

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2025-0052034
(43) 공개일자 2025년04월18일

(51) 국제특허분류(Int. Cl.)
G06F 3/01 (2006.01) G06F 3/00 (2006.01)
G06V 40/20 (2022.01)
(52) CPC특허분류
G06F 3/017 (2013.01)
G06F 3/005 (2013.01)
(21) 출원번호 10-2023-0135078
(22) 출원일자 2023년10월11일
심사청구일자 2023년10월11일

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
강형엽
경기도 용인시 기흥구 서그내로 30, 302호 (서천동)
한동현
경기도 수원시 영통구 덕영대로 1703, 908호(영통동)
(뒷면에 계속)
(74) 대리인
유미특허법인

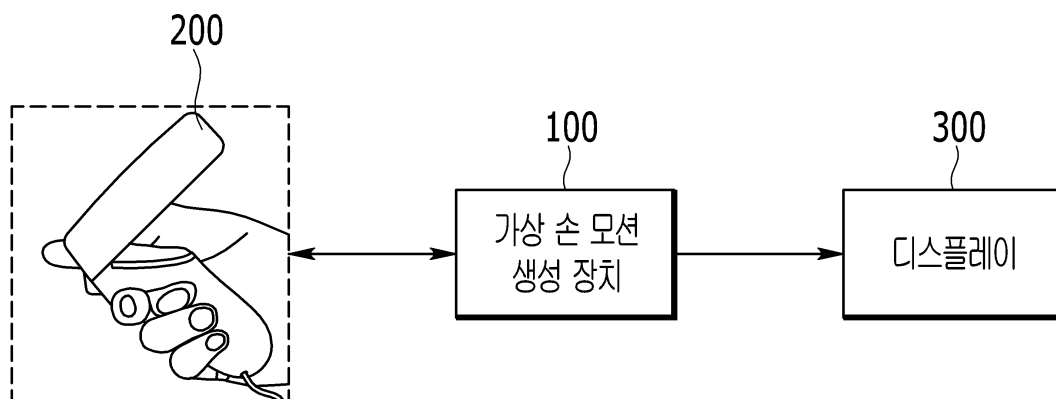
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 컨트롤러를 기반으로 한 가상 손의 모션 생성 장치 및 이의 동작 방법

(57) 요약

적어도 하나의 프로세서에 의해 동작하는 가상 손 모션 생성 장치의 동작 방법으로서, VR 컨트롤러를 가상 손으로 매핑하고, 가상 손과 상호 작용되는 가상 물체를 가상 환경에 생성한다. 가상 손의 포즈, 가상 손이 가상 물체를 쥐는 세기, 가상 물체의 정보를 포함하는 상태 정보를 획득하고, 상태 정보가 입력되면 가상 손이 가상 물체를 쥐는 모션을 수행하기 위해, 가상 손을 구성하는 손가락의 관절들 각각의 관절 각도 변화량이 출력되도록 인공지능 모델을 훈련시킨다.

대표도 - 도1



(52) CPC특허분류

G06F 3/011 (2022.02)

G06V 40/20 (2022.01)

(72) 발명자

이로운

서울특별시 강북구 송인로7길 42-5 (미아동)

김경민

경기도 수원시 영통구 청명로 16, 402호(영통동)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711179316
과제번호	RS-2022-00155911
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능융합혁신인재양성(R&D)
연구과제명	인공지능융합혁신인재양성(경희대학교)
기 여 율	1/1
과제수행기관명	경희대학교
연구기간	2022.07.01 ~ 2025.12.31

명세서

청구범위

청구항 1

적어도 하나의 프로세서에 의해 동작하는 가상 손 모션 생성 장치의 동작 방법으로서,

VR 컨트롤러를 가상 손으로 매핑하고, 상기 가상 손과 상호 작용되는 가상 물체를 가상 환경에 생성하는 단계,

상기 가상 손의 포즈, 상기 가상 손이 상기 가상 물체를 쥐는 세기, 상기 가상 물체의 정보를 포함하는 상태 정보를 획득하는 단계, 그리고

상기 상태 정보가 입력되면 상기 가상 손이 상기 가상 물체를 쥐는 모션을 수행하기 위해, 상기 가상 손을 구성하는 손가락의 관절들 각각의 관절 각도 변화량이 출력되도록 인공지능 모델을 훈련시키는 단계

를 포함하는, 동작 방법.

