



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0121089
(43) 공개일자 2024년08월08일

- (51) 국제특허분류(Int. Cl.)
G06T 13/40 (2011.01) G06N 3/045 (2023.01)
G06N 3/0455 (2023.01) G06N 3/08 (2023.01)
G06T 17/20 (2006.01) G06T 7/11 (2017.01)
G06T 7/50 (2017.01) G06T 7/70 (2017.01)
- (52) CPC특허분류
G06T 13/40 (2013.01)
G06N 3/045 (2023.01)
- (21) 출원번호 10-2023-0013915
(22) 출원일자 2023년02월01일
심사청구일자 2023년02월01일
- (71) 출원인
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
- (72) 발명자
주경돈
울산광역시 울주군 언양읍 유니스트길 50
김민석
울산광역시 울주군 언양읍 유니스트길 50
(뒷면에 계속)
- (74) 대리인
리앤목특허법인

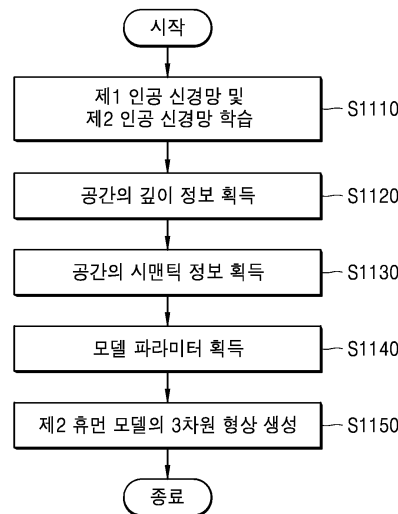
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법, 장치 및 컴퓨터 프로그램

(57) 요약

본 발명의 일 실시예에 따른 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법은, 공간의 깊이 정보를 획득하는 단계; 상기 공간의 시맨틱 정보를 획득하는 단계로써, 상기 시맨틱 정보는 상기 공간에 대한 영상으로부터 인식된 적어도 하나의 객체의 인식 정보 및 상기 영상에서 상기 적어도 하나의 객체 각각에 해당하는 영역을 포함하고; 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 획득하는 단계; 및 학습된 제1 인공 신경망에 상기 깊이 정보, 상기 시맨틱 정보 및 상기 모델 파라미터를 입력하여 상기 공간의 3차원 형상을 고려한 포즈를 갖는 제2 휴먼 모델의 3차원 형상을 생성하는 단계;를 포함할 수 있다.

대표도 - 도11



(52) CPC특허분류

G06N 3/0455 (2023.01)

G06N 3/08 (2023.01)

G06T 17/20 (2013.01)

G06T 7/11 (2017.01)

G06T 7/50 (2017.01)

G06T 7/70 (2017.01)

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

G06T 2207/30196 (2013.01)

(72) 발명자

강창우

울산광역시 울주군 언양읍 유니스트길 50

박정인

울산광역시 울주군 언양읍 유니스트길 50

이 발명을 지원한 국가연구개발사업

과제고유번호 1711160467

과제번호 2022-0-00612-001

부처명 과학기술정보통신부

과제관리(전문)기관명 정보통신기획평가원

연구사업명 사람중심인공지능핵심원천기술개발

연구과제명 일상생활에서 필요한 상식을 기반으로 새로운 사실을 추론하며 이해하는 인공지능 기술 개발

기 여 율 1/2

과제수행기관명 울산과학기술원

연구기간 2022.04.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호 1711152473

과제번호 2020-0-01336-003

부처명 과학기술정보통신부

과제관리(전문)기관명 정보통신기획평가원

연구사업명 정보통신방송혁신인재양성

연구과제명 인공지능대학원지원(울산과학기술원)

기 여 율 1/2

과제수행기관명 울산과학기술원

연구기간 2022.01.01 ~ 2022.12.31

명세서

청구범위

청구항 1

주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법에 있어서,

공간의 깊이 정보를 획득하는 단계;

상기 공간의 시맨틱 정보를 획득하는 단계로써, 상기 시맨틱 정보는 상기 공간에 대한 영상으로부터 인식된 적어도 하나의 객체의 인식 정보 및 상기 영상에서 상기 적어도 하나의 객체 각각에 해당하는 영역을 포함하고;

제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 획득하는 단계; 및

학습된 제1 인공 신경망에 상기 깊이 정보, 상기 시맨틱 정보 및 상기 모델 파라미터를 입력하여 상기 공간의 3차원 형상을 고려한 포즈를 갖는 제2 휴먼 모델의 3차원 형상을 생성하는 단계;를 포함하는, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법.

청구항 2

청구항 1에 있어서

상기 제1 인공 신경망은

하나 이상의 공간의 깊이 정보, 하나 이상의 공간에 대한 시맨틱 정보, 하나 이상의 휴먼 모델을 묘사하는 모델 파라미터 및 하나 이상의 휴먼 모델의 3차원 형상 간의 상관 관계를 학습한 신경망인, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법.

청구항 3

청구항 1에 있어서

상기 제1 인공 신경망은

복수의 레이어로 구성되고,

상기 복수의 레이어 중 적어도 일부 레이어는 공간의 깊이 정보와 상기 공간의 시맨틱 정보로부터 상기 공간을 묘사하는 제2 벡터를 생성하는 제2 인코더에 해당하고,

상기 복수의 레이어 중 적어도 일부 레이어는 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 포함하는 제1 벡터와 상기 제2 벡터로부터 상기 휴먼 모델의 특징과 상기 공간의 특징을 결합한 제3 벡터를 생성하는 제1 인코더에 해당하고,

상기 복수의 레이어 중 적어도 일부 레이어는 상기 제3 벡터 및 상기 제2 벡터로부터 상기 휴먼 모델의 특징과 상기 공간의 특징이 반영된 휴먼 모델의 3차원 형상을 생성하는 디코더에 해당하는, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법.

청구항 4

청구항 1에 있어서

상기 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법은

상기 공간의 깊이 정보를 획득하는 단계 이전에,

상기 제1 인공 신경망을 학습시키는 단계;를 더 포함하는, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법.

청구항 5

청구항 4에 있어서

상기 제1 인공 신경망을 학습시키는 단계는

제2 휴먼 모델의 3차원 형상과 상기 공간의 깊이 정보를 이용하여 제1 오차를 산출하는 단계;

학습된 제2 인공 신경망에 상기 제2 휴먼 모델의 3차원 형상을 입력하여 상기 제2 휴먼 모델의 포즈를 결정하는 단계로써, 상기 제2 인공 신경망은 복수의 휴먼 모델의 3차원 형상과 복수의 포즈 간의 상관 관계를 학습한 신경망이고;

상기 포즈를 참조하여, 상기 제2 휴먼 모델의 3차원 형상을 구성하는 적어도 하나의 지점 각각에 대한 가중치를 결정하는 단계;

상기 제1 오차에 상기 가중치를 적용하여 제2 오차를 산출하는 단계; 및

상기 제2 오차에 기반하여, 상기 제1 인공 신경망을 구성하는 적어도 하나의 파라미터를 갱신하는 단계;를 포함하는, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법.

청구항 6

청구항 5에 있어서

상기 제1 오차는

상기 제2 휴먼 모델의 3차원 형상을 구성하는 제1 지점과 상기 공간 상의 제2 지점 간의 거리 및 상기 제1 지점으로부터 정의되는 제1 방향과 상기 제2 지점으로부터 정의되는 제2 방향 간의 차이를 포함하는, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법.

청구항 7

컴퓨터를 이용하여 제1 항 내지 제6 항 중 어느 한 항의 방법을 실행하기 위하여 매체에 저장된 컴퓨터 프로그램.

청구항 8

주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 장치에 있어서, 상기 장치는 프로세서를 포함하고,

상기 프로세서는,

공간의 깊이 정보를 획득하고,

상기 공간의 시맨틱 정보를 획득하고, 상기 시맨틱 정보는 상기 공간에 대한 영상으로부터 인식된 적어도 하나의 객체의 인식 정보 및 상기 영상에서 상기 적어도 하나의 객체 각각에 해당하는 영역을 포함하고,

제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 획득하고,

학습된 제1 인공 신경망에 상기 깊이 정보, 상기 시맨틱 정보 및 상기 모델 파라미터를 입력하여 상기 공간의 3차원 형상을 고려한 포즈를 갖는 제2 휴먼 모델의 3차원 형상을 생성하는, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 장치.

청구항 9

청구항 8에 있어서

상기 제1 인공 신경망은

하나 이상의 공간의 깊이 정보, 하나 이상의 공간에 대한 시맨틱 정보, 하나 이상의 휴먼 모델을 묘사하는 모델 파라미터 및 하나 이상의 휴먼 모델의 3차원 형상 간의 상관 관계를 학습한 신경망인, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 장치.

청구항 10

청구항 8에 있어서

상기 제1 인공 신경망은

복수의 레이어로 구성되고,

상기 복수의 레이어 중 적어도 일부 레이어는 공간의 깊이 정보와 상기 공간의 시맨틱 정보로부터 상기 공간을 묘사하는 제2 벡터를 생성하는 제2 인코더에 해당하고,

상기 복수의 레이어 중 적어도 일부 레이어는 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 포함하는 제1 벡터와 상기 제2 벡터로부터 상기 휴먼 모델의 특징과 상기 공간의 특징을 결합한 제3 벡터를 생성하는 제1 인코더에 해당하고,

상기 복수의 레이어 중 적어도 일부 레이어는 상기 제3 벡터 및 상기 제2 벡터로부터 상기 휴먼 모델의 특징과 상기 공간의 특징이 반영된 휴먼 모델의 3차원 형상을 생성하는 디코더에 해당하는, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 장치.

