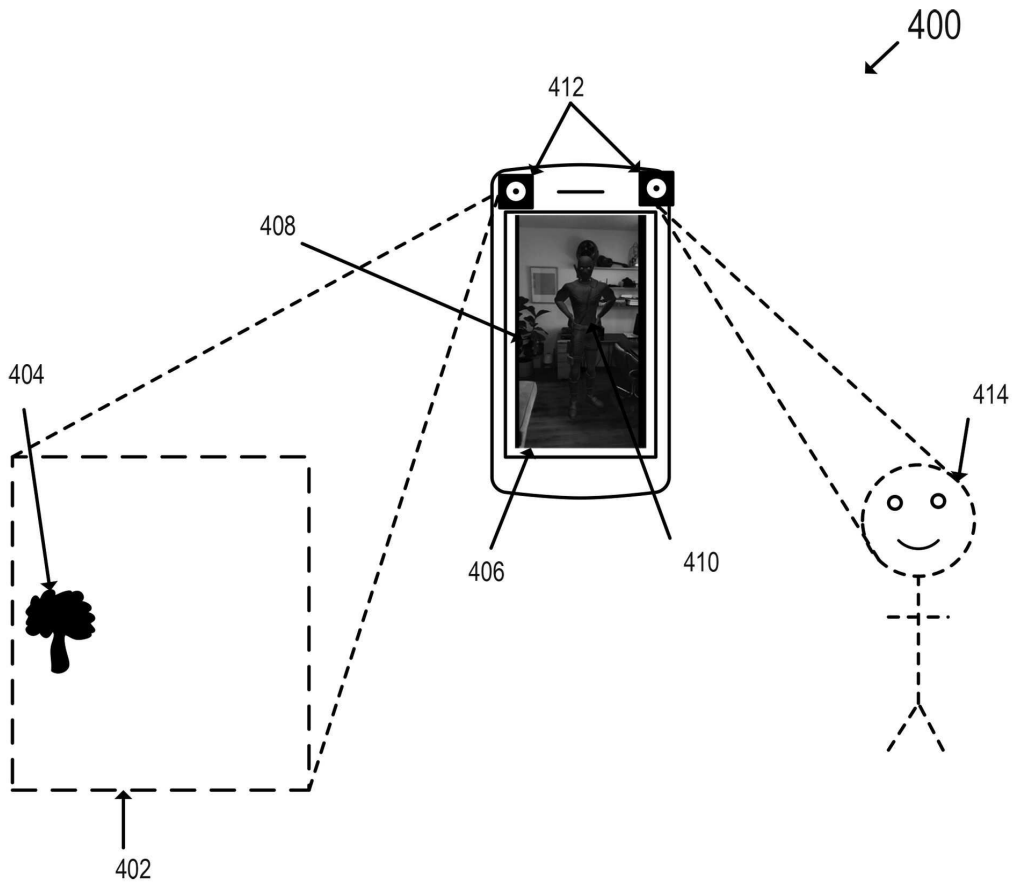
도면2



도면3



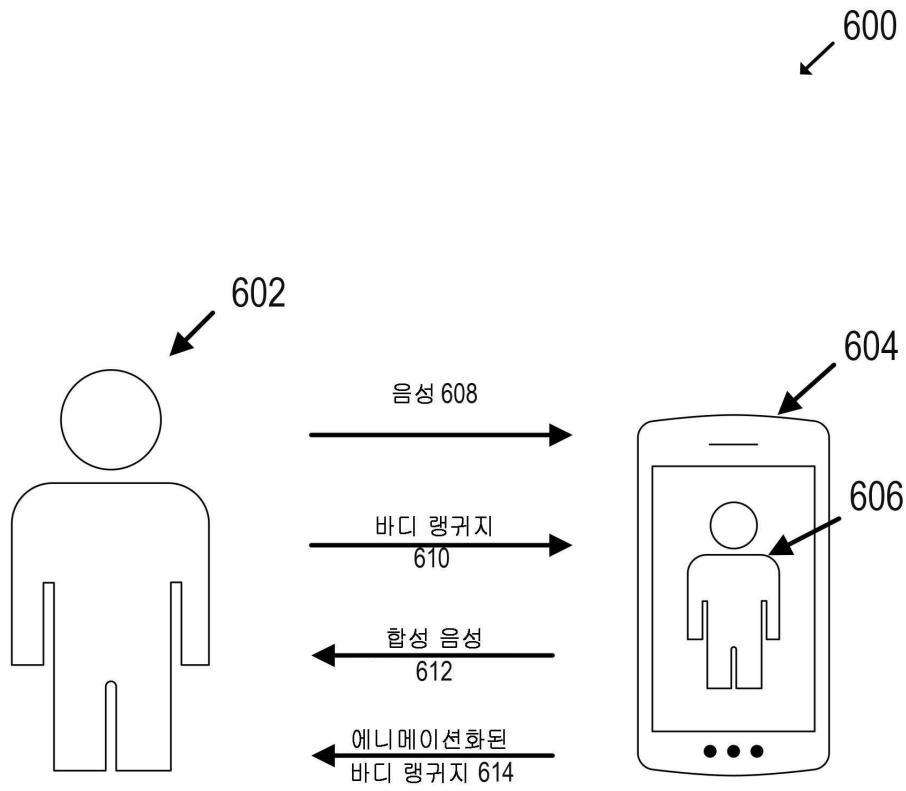
도면4