청구항 2

제1항에 있어서,

상기 상태 정보를 획득하는 단계는,

상기 VR 컨트롤러를 통해 수집한 상기 가상 손의 3차원 위치, 손목 방향, 손목 회전 정도를 상기 가상 손의 포즈로 획득하는 단계

를 포함하는, 동작 방법.

청구항 3

제2항에 있어서,

상기 상태 정보를 획득하는 단계는,

상기 손 모션 캡처 데이터로부터 상기 가상 물체를 쥐는 세기인 트리거 정보를 계산하는 단계

를 포함하는, 동작 방법.

청구항 4

제3항에 있어서,

상기 상태 정보를 획득하는 단계는,

상기 가상 환경에 생성된 복수의 복셀 센서들의 센싱 값들을 기초로 상기 가상 물체의 정보를 획득하는 단계

를 포함하는, 동작 방법.

청구항 5

제4항에 있어서,

상기 복수의 복셀 센서들은 글로벌 복셀 센서, 델타 글로벌 복셀 센서, 그리고 복수의 로컬 복셀 센서들을 포함하고,

상기 글로벌 복셀 센서는 상기 가상 손의 손목에 고정되어 상기 가상 손 앞 일정 범위 내에 위치한 상기 물체를 감지하고, 상기 델타 글로벌 복셀 센서는 상기 물체의 움직임과 이동 방향을 유추하며, 상기 복수의 로컬 복셀 센서들은 상기 물체의 세밀한 굴곡을 감지하는, 동작 방법.

청구항 6

제1항에 있어서,

상기 인공지능 모델을 훈련시키는 단계는,

상기 상태 정보를 상기 관절 각도 변화량으로 매핑하는 함수를 이용하여 상기 인공지능 모델을 훈련시키는, 동작 방법.

청구항 7

제6항에 있어서,

상기 인공지능 모델을 훈련시키는 단계는,

상기 가상 손의 포즈, 상기 가상 물체를 쥐는 세기 및 상기 가상 물체의 정보들 각각의 독립적인 특징이 연결된 출력 벡터를 추출하는 단계,

상기 출력 벡터의 차원보다 고차원의 특성 벡터를 생성하는 단계, 그리고

상기 특성 벡터로부터 상기 가상 손의 각 관절 각도 변화량에 해당하는 동작 벡터를 생성하는 단계를 포함하는, 동작 방법.

청구항 8

제7항에 있어서,

상기 동작 벡터를 생성하는 단계 이후에,

상기 동작 벡터로부터 상기 가상 손을 구성하는 상기 손가락의 복수의 관절들 각각의 토크를 산출하는 단계, 그리고

상기 산출한 토크를 물리 시뮬레이션에 적용하는 단계를

를 포함하는, 동작 방법.

청구항 9

적어도 하나의 프로세서에 의해 동작하는 모션 생성 장치의 동작 방법으로서,

제1 프레임에서 가상 손의 포즈를 획득하는 단계,

상기 가상 손에 매핑되어 있는 VR 컨트롤러의 컨트롤러 버튼으로부터 상기 제1 프레임에 트리거 정보를 획득하고, 상기 제1 프레임을 기준으로 미리 설정된 복수의 이전 프레임들의 이전 트리거 정보를 획득하는 단계,

상기 가상 환경에 생성된 복수의 복셀 센서들의 센싱 값들을 획득하는 단계, 그리고

상기 가상 손의 포즈, 상기 제1 프레임의 트리거 정보와 이전 트리거 정보, 그리고 상기 센싱 값들을 기초로, 상기 제1 프레임의 다음 프레임인 제2 프레임에서의 상기 가상 손을 구성하는 관절들 각각의 관절 각도 변화량을 출력하는 단계

를 포함하는, 동작 방법.

청구항 10

제9항에 있어서,

상기 관절 각도 변화량을 출력하는 단계 이후에,

상기 관절 각도 변화량으로부터 각 관절의 토크를 산출하는 단계, 그리고

상기 산출한 토크를 물리 시뮬레이션에 적용하는 단계를

를 포함하는, 동작 방법.