청구항 11

청구항 8에 있어서

상기 프로세서는

상기 공간의 깊이 정보를 획득하는 과정 이전에,

상기 제1 인공 신경망을 학습시키는, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 장치.

청구항 12

청구항 11에 있어서

상기 프로세서는

제2 휴먼 모델의 3차원 형상과 상기 공간의 깊이 정보를 이용하여 제1 오차를 산출하고,

학습된 제2 인공 신경망에 상기 제2 휴먼 모델의 3차원 형상을 입력하여 상기 제2 휴먼 모델의 포즈를 결정하고, 상기 제2 인공 신경망은 복수의 휴먼 모델의 3차원 형상과 복수의 포즈 간의 상관 관계를 학습한 신경망이고,

상기 포즈를 참조하여, 상기 제2 휴먼 모델의 3차원 형상을 구성하는 적어도 하나의 지점 각각에 대한 가중치를 결정하고,

상기 제1 오차에 상기 가중치를 적용하여 제2 오차를 산출하고,

상기 제2 오차에 기반하여, 상기 제1 인공 신경망을 구성하는 적어도 하나의 파라미터를 갱신하는, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 장치.

청구항 13

청구항 12에 있어서

상기 제1 오차는

상기 제2 휴먼 모델의 3차원 형상을 구성하는 제1 지점과 상기 공간 상의 제2 지점 간의 거리 및 상기 제1 지점으로부터 정의되는 제1 방향과 상기 제2 지점으로부터 정의되는 제2 방향 간의 차이를 포함하는, 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 장치.

발명의 설명

기술 분야

본 발명은 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법, 장치 및 컴퓨터 프로그램에 관한 것이다.

배경 기술

[0001]

- [0002] 인공지능의 등장으로 가상의 공간 상에서 다양한 콘텐츠를 제공하는 기술들은 비약적으로 발전하고 있는 추세이다.
- [0003] 특히, 3차원 공간에 3차원 휴먼 모델을 생성하는 것은 다가오는 메타버스 시대에 인간을 묘사하고 타인과 소통하는 매체로 많은 주목을 받기 시작했다.
- [0004] 주어진 환경 또는 장면에서 자연스러운 3차원 휴먼 모델을 생성하려면 장면 컨텍스트 정보와 3차원 휴먼 모델이 운동학적으로 실현 가능한 포즈를 고려하는 것이 필수적이다.

발명의 내용

해결하려는 과제

- [0005] 본 발명은 주어진 3차원 공간에서 보다 자연스러운 3차원 휴먼 모델을 생성할 수 있도록 하고자 한다.
- [0006] 또한 본 발명은 3차원 휴먼 모델을 생성함에 있어서 사용되는 신경망들을 효과적으로 학습하는 과정을 제공하고자 한다.

과제의 해결 수단

- [0007] 본 발명의 일 실시예에 따른 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법은, 공간의 깊이 정보를 획득하는 단계; 상기 공간의 시맨틱 정보를 획득하는 단계로써, 상기 시맨틱 정보는 상기 공간에 대한 영상으로부터 인식된 적어도 하나의 객체의 인식 정보 및 상기 영상에서 상기 적어도 하나의 객체 각각에 해당하는 영역을 포함하고; 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 획득하는 단계; 및 학습된 제1 인공 신경망에 상기 깊이 정보, 상기 시맨틱 정보 및 상기 모델 파라미터를 입력하여 상기 공간의 3차원 형상을 고려한 포즈를 갖는 제2 휴먼 모델의 3차원 형상을 생성하는 단계;를 포함할 수 있다.
- [0008] 상기 제1 인공 신경망은 하나 이상의 공간의 깊이 정보, 하나 이상의 공간에 대한 시맨틱 정보, 하나 이상의 휴먼 모델을 묘사하는 모델 파라미터 및 하나 이상의 휴먼 모델의 3차원 형상 간의 상관 관계를 학습한 신경망일 수 있다.
- [0009] 상기 제1 인공 신경망은 복수의 레이어로 구성되고, 상기 복수의 레이어 중 적어도 일부 레이어는 공간의 깊이 정보와 상기 공간의 시맨틱 정보로부터 상기 공간을 묘사하는 제2 벡터를 생성하는 제2 인코더에 해당하고, 상기 복수의 레이어 중 적어도 일부 레이어는 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 포함하는 제1 벡터와 상기 제2 벡터로부터 상기 휴먼 모델의 특징과 상기 공간의 특징을 결합한 제3 벡터를 생성하는 제1 인코더에 해당하고, 상기 복수의 레이어 중 적어도 일부 레이어는 상기 제3 벡터 및 상기 제2 벡터로부터 상기 휴먼 모델의 특징과 상기 공간의 특징이 반영된 휴먼 모델의 3차원 형상을 생성하는 디코더에 해당할 수 있다.
- [0010] 상기 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 방법은 상기 공간의 깊이 정보를 획득하는 단계 이전에, 상기 제1 인공 신경망을 학습시키는 단계;를 더 포함할 수 있다.
- [0011] 상기 제1 인공 신경망을 학습시키는 단계는 제2 휴먼 모델의 3차원 형상과 상기 공간의 깊이 정보를 이용하여 제1 오차를 산출하는 단계; 학습된 제2 인공 신경망에 상기 제2 휴먼 모델의 3차원 형상을 입력하여 상기 제2 휴먼 모델의 포즈를 결정하는 단계로써, 상기 제2 인공 신경망은 복수의 휴먼 모델의 3차원 형상과 복수의 포즈 간의 상관 관계를 학습한 신경망이고; 상기 포즈를 참조하여, 상기 제2 휴먼 모델의 3차원 형상을 구성하는 적어도 하나의 지점 각각에 대한 가중치를 결정하는 단계; 상기 제1 오차에 상기 가중치를 적용하여 제2 오차를 산출하는 단계; 및 상기 제2 오차에 기반하여, 상기 제1 인공 신경망을 구성하는 적어도 하나의 파라미터를 갱신하는 단계;를 포함할 수 있다.
- [0012] 상기 제1 오차는 상기 제2 휴먼 모델의 3차원 형상을 구성하는 제1 지점과 상기 공간 상의 제2 지점 간의 거리 및 상기 제1 지점으로부터 정의되는 제1 방향과 상기 제2 지점으로부터 정의되는 제2 방향 간의 차이를 포함할 수 있다.
- [0013] 본 발명의 일 실시예에 따른 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 장치에 있어서, 상기 장치는 프로세서를 포함하고, 상기 프로세서는, 공간의 깊이 정보를 획득하고, 상기 공간의 시맨틱 정보를 획득하고, 상기 시맨틱 정보는 상기 공간에 대한 영상으로부터 인식된 적어도 하나의 객체의 인식 정보 및 상기 영상에서 상기 적어도 하나의 객체 각각에 해당하는 영역을 포함하고, 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 획득하고, 학습된 제1 인공 신경망에 상기 깊이 정보, 상기 시맨틱 정보 및 상기 모델 파라미터

를 입력하여 상기 공간의 3차원 형상을 고려한 포즈를 갖는 제2 휴먼 모델의 3차원 형상을 생성할 수 있다.

[0014] 상기 제1 인공 신경망은 하나 이상의 공간의 깊이 정보, 하나 이상의 공간에 대한 시맨틱 정보, 하나 이상의 휴먼 모델을 묘사하는 모델 파라미터 및 하나 이상의 휴먼 모델의 3차원 형상 간의 상관 관계를 학습한 신경망일 수 있다.

[0015] 상기 제1 인공 신경망은 복수의 레이어로 구성되고, 상기 복수의 레이어 중 적어도 일부 레이어는 공간의 깊이 정보와 상기 공간의 시맨틱 정보로부터 상기 공간을 묘사하는 제2 벡터를 생성하는 제2 인코더에 해당하고, 상기 복수의 레이어 중 적어도 일부 레이어는 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 포함하는 제1 벡터와 상기 제2 벡터로부터 상기 휴먼 모델의 특징과 상기 공간의 특징을 결합한 제3 벡터를 생성하는 제1 인코더에 해당하고, 상기 복수의 레이어 중 적어도 일부 레이어는 상기 제3 벡터 및 상기 제2 벡터로부터 상기 휴먼 모델의 특징과 상기 공간의 특징이 반영된 휴먼 모델의 3차원 형상을 생성하는 디코더에 해당할 수 있다.

[0016] 상기 프로세서는 상기 공간의 깊이 정보를 획득하는 과정 이전에, 상기 제1 인공 신경망을 학습시킬 수 있다.

[0017] 상기 프로세서는 제2 휴먼 모델의 3차원 형상과 상기 공간의 깊이 정보를 이용하여 제1 오차를 산출하고, 학습된 제2 인공 신경망에 상기 제2 휴먼 모델의 3차원 형상을 입력하여 상기 제2 휴먼 모델의 포즈를 결정하고, 상기 제2 인공 신경망은 복수의 휴먼 모델의 3차원 형상과 복수의 포즈 간의 상관 관계를 학습한 신경망이고, 상기 포즈를 참조하여, 상기 제2 휴먼 모델의 3차원 형상을 구성하는 적어도 하나의 지점 각각에 대한 가중치를 결정하고, 상기 제1 오차에 상기 가중치를 적용하여 제2 오차를 산출하고, 상기 제2 오차에 기반하여, 상기 제1 인공 신경망을 구성하는 적어도 하나의 파라미터를 갱신할 수 있다.

[0018] 상기 제1 오차는 상기 제2 휴먼 모델의 3차원 형상을 구성하는 제1 지점과 상기 공간 상의 제2 지점 간의 거리 및 상기 제1 지점으로부터 정의되는 제1 방향과 상기 제2 지점으로부터 정의되는 제2 방향 간의 차이를 포함할 수 있다.