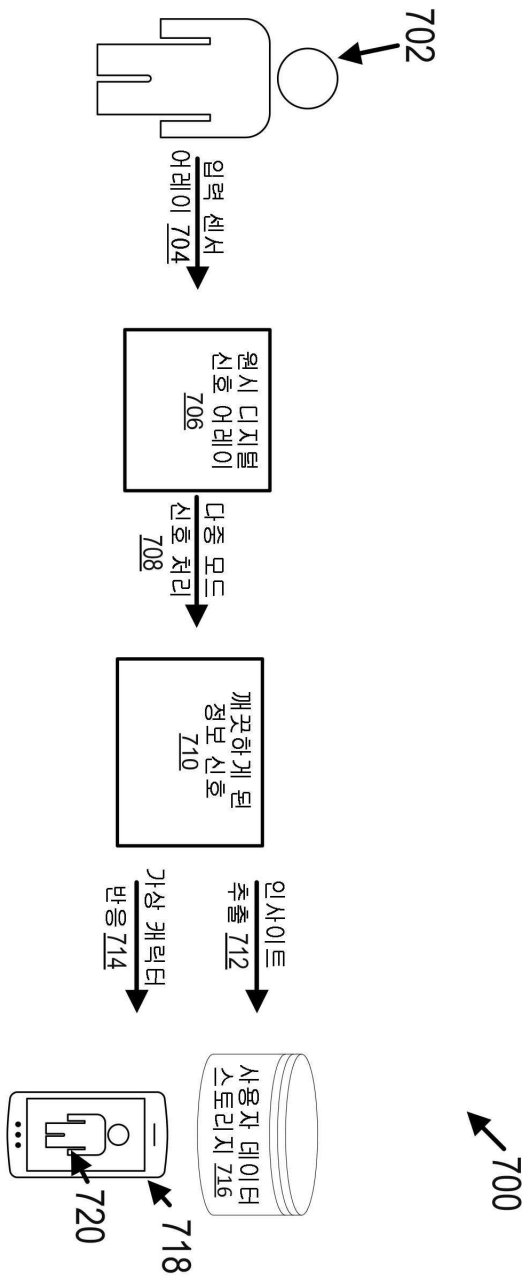
도면5



도면6



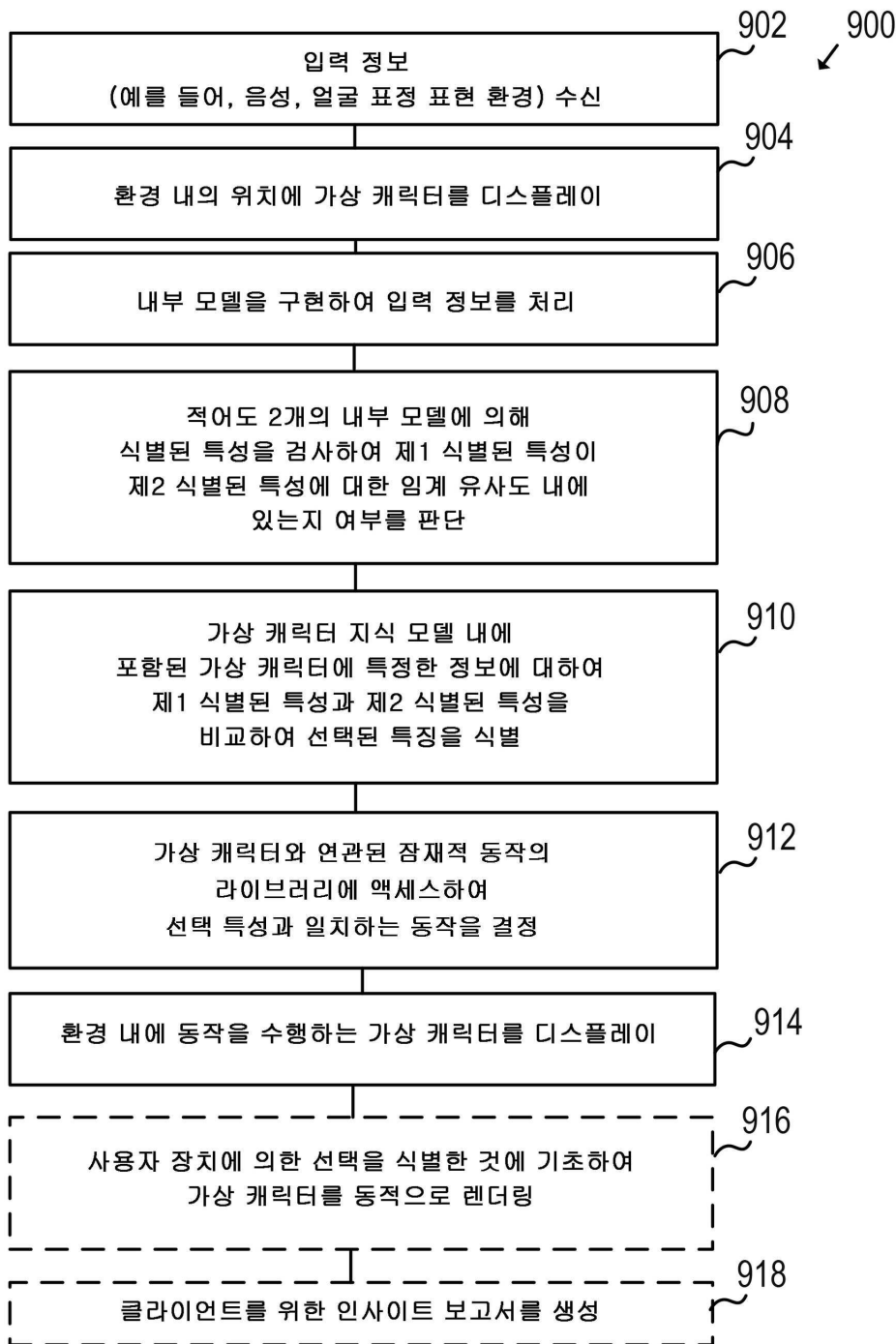
도면7



도면8



도면9



도면10