청구항 11

제9항에 있어서,

상기 복수의 복셀 센서들은 글로벌 복셀 센서, 델타 글로벌 복셀 센서, 그리고 복수의 로컬 복셀 센서들을 포함

하고,

상기 글로벌 복셀 센서는 상기 가상 손의 손목에 고정되어 상기 가상 손 앞 일정 범위 내에 위치한 상기 물체를 감지하고, 상기 델타 글로벌 복셀 센서는 상기 물체의 움직임과 이동 방향을 유추하며, 상기 복수의 로컬 복셀 센서들은 상기 물체의 세밀한 굴곡을 감지하는, 동작 방법.

청구항 12

적어도 하나의 프로세서에 의해 동작하는 모션 생성 장치로서,

VR 컨트롤러로부터 트리거 정보와 포즈 정보를 수집하는 인터페이스, 그리고

프로세서

를 포함하고,

상기 프로세서는,

제1 프레임에서 상기 VR 컨트롤러가 매핑된 가상 손의 포즈, 상기 제1 프레임에서의 트리거 정보와 상기 제1 프레임을 기준으로 미리 설정된 복수의 이전 프레임들의 이전 트리거 정보, 가상 환경에 생성된 복수의 복셀 센서들의 센싱 값들을 포함하는 상태 정보를 기초로, 상기 제1 프레임의 다음 프레임인 제2 프레임에서의 상기 가상 손이 취할 모션을 위해 상기 가상 손을 구성하는 관절들 각각의 관절 각도 변화량을 계산하는, 가상 손 모션 생성 장치.

청구항 13

제12항에 있어서,

상기 프로세서는,

상기 상태 정보가 입력되면 상기 가상 손을 구성하는 관절들 각각의 관절 각도 변화량이 출력되도록 인공지능 모델을 학습시키는, 가상 손 모션 생성 장치.

청구항 14

제13항에 있어서,

상기 프로세서는,

상기 가상 손의 포즈, 상기 트리거 정보 및 상기 센싱 값들 각각의 독립적인 특징이 연결된 출력 벡터를 추출하고, 상기 출력 벡터의 차원보다 고차원의 특성 벡터를 생성하며, 상기 특성 벡터로부터 상기 가상 손의 각 관절 각도 변화량에 해당하는 동작 벡터를 생성하는, 가상 손 모션 생성 장치.

청구항 15

제14항에 있어서,

상기 프로세서는,

상기 관절 각도 변화량으로부터 각 관절의 토크를 산출하고, 상기 산출한 토크를 물리 시뮬레이션에 적용하는, 가상 손 모션 생성 장치.

청구항 16

제12항에 있어서,

상기 복수의 복셀 센서들은 글로벌 복셀 센서, 델타 글로벌 복셀 센서, 그리고 복수의 로컬 복셀 센서들을 포함하고,

상기 글로벌 복셀 센서는 상기 가상 손의 손목에 고정되어 상기 가상 손 앞 일정 범위 내에 위치한 상기 물체를 감지하고, 상기 델타 글로벌 복셀 센서는 상기 물체의 움직임과 이동 방향을 유추하며, 상기 복수의 로컬 복셀 센서들은 상기 물체의 세밀한 굴곡을 감지하는, 모션 생성 장치.

발명의 설명

기술 분야

- [0001] 본 발명은 컨트롤러를 기반으로 한 가상 손의 모션 생성 장치 및 이의 동작 방법에 관한 것으로, 상세하게는 손 추적 센서를 사용하기 어려운 가상현실 환경에서 사용자가 손으로 가상현실 환경의 물체와 자연스럽게 상호 작용할 수 있도록 지원하는 기술에 관한 것이다.