발명의 효과

[0019] 본 발명에 따르면 주어진 3차원 공간에서 보다 자연스러운 3차원 휴먼 모델을 생성할 수 있다.

[0020] 또한 3차원 휴먼 모델을 생성함에 있어서 사용되는 신경망들을 효과적으로 학습시킬 수 있다.

도면의 간단한 설명

[0021] 도 1은 본 발명의 일 실시예에 따른 휴먼 모델 생성 시스템의 구성을 개략적으로 도시한 도면이다.

도 2는 본 발명의 일 실시예에 따른 서버(100)의 구성을 개략적으로 도시한 도면이다.

도 3은 본 발명의 일 실시예에 따른 사용자 단말(300)의 구성을 개략적으로 도시한 도면이다.

도 4는 본 발명의 일 실시예에 따른 제1 인공 신경망(500)의 예시적인 구조를 설명하기 위한 도면이다.

도 5는 본 발명의 일 실시예에 따른 제2 인공 신경망(600)의 입력 및 출력을 설명하기 위한 도면이다.

도 6은 제2 인공 신경망(600)의 예시적인 학습 데이터를 도시한 도면이다.

도 7 내지 도 10은 제1 프로세서(120)가 제1 인공 신경망(500)을 학습시키는 과정을 설명하기 위한 도면이다.

도 11은 본 발명의 제1 프로세서(120)에 의해 수행되는 모델 생성 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0022] 본 발명은 다양한 변환을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 본 발명의 효과 및 특징, 그리고 그것들을 달성하는 방법은 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 다양한 형태로 구현될 수 있다.

[0023] 이하, 첨부된 도면을 참조하여 본 발명의 실시예들을 상세히 설명하기로 하며, 도면을 참조하여 설명할 때 동일하거나 대응하는 구성 요소는 동일한 도면부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다.

[0024] 이하의 실시예에서, 제1, 제2 등의 용어는 한정적인 의미가 아니라 하나의 구성 요소를 다른 구성 요소와 구별하는 목적으로 사용되었다. 이하의 실시예에서, 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의

표현을 포함한다. 이하의 실시예에서, 포함하다 또는 가지다 등의 용어는 명세서상에 기재된 특징, 또는 구성요소가 존재함을 의미하는 것이고, 하나 이상의 다른 특징들 또는 구성요소가 부가될 가능성을 미리 배제하는 것은 아니다. 도면에서는 설명의 편의를 위하여 구성 요소들이 그 크기가 과장 또는 축소될 수 있다. 예컨대, 도면에서 나타난 각 구성의 크기 및 형태는 설명의 편의를 위해 임의로 나타내었으므로, 본 발명이 반드시 도시된 바에 한정되지 않는다.

- [0026] 도 1은 본 발명의 일 실시예에 따른 휴먼 모델 생성 시스템의 구성을 개략적으로 도시한 도면이다.
- [0027] 본 발명의 일 실시예에 따른 휴먼 모델 생성 시스템은 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성할 수 있다.
- [0028] 또한 본 발명의 일 실시예에 따른 휴먼 모델 생성 시스템은 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하기 위한 적어도 하나의 인공 신경망을 학습시킬 수 있다.
- [0029] 본 발명에서 '휴먼 모델'은 인체의 형상을 갖는 3차원 객체로써, 컴퓨팅 장치 상에서 구현되는 것을 의미할 수 있다.
- [0030] 본 발명에서 '공간'은 영상 획득 장치 등이 획득한 영상에 기반하여 컴퓨팅 장치 등에서 전송한 휴먼 모델과 함께 표시되는 3차원 공간을 의미할 수 있다. 이와 같은 공간은 영상 획득 장치에 획득한 깊이 영상, RGB 영상 등에 의해 컴퓨팅 장치에 구현될 수 있다. 다만 이는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0031] 본 발명의 다른 실시예에서 '공간'은 컴퓨팅 장치 상에서 개발자 등에 의해 구현된 가상의 공간을 의미할 수도 있다. 가령 공간은 메타버스 서비스 상에서 서비스가 제공되는 가상의 공간이나, 게임 서비스, SNS 서비스 등에서 서비스가 제공되는 가상의 공간일 수도 있다. 다만 이는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0032] 본 발명에서 휴먼 모델의 '공간에 따르는 포즈'는 휴먼 모델이 공간의 형상에 부합되는 자세인 것을 의미할 수 있다. 가령 공간 상에 의자가 놓여진 상황을 가정할 경우, 휴먼 모델의 공간에 따르는 포즈는 휴먼 모델이 공간상의 의자에 앉아있는 포즈를 의미할 수 있다. 다만 이와 같은 공간 및 포즈는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0033] 본 발명의 일 실시예에 따른 휴먼 모델 생성 시스템은 도 1에 도시된 바와 같이 서버(100), 서비스 서버(200), 사용자 단말(300) 및 통신망(400)을 포함할 수 있다.
- [0035] 본 발명의 일 실시예에 따른 서비스 서버(200)는 서버(100)가 생성한 휴먼 모델에 기반한 서비스를 제공하는 장치일 수 있다. 가령 서비스 서버(200)는 메타버스 공간 상에서 사용자들이 자신이 실제 거주하는 주택의 형상과 동일하거나 유사한 형상을 갖는 주거 공간을 구축하고, 구축된 공간 상에서 다양한 활동을 수행할 수 있도록 하는 서비스를 제공하는 장치일 수 있다. 다만 이는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0036] 이와 같은 서비스 서버(200)는 휴먼 모델을 생성하는 서버(100)와 동일한 주체에 의해 운영되는 장치일 수도 있고, 서버(100)와 상이한 주체에 의해 운영되는 장치일 수도 있다.
- [0037] 본 발명의 선택적 실시예에서 서비스 서버(200)는 서버(100)와 일체로써 구성될 수도 있다.
- [0039] 본 발명의 일 실시예에 따른 통신망(400)은 휴먼 모델 생성 시스템의 각 구성 간의 데이터 송수신을 매개하는 통신망을 의미할 수 있다. 가령 통신망(400)은 LANs(Local Area Networks), WANs(Wide Area Networks), MANs(Metropolitan Area Networks), ISDNs(Integrated Service Digital Networks) 등의 유선 네트워크나, 무선 LANs, CDMA, 블루투스, 위성 통신 등의 무선 네트워크를 망라할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다.
- [0040] 예를 들어 통신망(400)은 사용자 단말(300)이 획득한 사용자의 공간 영상을 서버(100)로 전송하는 경로를 제공할거나, 서버(100)에 의해 생성된 휴먼 모델을 사용자 단말(300) 및/또는 서비스 서버(200)로 전송하는 경로를

제공할 수 있다. 다만 이는 예시적인 것으로 본 발명의 사상이 이에 한정되는 것은 아니다.

- [0042] 본 발명의 일 실시예에 따른 서버(100)는 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성할 수 있다. 또한 서버(100)는 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하기 위한 적어도 하나의 인공 신경망을 학습시킬 수 있다.
- [0043] 도 2는 본 발명의 일 실시예에 따른 서버(100)의 구성을 개략적으로 도시한 도면이다.
- [0044] 도 2를 참조하면, 본 발명의 일 실시예에 따른 서버(100)는 통신부(110), 제1 프로세서(120), 메모리(130) 및 제2 프로세서(140)를 포함할 수 있다. 또한 도면에는 도시되지 않았으나, 본 발명의 일 실시예에 따른 서버(100)는 입/출력부, 프로그램 저장부 등을 더 포함할 수 있다.
- [0045] 통신부(110)는 서버(100)가 서비스 서버(200) 및/또는 사용자 단말(300)과 같은 다른 네트워크 장치와 유무선 연결을 통해 제어 신호 또는 데이터 신호와 같은 신호를 송수신하기 위해 필요한 하드웨어 및 소프트웨어를 포함하는 장치일 수 있다.
- [0046] 제1 프로세서(120)는 학습된 하나 이상의 인공 신경망을 이용하여 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성하는 일련의 과정을 수행하는 장치일 수 있다.
- [0047] 이때 프로세서(Processor)는, 예를 들어 프로그램 내에 포함된 코드 또는 명령으로 표현된 기능을 수행하기 위해 물리적으로 구조화된 회로를 갖는, 하드웨어에 내장된 데이터 처리 장치를 의미할 수 있다. 이와 같이 하드웨어에 내장된 데이터 처리 장치의 일 예로써, 마이크로프로세서(Microprocessor), 중앙처리장치(Central Processing Unit: CPU), 프로세서 코어(Processor Core), 멀티프로세서(Multiprocessor), ASIC(Application-Specific Integrated Circuit), FPGA(Field Programmable Gate Array) 등의 처리 장치를 망라할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다.
- [0048] 메모리(130)는 서버(100)가 처리하는 데이터를 일시적 또는 영구적으로 저장하는 기능을 수행한다. 메모리는 자기 저장 매체(Magnetic Storage Media) 또는 플래시 저장 매체(Flash Storage Media)를 포함할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다.
- [0049] 가령 메모리(130)는 학습된 인공 신경망을 구성하는 데이터들(가령 계수들)을 일시적 및/또는 영구적으로 저장할 수 있다. 물론 메모리(130)는 인공 신경망을 학습하기 위한 학습 데이터 또는 사용자 단말(300)로부터 수신된 데이터를 저장할 수도 있다. 다만 이는 예시적인 것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0050] 제2 프로세서(140)는 전술한 제1 프로세서(120)의 제어에 따라 연산을 수행하는 장치를 의미할 수 있다. 이때 제2 프로세서(140)는 전술한 제1 프로세서(120)보다 높은 연산 능력을 갖는 장치일 수 있다. 가령 제2 프로세서(140)는 GPU(Graphics Processing Unit)로 구성될 수 있다.
- [0051] 본 발명의 일 실시예에 따른 제2 프로세서(140)는 제1 프로세서(120)의 제어에 따라 학습된 인공 신경망을 이용하여 주어진 공간에 따르는 포즈를 갖는 휴먼 모델을 생성할 수 있다. 다만 이는 예시적인 것으로 본 발명의 사상이 이에 한정되는 것은 아니다. 본 발명의 일 실시예에서, 제2 프로세서(140)는 복수일 수도 있고, 단수일 수도 있다.
- [0053] 본 발명의 일 실시예에 따른 사용자 단말(300)은 공간의 영상을 획득하여 서버(100) 및/또는 서비스 서버(200)로 전송하고, 서버(100) 및/또는 서비스 서버(200)로부터 공간에 따르는 포즈를 갖는 휴먼 모델을 수신하여 사용자에게 제공하는 장치일 수 있다.
- [0054] 도 3은 본 발명의 일 실시예에 따른 사용자 단말(300)의 구성을 개략적으로 도시한 도면이다.
- [0055] 도 3을 참조하면, 본 발명의 일 실시예에 따른 사용자 단말(300)은 통신부(310), 제3 프로세서(320), 메모리(330), 영상 획득부(340), 제4 프로세서(350) 및 디스플레이부(360)를 포함할 수 있다.
- [0056] 통신부(310)는 사용자 단말(300)이 서버(100) 및/또는 서비스 서버(200)와 같은 다른 네트워크 장치와 유무선 연결을 통해 제어 신호 또는 데이터 신호와 같은 신호를 송수신하기 위해 필요한 하드웨어 및 소프트웨어를 포함하는 장치일 수 있다.
- [0057] 제3 프로세서(320)는 사용자의 조작에 따라 공간의 영상을 획득하고, 수신된 휴먼 모델을 사용자에게 제공하는

장치일 수 있다. 이때 프로세서(Processor)는, 예를 들어 프로그램 내에 포함된 코드 또는 명령으로 표현된 기능을 수행하기 위해 물리적으로 구조화된 회로를 갖는, 하드웨어에 내장된 데이터 처리 장치를 의미할 수 있다. 이와 같이 하드웨어에 내장된 데이터 처리 장치의 일 예로써, 마이크로프로세서(Microprocessor), 중앙처리장치(Central Processing Unit: CPU), 프로세서 코어(Processor Core), 멀티프로세서(Multiprocessor), ASIC(Application-Specific Integrated Circuit), FPGA(Field Programmable Gate Array) 등의 처리 장치를 망라할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다.

[0058] 메모리(330)는 사용자 단말(300)이 처리하는 데이터를 일시적 또는 영구적으로 저장하는 기능을 수행한다. 메모리(330)는 자기 저장 매체(Magnetic Storage Media) 또는 플래시 저장 매체(Flash Storage Media)를 포함할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다. 가령 메모리(330)는 영상 획득부(340)가 획득한 영상을 일시적 및/또는 영구적으로 저장할 수 있다. 다만 이는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.

[0059] 영상 획득부(340)는 공간에 대한 영상을 획득하고 제3 프로세서(320)에 제공하는 장치일 수 있다. 이와 같은 영상 획득부(340)는 사용자 단말(300)과 일체로써 구성될 수도 있고, 사용자 단말(300)과 별도로 구성될 수도 있다. 영상 획득부(340)가 사용자 단말(300)과 별도로 구성되는 경우, 영상 획득부(340)와 사용자 단말(300)은 소정의 통신 방식으로 연결될 수 있다. 가령 영상 획득부(340)와 사용자 단말(300)은 무선통신 방식으로 연결될 수 있다. 다만 이는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.

[0060] 한편 영상 획득부(340)는 공간에 대한 깊이 영상과 RGB 영상을 모두 획득할 수 있다. 이때 깊이 영상은 사용자 단말(300)과 공간을 구성하는 복수의 지점 간의 거리 정보를 갖는 영상일 수 있고, RGB 영상은 공간에 대한 색상 정보를 갖는 영상일 수 있다.

[0061] 제4 프로세서(350)는 전술한 제3 프로세서(320)의 제어에 따라 연산을 수행하는 장치를 의미할 수 있다. 이때 제4 프로세서(350)는 전술한 제3 프로세서(320)보다 높은 연산 능력을 갖는 장치일 수 있다. 가령 제4 프로세서(350)는 GPU(Graphics Processing Unit)로 구성될 수 있다. 다만 이는 예시적인 것으로 본 발명의 사상이 이에 한정되는 것은 아니다. 본 발명의 일 실시예에서, 제4 프로세서(350)는 복수일 수도 있고, 단수일 수도 있다.

[0062] 디스플레이부(360)는 서버(100)로부터 수신된 휴먼 모델을 표시하는 장치일 수 있다. 가령 디스플레이부(360)는 사용자 단말(300)이 획득한 영상 상에 서버(100)로부터 수신된 휴먼 모델을 함께 표시할 수 있다. 다만 이와 같은 표시 내용은 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.

[0063] 본 발명의 일 실시예에서 사용자 단말(300)은 휴대폰, 태블릿 PC 및 PC 등과 같은 범용적 정보처리장치로 구현될 수도 있다. 다만 이는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.

[0065] 본 발명에서 '제1 인공 신경망'은 하나 이상의 공간의 깊이 정보, 하나 이상의 공간에 대한 시맨틱 정보, 하나 이상의 휴먼 모델을 묘사하는 모델 파라미터 및 하나 이상의 휴먼 모델의 3차원 형상 간의 상관 관계를 학습한 신경망 일 수 있다.

[0066] 바꾸어 말하면 제1 인공 신경망은 특정 공간의 깊이 정보, 해당 공간에 대한 시맨틱 정보 및 휴먼 모델을 묘사하는 파라미터가 입력 됨에 따라, 해당 공간에 따르는 휴먼 모델의 3차원 형상을 출력하도록 학습된 신경망 일 수 있다.

[0067] 한편 본 발명에서 공간의 '시맨틱 정보'는 공간에 대한 RGB 영상으로부터 인식된 적어도 하나의 객체의 인식 정보 및 영상에서 적어도 하나의 객체 각각에 해당하는 영역을 포함하는 정보일 수 있다. 가령 영상 내에 의자가 포함된 경우, 시맨틱 정보에는 객체의 인식 정보로써 '의자'와 해당 의자가 영상 내에서 위치하는 영역을 나타내는 정보(예를 들어 픽셀정보 등)를 포함할 수 있다.

[0068] 본 발명에서 휴먼 모델을 묘사하는 '파라미터'는 휴먼 모델의 외형을 묘사하는 수치들을 의미할 수 있다. 가령 파라미터는 휴먼 모델의 신장, 허리둘레, 다리 길이, 팔 길이 등에 해당하는 수치들에 해당할 수 있다. 다만 이와 같은 파라미터의 구성은 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.

[0069] 본 발명에서 휴먼 모델의 '3차원 형상'은 가상의 3차원 공간 상에서 묘사되는 휴먼 모델의 3차원적 외형을 의미할 수 있다. 이와 같은 3차원 형상은 가령 포인트 클라우드의 형태로 정의될 수 있다. 다만 이는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.