배경 기술

- [0002] 가상현실(VR: Virtual Reality) 기술이 빠르게 발전함에 따라, VR 기술을 활용한 다양한 콘텐츠와 서비스들이 제공되고 있다. 그러나, VR 기술은 아직 완전하게 성숙하지 않은 기술이라 여전히 다양한 제약 사항이 존재하며, 이는 VR 기술의 상호작용 기술에도 큰 영향을 미치고 있다.
- [0003] 현재 가장 많이 사용되는 상호작용 방법은 VR 컨트롤러를 사용한 방식이다. 이 방식은 컨트롤러의 위치를 3D 공간에서 완전히 추적하며, 사용자가 컨트롤러의 버튼을 누를 때마다 각 버튼에 매핑된 미리 정의된 상호작용을 수행한다. 이러한 방식은 VR 환경에서 사용자의 행동을 미리 주어진 행동 들로만 강제하고 복잡하고 섬세한 상호작용을 허용하지 않아, 사용자의 몰입감을 크게 저하시킨다.
- [0004] 이 문제를 해결하기 위해, 최근에는 손 추적을 기반으로 가상 물체를 조작할 수 있는 방법들이 연구되고 있다. RGB 카메라나, 깊이 센서(depth sensor), 광학 카메라 등을 이용한 손 추적 기술을 이용하면 사용자의 손 모션을 실시간으로 추적하고, 추적한 손 모션을 그대로 가상 환경에 매핑해 가상 물체와 상호 작용하는 것이 가능해진다.
- [0005] 기존의 손 추적 기술은 손의 골격 구조만을 매핑했으나, 최근에는 가상현실 HMD(Head Mounted Display)에 다수의 깊이 센서를 달아, 부피를 포함한 손의 3D 메시 자체를 매핑하는 기술도 연구되고 있다. 그러나, 손 추적을 위한 깊이 센서를 사용하는 것은 비용 문제 이외에도, 센서 설치 환경 및 캘리브레이션 문제로 인해 사용자에게 어려움을 줄 수 있다. 또한 손 추적만을 이용한 상호 작용은 허공에서 손을 움직이는 방식이라, 사용자에게 물리적인 피드백을 줄 수 없다. 따라서 사용자가 불편함을 느끼고 정확성이 감소되는 단점이 있다.
- [0006] 또한, VR 환경에서의 상호작용은 버튼 클릭으로 정해진 동작을 수행하거나, 추가적인 시각 센서 혹은 VR 글러브 등을 이용해 사용자 손의 움직임을 추적하는 방식으로 제한되어 있다. 이러한 방식은 상호작용의 다양성을 제한할 뿐만 아니라, 센서를 구비하기 위한 추가 비용을 야기한다.

발명의 내용

해결하려는 과제

- [0007] 따라서, 본 발명은 별도의 센서 없이도 VR 컨트롤러를 쥐고 있는 손의 포즈, VR 컨트롤러의 트리거 버튼 입력 정도, 그리고 가상 환경에 구현된 센서들이 센싱한 센싱 값에 따라 가상 손의 모션을 실시간으로 생성하여, 어떤 상황에서도 사용자가 가상 물체를 잡고 조작할 수 있도록 하는 컨트롤러를 기반으로 한 가상 손의 모션 생성 장치 및 이의 동작 방법을 제공한다.

과제의 해결 수단

- [0008] 상기 본 발명의 기술적 과제를 달성하기 위한 본 발명의 하나의 특징인 적어도 하나의 프로세서에 의해 동작하는 가상 손 모션 생성 장치의 동작 방법으로서, VR 컨트롤러를 가상 손으로 매핑하고, 상기 가상 손과 상호 작용되는 가상 물체를 가상 환경에 생성하는 단계, 상기 가상 손의 포즈, 상기 가상 손이 상기 가상 물체를 쥐는 세기, 상기 가상 물체의 정보를 포함하는 상태 정보를 획득하는 단계, 그리고 상기 상태 정보가 입력되면 상기 가상 손이 상기 가상 물체를 쥐는 모션을 수행하기 위해, 상기 가상 손을 구성하는 손가락의 관절들 각각의 관절 각도 변화량이 출력되도록 인공지능 모델을 학습시키는 단계를 포함한다.
- [0009] 상기 상태 정보를 획득하는 단계는, 상기 VR 컨트롤러를 통해 수집한 상기 가상 손의 3차원 위치, 손목 방향, 손목 회전 정도를 상기 가상 손의 포즈로 획득하는 단계를 포함할 수 있다.
- [0010] 상기 상태 정보를 획득하는 단계는, 손 모션 캡처 데이터로부터 상기 가상 물체의 쥐는 세기에 해당하는 트리거 정보를 계산하는 단계를 포함할 수 있다.