- [0071] 도 4는 본 발명의 일 실시예에 따른 제1 인공 신경망(500)의 예시적인 구조를 설명하기 위한 도면이다.
- [0072] 본 발명의 일 실시예에 따른 제1 인공 신경망(500)은 복수의 레이어로 구성되며 복수의 레이어 각각은 서로 구분되는 역할을 수행하도록 구성될 수 있다.
- [0073] 가령 복수의 레이어 중 적어도 일부 레이어는 공간의 깊이 정보(Depth map)와 공간의 시맨틱 정보(Semantic Seg.)로부터 공간을 묘사하는 제2 벡터를 생성하는 제2 인코더(520)에 해당할 수 있다.
- [0074] 또한 복수의 레이어 중 적어도 일부 레이어는 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터(Xh)를 포함하는 제1 벡터와 전술한 제2 인코더(520)가 생성한 제2 벡터로부터 휴먼 모델의 특징과 공간의 특징을 결합한 제3 벡터(540)를 생성하는 제1 인코더(510)에 해당할 수 있다.
- [0075] 또한 복수의 레이어 중 적어도 일부 레이어는 제2 인코더(520)가 생성한 제2 벡터와 제1 인코더(510)가 생성한 제3 벡터(540)로부터 휴먼 모델의 특징과 공간의 특징이 반영된 휴먼 모델의 3차원 형상(Xh rec)을 생성하는 디코더(530)에 해당할 수 있다.
- [0076] 다만 이와 같은 인코더(510, 520) 및 디코더(530) 구성은 예시적인 것으로, 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0078] 본 발명에서 '제2 인공 신경망'은 복수의 휴먼 모델의 3차원 형상과 복수의 포즈 간의 상관 관계를 학습한 신경망 일 수 있다.
- [0079] 바꾸어 말하면 제2 인공 신경망은 휴먼 모델의 3차원 형상이 입력 됨에 따라 해당 형상의 포즈를 출력하도록 학습된 신경망 일 수 있다.
- [0080] 도 5는 본 발명의 일 실시예에 따른 제2 인공 신경망(600)의 입력 및 출력을 설명하기 위한 도면이다. 도 6은 제2 인공 신경망(600)의 예시적인 학습 데이터를 도시한 도면이다.
- [0081] 도 5를 참조하면, 제2 인공 신경망(600)은 휴먼 모델의 3차원 형상(Xh rec)이 입력 됨에 따라 해당 형상의 포즈를 출력하도록 학습된 신경망 일 수 있다. 가령 제2 인공 신경망(600)은 휴먼 모델의 3차원 형상(Xh rec)이 입력됨에 따라 모델이 앉아있는 포즈, 모델이 서 있는 포즈, 모델이 누워 있는 포즈 각각에 해당할 확률 값을 출력할 수 있다. 다만 이와 같은 포즈의 수는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0083] 이와 같은 제2 인공 신경망(600)은 복수의 학습 데이터가 기반하여 학습될 수 있다.
- [0084] 도 6을 참조하면, 제2 인공 신경망(600)의 학습에 사용되는 복수의 학습 데이터(810) 각각은 휴먼 모델의 3차원 형상과 해당 형상의 포즈를 포함할 수 있다. 가령 첫 번째 학습 데이터(811)의 경우 벡터의 형태로 생성된 휴먼 모델의 3차원 형상(811A)과 해당 형상의 포즈(811B)를 포함할 수 있다. 이와 유사하게 두 번째 학습 데이터(812) 및 세 번째 학습 데이터(813)도 각각 상술한 항목들을 포함할 수 있다. 다만 이와 같은 학습 데이터의 구성은 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0086] 도 7 내지 도 10은 제1 프로세서(120)가 제1 인공 신경망(500)을 학습시키는 과정을 설명하기 위한 도면이다.
- [0087] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 학습 데이터를 이용하여 제1 인공 신경망(500)을 학습시킬 수 있다. 이때 학습 데이터는 공간의 깊이 정보(Xs), 공간의 시맨틱 정보(Xs) 및 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터(Xh)를 포함할 수 있다.
- [0088] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제1 인공 신경망(500)에 공간의 깊이 정보(Xs), 공간의 시맨틱 정보(Xs) 및 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터(Xh)를 입력하여 공간의 3차원 형상을 고려한 포즈를 갖는 제2 휴먼 모델의 3차원 형상(Xh rec)을 생성할 수 있다.
- [0089] 이때 제1 인공 신경망(500)은 전술한 바와 같이 특정 공간의 깊이 정보, 해당 공간에 대한 시맨틱 정보 및 휴먼 모델을 묘사하는 파라미터가 입력 됨에 따라, 해당 공간에 따르는 휴먼 모델의 3차원 형상을 출력하도록 학습된

신경망 일 수 있다.

- [0090] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제2 휴먼 모델의 3차원 형상(Xh rec)과 공간의 깊이 정보(Xs)를 이용하여 제1 오차(Err1)를 산출할 수 있다.
- [0092] 도 8은 예시적인 제2 휴먼 모델의 3차원 형상(710)과 공간의 깊이 정보에 따르는 공간의 형상(720)을 도시한 도면이다. 이하에서는 도 8의 상황을 가정하여 도 9 및 도 10을 함께 참조하여 설명한다.
- [0093] 도 9에 도시된 바와 같이, 제1 프로세서(120)는 제2 휴먼 모델의 3차원 형상(710)을 구성하는 제1 지점과 상기 공간의 형상(720) 상의 제2 지점 간의 거리(731)를 포함하도록 제1 오차(Err1)를 산출할 수 있다.
- [0094] 또한 제1 프로세서(120)는 도 10에 도시된 바와 같이 제1 지점으로부터 정의되는 제1 방향(733)과 제2 지점으로부터 정의되는 제2 방향(732) 간의 차이를 더 포함하도록 제1 오차(Err1)를 산출할 수도 있다.
- [0096] 다시 도 7을 참조하면, 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제2 인공 신경망(600)에 제2 휴먼 모델의 3차원 형상(Xh rec)을 입력하여 포즈(Pose)를 결정할 수 있다. 이때 제2 인공 신경망(600)은 전술한 바와 같이 휴먼 모델의 3차원 형상이 입력 됨에 따라 해당 형상의 포즈를 출력하도록 학습된 신경망 일 수 있다.
- [0098] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제2 인공 신경망(600)에 의해 결정된 포즈(Pose)를 참조하여, 제2 휴먼 모델의 3차원 형상(Xh rec)을 구성하는 적어도 하나의 지점 각각에 대한 가중치(W)를 결정할 수 있다.
- [0099] 이때 본 발명의 일 실시예에 따른 제1 프로세서(120)는 포즈 별로 제2 휴먼 모델의 3차원 형상(Xh rec)의 각 부위 별 가중치를 상이하게 결정할 수 있다. 가령 도 8에 도시된 상황에 따라 포즈(Pose)가 '앉은 자세'로 결정된 경우, 제1 프로세서(120)는 제2 휴먼 모델의 3차원 형상(Xh rec)의 엉덩이 부위(711)를 구성하는 지점의 가중치를 다른 부위를 구성하는 지점에 비해 높게 설정할 수 있다. 여기서 가중치가 '높게 설정'될 경우, 후술하는 제1 인공 신경망(500)의 학습에 가중치가 높게 설정된 지점의 오차가 다른 지점의 오차에 비해 더 많이 반영될 수 있다.
- [0100] 이와 같이 본 발명은 포즈 별로 휴먼 모델이 공간에 접하는 지점(또는 부위)의 가중치를 상이하게 설정함으로써, 휴먼 모델이 공간에 접함에 있어서 보다 자연스러운 포즈가 가능하도록 할 수 있다.
- [0102] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제1 오차(Err1)에 가중치(W)를 적용하여 제2 오차(Err2)를 산출할 수 있다. 또한 제1 프로세서(120)는 제2 오차(Err2)에 기반하여, 제1 인공 신경망(500)을 구성하는 적어도 하나의 파라미터를 갱신할 수 있다.
- [0103] 이하에서는 도 4 내지 도 10에서 설명한 과정에 따라 제1 인공 신경망(500)과 제2 인공 신경망(600)이 학습되었음을 전제로 제1 프로세서(120)가 휴먼 모델을 생성하는 과정을 설명한다.
- [0105] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 공간의 깊이 정보를 획득할 수 있다. 또한 제1 프로세서(120)는 공간의 시맨틱 정보를 획득할 수 있다.
- [0106] 가령 전술한 바와 같이 사용자 단말(300)의 영상 획득부(340)가 공간에 대한 깊이 영상과 RGB 영상을 모두 획득하는 경우, 제1 프로세서(120)는 사용자 단말(300)로부터 공간의 깊이 영상을 수신하여 깊이 정보를 생성하고, RGB 영상을 수신하여 시맨틱 정보를 생성할 수 있다.
- [0107] 본 발명의 다른 실시예에서, 제1 프로세서(120)는 서비스 서버(200)로부터 공간의 깊이 정보 및 시맨틱 정보를 획득할 수 있다. 가령 서비스 서버(200)에 의해 제공되는 메타버스 서비스 상에 가상의 공간이 구현된 경우, 제1 프로세서(120)는 서비스 서버(200)로부터 가상 공간의 깊이 정보 및 가상의 공간에 대한 시맨틱 정보를 획득할 수도 있다. 이와 같은 경우 가상 공간의 깊이 정보는 깊이 영상에 의해 획득되는 것이 아니라, 소정의 알고리즘에 따라 산출된 것 일 수 있다. 이와 유사하게 가상 공간에 대한 시맨틱 정보 또한 영상의 분석을 통해 획득되는 것이 아니라, 가상 공간의 렌더링을 위한 제어 명령 등으로부터 획득되는 것 일 수 있다.

- [0109] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 획득할 수 있다. 여기서 '모델 파라미터'는 전술한 바와 같이 제1 휴먼 모델의 외형을 묘사하는 수치들로서 가령 휴먼 모델의 신장, 허리둘레, 다리 길이, 팔 길이 등에 해당하는 수치들에 해당할 수 있다. 다만 이와 같은 파라미터의 구성은 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0110] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 서비스 서버(200) 및/또는 사용자 단말(300)로부터 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 획득할 수 있다.
- [0112] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 학습된 제1 인공 신경망(500)에 깊이 정보, 시맨틱 정보 및 모델 파라미터를 입력하여 공간의 3차원 형상을 고려한 포즈를 갖는 제2 휴먼 모델의 3차원 형상을 생성할 수 있다. 여기서 제1 인공 신경망(500)은 도 4에서 설명한 바와 같은 구조로 구성되며, 도 7 내지 도 10에서 설명한 과정에 따라 학습된 신경망 일 수 있다.
- [0113] 이로써 본 발명은 공간의 형상에 보다 부합하는 포즈를 갖는 휴먼 모델을 생성할 수 있다.
- [0115] 도 11은 본 발명의 제1 프로세서(120)에 의해 수행되는 모델 생성 방법을 설명하기 위한 흐름도이다. 이하에서는 도 1 내지 도 10을 함께 참조하여 설명한다.
- [0116] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제1 인공 신경망(500) 및 제2 인공 신경망(600)을 학습시킬 수 있다.(S1110)
- [0117] 본 발명에서 '제1 인공 신경망'은 하나 이상의 공간의 깊이 정보, 하나 이상의 공간에 대한 시맨틱 정보, 하나 이상의 휴먼 모델을 묘사하는 모델 파라미터 및 하나 이상의 휴먼 모델의 3차원 형상 간의 상관 관계를 학습한 신경망 일 수 있다.
- [0118] 바꾸어 말하면 제1 인공 신경망은 특정 공간의 깊이 정보, 해당 공간에 대한 시맨틱 정보 및 휴먼 모델을 묘사하는 파라미터가 입력 됨에 따라, 해당 공간에 따르는 휴먼 모델의 3차원 형상을 출력하도록 학습된 신경망 일 수 있다.
- [0119] 한편 본 발명에서 공간의 '시맨틱 정보'는 공간에 대한 RGB 영상으로부터 인식된 적어도 하나의 객체의 인식 정보 및 영상에서 적어도 하나의 객체 각각에 해당하는 영역을 포함하는 정보일 수 있다. 가령 영상 내에 의자가 포함된 경우, 시맨틱 정보에는 객체의 인식 정보로써 '의자'와 해당 의자가 영상 내에서 위치하는 영역을 나타내는 정보(예를 들어 픽셀정보 등)를 포함할 수 있다.
- [0120] 본 발명에서 휴먼 모델을 묘사하는 '파라미터'는 휴먼 모델의 외형을 묘사하는 수치들을 의미할 수 있다. 가령 파라미터는 휴먼 모델의 신장, 허리둘레, 다리 길이, 팔 길이 등에 해당하는 수치들에 해당할 수 있다. 다만 이와 같은 파라미터의 구성은 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0121] 본 발명에서 휴먼 모델의 '3차원 형상'은 가상의 3차원 공간 상에서 묘사되는 휴먼 모델의 3차원적 외형을 의미할 수 있다. 이와 같은 3차원 형상은 가령 포인트 클라우드의 형태로 정의될 수 있다. 다만 이는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0123] 도 4는 본 발명의 일 실시예에 따른 제1 인공 신경망(500)의 예시적인 구조를 설명하기 위한 도면이다.
- [0124] 본 발명의 일 실시예에 따른 제1 인공 신경망(500)은 복수의 레이어로 구성되며 복수의 레이어 각각은 서로 구분되는 역할을 수행하도록 구성될 수 있다.
- [0125] 가령 복수의 레이어 중 적어도 일부 레이어는 공간의 깊이 정보(Depth map)와 공간의 시맨틱 정보(Semantic Seg.)로부터 공간을 묘사하는 제2 벡터를 생성하는 제2 인코더(520)에 해당할 수 있다.
- [0126] 또한 복수의 레이어 중 적어도 일부 레이어는 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터(Xh)를 포함하는 제1 벡터와 전술한 제2 인코더(520)가 생성한 제2 벡터로부터 휴먼 모델의 특징과 공간의 특징을 결합한 제3 벡터(540)를 생성하는 제1 인코더(510)에 해당할 수 있다.

- [0127] 또한 복수의 레이어 중 적어도 일부 레이어는 제2 인코더(520)가 생성한 제2 벡터와 제1 인코더(510)가 생성한 제3 벡터(540)로부터 휴먼 모델의 특징과 공간의 특징이 반영된 휴먼 모델의 3차원 형상(Xh rec)을 생성하는 디코더(530)에 해당할 수 있다.
- [0128] 다만 이와 같은 인코더(510, 520) 및 디코더(530) 구성은 예시적인 것으로, 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0130] 본 발명에서 '제2 인공 신경망'은 복수의 휴먼 모델의 3차원 형상과 복수의 포즈 간의 상관 관계를 학습한 신경망 일 수 있다.
- [0131] 바꾸어 말하면 제2 인공 신경망은 휴먼 모델의 3차원 형상이 입력 됨에 따라 해당 형상의 포즈를 출력하도록 학습된 신경망 일 수 있다.
- [0132] 도 5는 본 발명의 일 실시예에 따른 제2 인공 신경망(600)의 입력 및 출력을 설명하기 위한 도면이다. 도 6은 제2 인공 신경망(600)의 예시적인 학습 데이터를 도시한 도면이다.
- [0133] 도 5를 참조하면, 제2 인공 신경망(600)은 휴먼 모델의 3차원 형상(Xh rec)이 입력 됨에 따라 해당 형상의 포즈를 출력하도록 학습된 신경망 일 수 있다. 가령 제2 인공 신경망(600)은 휴먼 모델의 3차원 형상(Xh rec)이 입력됨에 따라 모델이 앉아있는 포즈, 모델이 서 있는 포즈, 모델이 누워 있는 포즈 각각에 해당할 확률 값을 출력할 수 있다. 다만 이와 같은 포즈의 수는 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0135] 이와 같은 제2 인공 신경망(600)은 복수의 학습 데이터가 기반하여 학습될 수 있다.
- [0136] 도 6을 참조하면, 제2 인공 신경망(600)의 학습에 사용되는 복수의 학습 데이터(810) 각각은 휴먼 모델의 3차원 형상과 해당 형상의 포즈를 포함할 수 있다. 가령 첫 번째 학습 데이터(811)의 경우 벡터의 형태로 생성된 휴먼 모델의 3차원 형상(811A)과 해당 형상의 포즈(811B)를 포함할 수 있다. 이와 유사하게 두 번째 학습 데이터(812) 및 세 번째 학습 데이터(813)도 각각 상술한 항목들을 포함할 수 있다. 다만 이와 같은 학습 데이터의 구성은 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0138] 도 7 내지 도 10은 제1 프로세서(120)가 제1 인공 신경망(500)을 학습시키는 과정을 설명하기 위한 도면이다.
- [0139] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 학습 데이터를 이용하여 제1 인공 신경망(500)을 학습시킬 수 있다. 이때 학습 데이터는 공간의 깊이 정보(X_s), 공간의 시맨틱 정보(X_s) 및 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터(X_h)를 포함할 수 있다.
- [0140] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제1 인공 신경망(500)에 공간의 깊이 정보(X_s), 공간의 시맨틱 정보(X_s) 및 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터(X_h)를 입력하여 공간의 3차원 형상을 고려한 포즈를 갖는 제2 휴먼 모델의 3차원 형상(Xh rec)을 생성할 수 있다.
- [0141] 이때 제1 인공 신경망(500)은 전술한 바와 같이 특정 공간의 깊이 정보, 해당 공간에 대한 시맨틱 정보 및 휴먼 모델을 묘사하는 파라미터가 입력 됨에 따라, 해당 공간에 따르는 휴먼 모델의 3차원 형상을 출력하도록 학습된 신경망 일 수 있다.
- [0142] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제2 휴먼 모델의 3차원 형상(Xh rec)과 공간의 깊이 정보(X_s)를 이용하여 제1 오차(Err1)를 산출할 수 있다.
- [0144] 도 8은 예시적인 제2 휴먼 모델의 3차원 형상(710)과 공간의 깊이 정보에 따르는 공간의 형상(720)을 도시한 도면이다. 이하에서는 도 8의 상황을 가정하여 도 9 및 도 10을 함께 참조하여 설명한다.
- [0145] 도 9에 도시된 바와 같이, 제1 프로세서(120)는 제2 휴먼 모델의 3차원 형상(710)을 구성하는 제1 지점과 상기 공간의 형상(720) 상의 제2 지점 간의 거리(731)를 포함하도록 제1 오차(Err1)를 산출할 수 있다.
- [0146] 또한 제1 프로세서(120)는 도 10에 도시된 바와 같이 제1 지점으로부터 정의되는 제1 방향(733)과 제2 지점으로

부터 정의되는 제2 방향(732) 간의 차이를 더 포함하도록 제1 오차(Err1)를 산출할 수도 있다.

- [0148] 다시 도 7을 참조하면, 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제2 인공 신경망(600)에 제2 휴먼 모델의 3차원 형상(Xh rec)을 입력하여 포즈(Pose)를 결정할 수 있다. 이때 제2 인공 신경망(600)은 전술한 바와 같이 휴먼 모델의 3차원 형상이 입력 됨에 따라 해당 형상의 포즈를 출력하도록 학습된 신경망 일 수 있다.
- [0150] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제2 인공 신경망(600)에 의해 결정된 포즈(Pose)를 참조하여, 제2 휴먼 모델의 3차원 형상(Xh rec)을 구성하는 적어도 하나의 지점 각각에 대한 가중치(W)를 결정할 수 있다.
- [0151] 이때 본 발명의 일 실시예에 따른 제1 프로세서(120)는 포즈 별로 제2 휴먼 모델의 3차원 형상(Xh rec)의 각 부위 별 가중치를 상이하게 결정할 수 있다. 가령 도 8에 도시된 상황에 따라 포즈(Pose)가 '앉은 자세'로 결정된 경우, 제1 프로세서(120)는 제2 휴먼 모델의 3차원 형상(Xh rec)의 엉덩이 부위(711)를 구성하는 지점의 가중치를 다른 부위를 구성하는 지점에 비해 높게 설정할 수 있다. 여기서 가중치가 '높게 설정'될 경우, 후술하는 제1 인공 신경망(500)의 학습에 가중치가 높게 설정된 지점의 오차가 다른 지점의 오차에 비해 더 많이 반영될 수 있다.
- [0152] 이와 같이 본 발명은 포즈 별로 휴먼 모델이 공간에 접하는 지점(또는 부위)의 가중치를 상이하게 설정함으로써, 휴먼 모델이 공간에 접함에 있어서 보다 자연스러운 포즈가 가능하도록 할 수 있다.
- [0154] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제1 오차(Err1)에 가중치(W)를 적용하여 제2 오차(Err2)를 산출할 수 있다. 또한 제1 프로세서(120)는 제2 오차(Err2)에 기반하여, 제1 인공 신경망(500)을 구성하는 적어도 하나의 파라미터를 갱신할 수 있다.
- [0155] 이하에서는 도 4 내지 도 10에서 설명한 과정에 따라 제1 인공 신경망(500)과 제2 인공 신경망(600)이 학습되었음을 전제로 제1 프로세서(120)가 휴먼 모델을 생성하는 과정을 설명한다.
- [0157] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 공간의 깊이 정보를 획득할 수 있다.(S1120) 또한 제1 프로세서(120)는 공간의 시맨틱 정보를 획득할 수 있다.(S1130)
- [0158] 가령 전술한 바와 같이 사용자 단말(300)의 영상 획득부(340)가 공간에 대한 깊이 영상과 RGB 영상을 모두 획득하는 경우, 제1 프로세서(120)는 사용자 단말(300)로부터 공간의 깊이 영상을 수신하여 깊이 정보를 생성하고, RGB 영상을 수신하여 시맨틱 정보를 생성할 수 있다.
- [0159] 본 발명의 다른 실시예에서, 제1 프로세서(120)는 서비스 서버(200)로부터 공간의 깊이 정보 및 시맨틱 정보를 획득할 수 있다. 가령 서비스 서버(200)에 의해 제공되는 메타버스 서비스 상에 가상의 공간이 구현된 경우, 제1 프로세서(120)는 서비스 서버(200)로부터 가상 공간의 깊이 정보 및 가상의 공간에 대한 시맨틱 정보를 획득할 수도 있다. 이와 같은 경우 가상 공간의 깊이 정보는 깊이 영상에 의해 획득되는 것이 아니라, 소정의 알고리즘에 따라 산출된 것 일 수 있다. 이와 유사하게 가상 공간에 대한 시맨틱 정보 또한 영상의 분석을 통해 획득되는 것이 아니라, 가상 공간의 렌더링을 위한 제어 명령 등으로부터 획득되는 것 일 수 있다.
- [0161] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 획득할 수 있다.(S1140) 여기서 '모델 파라미터'는 전술한 바와 같이 제1 휴먼 모델의 외형을 묘사하는 수치들로서 가령 휴먼 모델의 신장, 허리둘레, 다리 길이, 팔 길이 등에 해당하는 수치들에 해당할 수 있다. 다만 이와 같은 파라미터의 구성은 예시적인것으로 본 발명의 사상이 이에 한정되는 것은 아니다.
- [0162] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 서비스 서버(200) 및/또는 사용자 단말(300)로부터 제1 휴먼 모델을 묘사하는 적어도 하나의 모델 파라미터를 획득할 수 있다.
- [0164] 본 발명의 일 실시예에 따른 제1 프로세서(120)는 학습된 제1 인공 신경망(500)에 깊이 정보, 시맨틱 정보 및

모델 파라미터를 입력하여 공간의 3차원 형상을 고려한 포즈를 갖는 제2 휴먼 모델의 3차원 형상을 생성할 수 있다. (S1150) 여기서 제1 인공 신경망(500)은 도 4에서 설명한 바와 같은 구조로 구성되며, 도 7 내지 도 10에서 설명한 과정에 따라 학습된 신경망 일 수 있다.

[0165] 이로써 본 발명은 공간의 형상에 보다 부합하는 포즈를 갖는 휴먼 모델을 생성할 수 있다.

[0167] 이상 설명된 본 발명에 따른 실시예는 컴퓨터 상에서 다양한 구성요소를 통하여 실행될 수 있는 컴퓨터 프로그램의 형태로 구현될 수 있으며, 이와 같은 컴퓨터 프로그램은 컴퓨터로 판독 가능한 매체에 기록될 수 있다. 이때, 매체는 컴퓨터로 실행 가능한 프로그램을 저장하는 것일 수 있다. 매체의 예시로는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등을 포함하여 프로그램 명령어가 저장되도록 구성된 것이 있을 수 있다.

[0168] 한편, 상기 컴퓨터 프로그램은 본 발명을 위하여 특별히 설계되고 구성된 것이거나 컴퓨터 소프트웨어 분야의 당업자에게 공지되어 사용 가능한 것일 수 있다. 컴퓨터 프로그램의 예에는, 컴파일러에 의하여 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용하여 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드도 포함될 수 있다.

[0169] 본 발명에서 설명하는 특정 실행들은 일 실시 예들로서, 어떠한 방법으로도 본 발명의 범위를 한정하는 것은 아니다. 명세서의 간결함을 위하여, 종래 전자적인 구성들, 제어 시스템들, 소프트웨어, 상기 시스템들의 다른 기능적인 측면들의 기재는 생략될 수 있다. 또한, 도면에 도시된 구성 요소들 간의 선들의 연결 또는 연결 부재들은 기능적인 연결 및/또는 물리적 또는 회로적 연결들을 예시적으로 나타낸 것으로서, 실제 장치에서는 대체 가능하거나 추가의 다양한 기능적인 연결, 물리적인 연결, 또는 회로 연결들로서 나타내어질 수 있다. 또한, "필수적인", "중요하게" 등과 같이 구체적인 언급이 없다면 본 발명의 적용을 위하여 반드시 필요한 구성 요소가 아닐 수 있다.

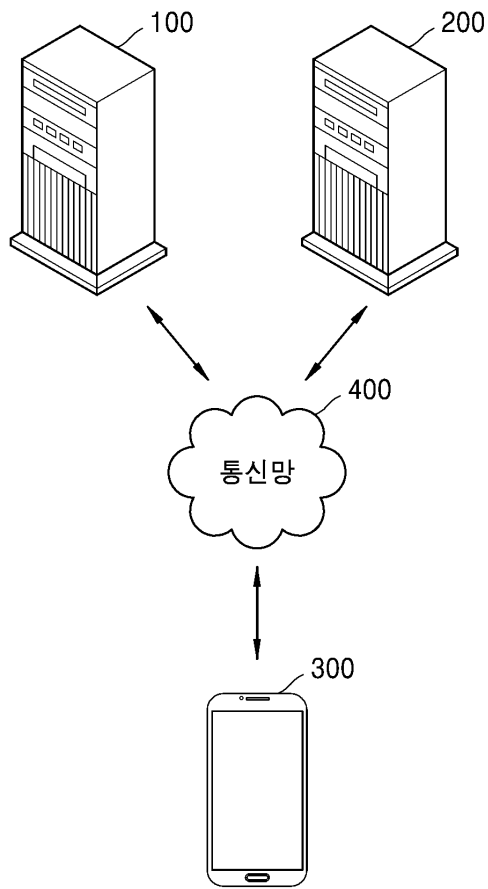
[0170] 따라서, 본 발명의 사상은 상기 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐만 아니라 이 특허청구범위와 균등한 또는 이로부터 등가적으로 변경된 모든 범위는 본 발명의 사상의 범주에 속한다고 할 것이다.

부호의 설명

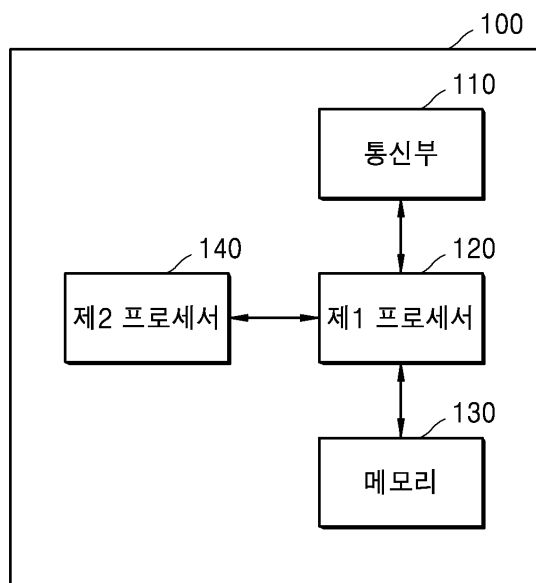
[0171] 100: 서버
110: 통신부
120: 제1 프로세서
130: 메모리
140: 제2 프로세서
200: 서비스 서버
300: 사용자 단말
310: 통신부
320: 제3 프로세서
330: 메모리
340: 영상 획득부
350: 제4 프로세서
360: 디스플레이부
400: 통신망

도면

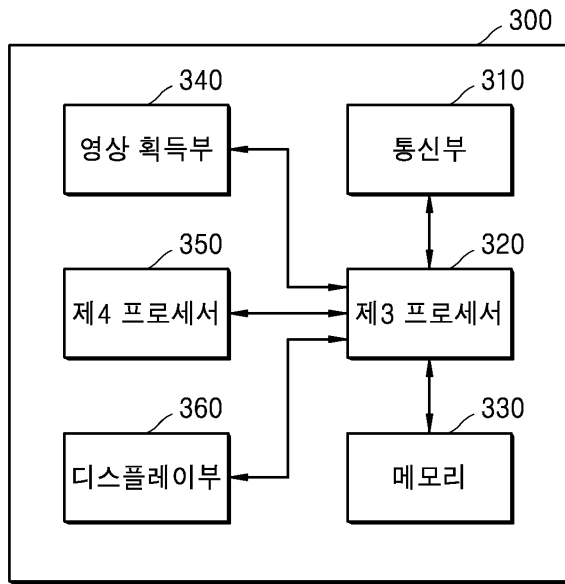
도면1



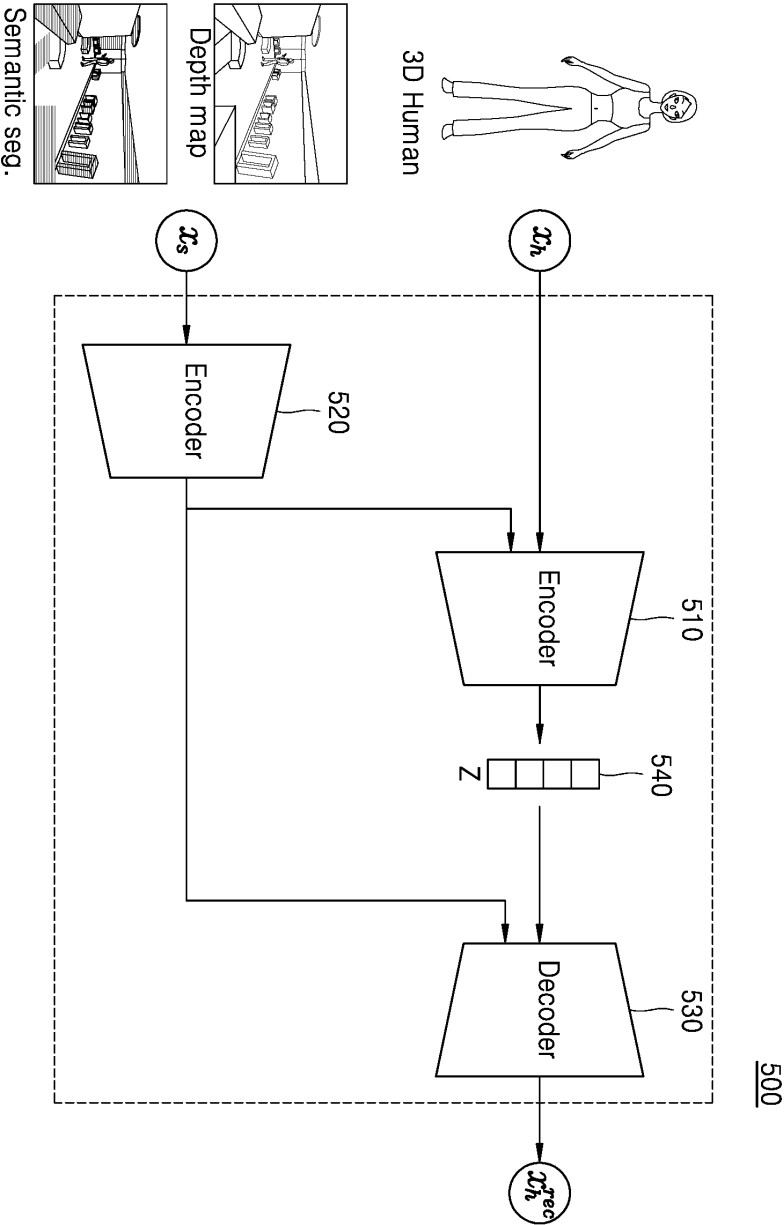
도면2



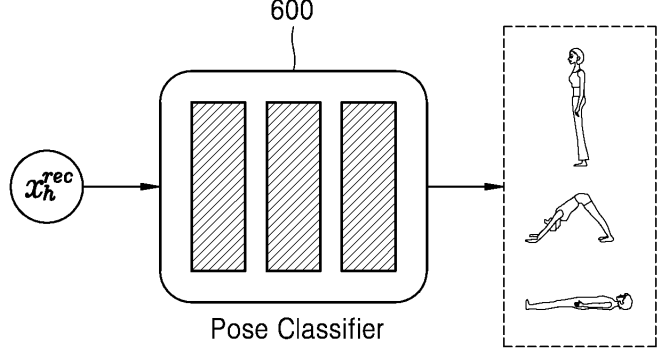
도면3



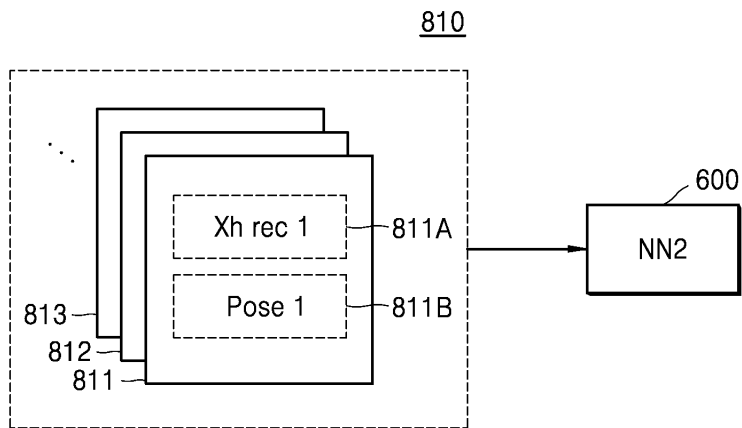
도면4



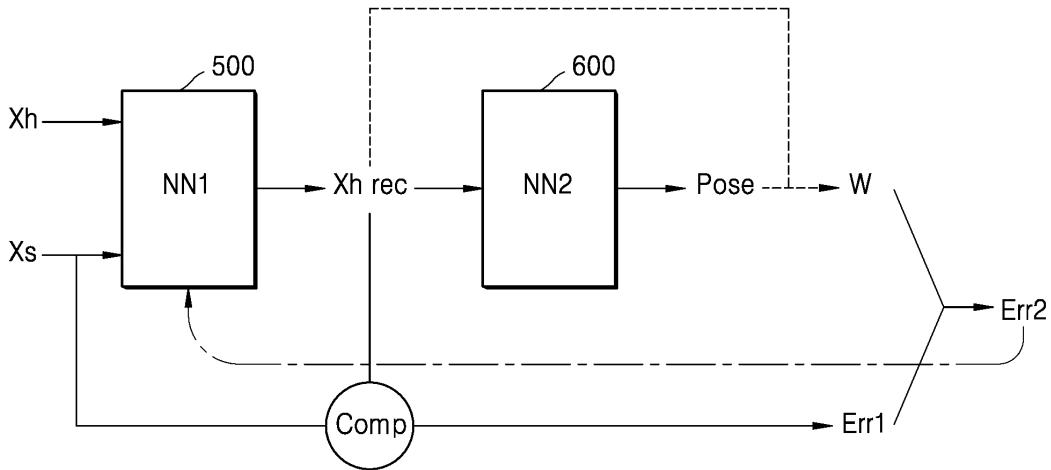
도면5



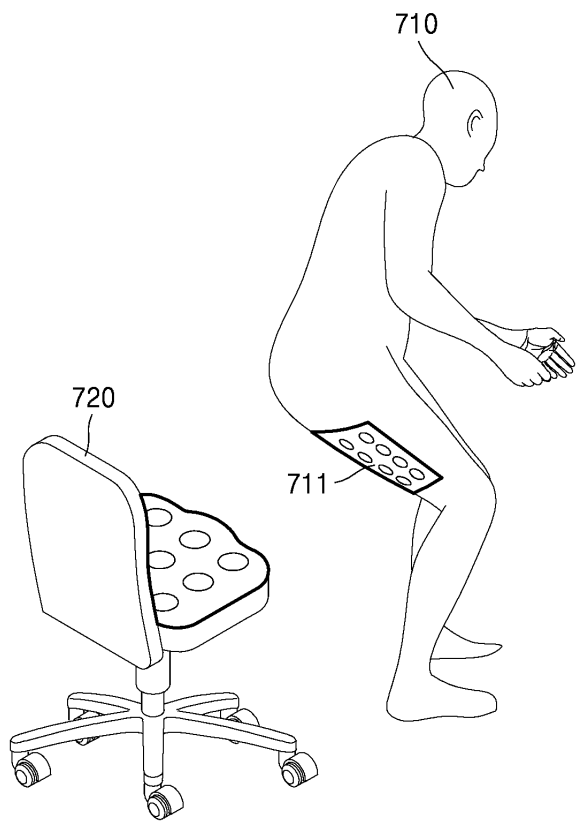
도면6



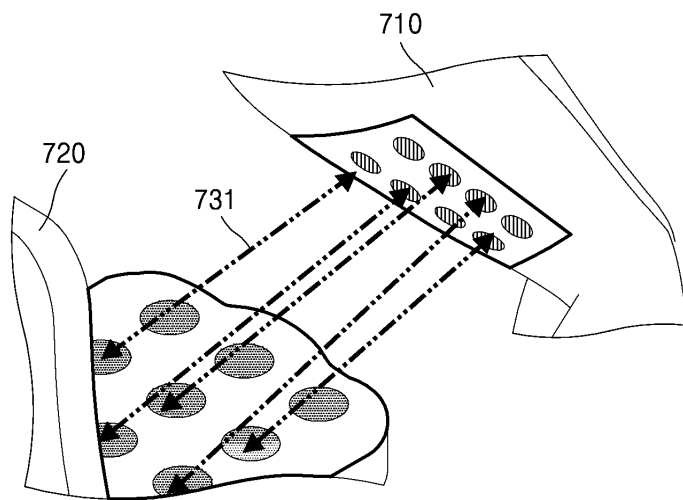
도면7



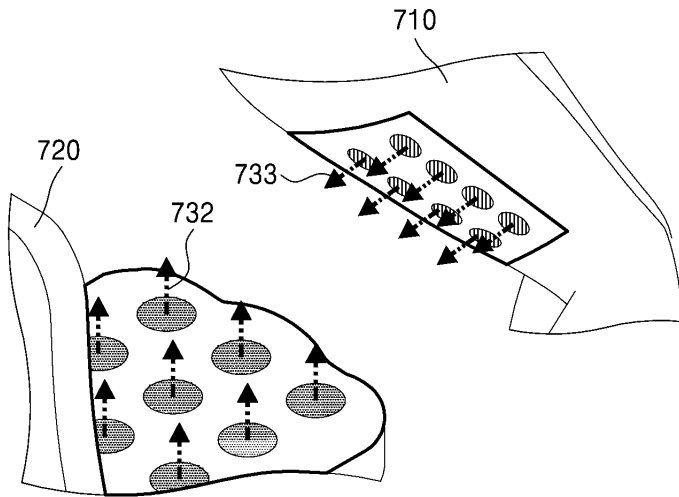
도면8



도면9



도면10



도면11