- [0011] 상기 상태 정보를 획득하는 단계는, 상기 가상 환경에 생성된 복수의 복셀 센서들의 센싱 값들을 기초로 상기 가상 물체의 정보를 획득하는 단계를 포함할 수 있다.
- [0012] 상기 복수의 복셀 센서들은 글로벌 복셀 센서, 델타 글로벌 복셀 센서, 그리고 복수의 로컬 복셀 센서들을 포함하고, 상기 글로벌 복셀 센서는 상기 가상 손의 손목에 고정되어 상기 가상 손 앞 일정 범위 내에 위치한 상기 물체를 감지하고, 상기 델타 글로벌 복셀 센서는 상기 물체의 움직임과 이동 방향을 유추하며, 상기 복수의 로컬 복셀 센서들은 상기 물체의 세밀한 굴곡을 감지할 수 있다.
- [0013] 상기 인공지능 모델을 학습시키는 단계는, 상기 상태 정보를 상기 관절 각도 변화량으로 매핑하는 함수를 이용하여 상기 인공지능 모델을 학습시킬 수 있다.
- [0014] 상기 인공지능 모델을 학습시키는 단계는, 상기 가상 손의 포즈, 상기 가상 물체를 쥐는 세기 및 상기 가상 물체의 정보들 각각의 독립적인 특징이 연결된 출력 벡터를 추출하는 단계, 상기 출력 벡터의 차원보다 고차원의 특성 벡터를 생성하는 단계, 그리고 상기 특성 벡터로부터 상기 가상 손의 각 관절 각도 변화량에 해당하는 동작 벡터를 생성하는 단계를 포함할 수 있다.
- [0015] 상기 동작 벡터를 생성하는 단계 이후에, 상기 동작 벡터로부터 상기 가상 손을 구성하는 상기 손가락의 복수의 관절들 각각의 토크를 산출하는 단계, 그리고 상기 산출한 토크를 물리 시뮬레이션에 적용하는 단계를 포함할 수 있다.
- [0016] 상기 본 발명의 기술적 과제를 달성하기 위한 본 발명의 또 다른 특징인 적어도 하나의 프로세서에 의해 동작하는 모션 생성 장치의 동작 방법으로서, 제1 프레임에서 가상 손의 포즈를 획득하는 단계, 상기 가상 손에 매핑되어 있는 VR 컨트롤러의 컨트롤러 버튼으로부터 상기 제1 프레임에 트리거 정보를 획득하고, 상기 제1 프레임을 기준으로 미리 설정된 복수의 이전 프레임들의 이전 트리거 정보를 획득하는 단계, 상기 가상 환경에 생성된 복수의 복셀 센서들의 센싱 값들을 획득하는 단계, 그리고 상기 가상 손의 포즈, 상기 제1 프레임의 트리거 정보와 이전 트리거 정보, 그리고 상기 센싱 값들을 기초로, 상기 제1 프레임의 다음 프레임인 제2 프레임에서의 상기 가상 손을 구성하는 관절들 각각의 관절 각도 변화량을 출력하는 단계를 포함한다.
- [0017] 상기 관절 각도 변화량을 출력하는 단계 이후에, 상기 관절 각도 변화량으로부터 각 관절의 토크를 산출하는 단계, 그리고 상기 산출한 토크를 물리 시뮬레이션에 적용하는 단계를 포함할 수 있다.
- [0018] 상기 복수의 복셀 센서들은 글로벌 복셀 센서, 델타 글로벌 복셀 센서, 그리고 복수의 로컬 복셀 센서들을 포함하고, 상기 글로벌 복셀 센서는 상기 가상 손의 손목에 고정되어 상기 가상 손 앞 일정 범위 내에 위치한 상기 물체를 감지하고, 상기 델타 글로벌 복셀 센서는 상기 물체의 움직임과 이동 방향을 유추하며, 상기 복수의 로컬 복셀 센서들은 상기 물체의 세밀한 굴곡을 감지할 수 있다.
- [0019] 상기 본 발명의 기술적 과제를 달성하기 위한 본 발명의 또 다른 특징인 적어도 하나의 프로세서에 의해 동작하는 모션 생성 장치로서, VR 컨트롤러로부터 트리거 정보와 포즈 정보를 수집하는 인터페이스, 그리고 프로세서를 포함하고, 상기 프로세서는, 제1 프레임에서 상기 VR 컨트롤러가 매핑된 가상 손의 포즈, 상기 제1 프레임에서의 트리거 정보와 상기 제1 프레임을 기준으로 미리 설정된 복수의 이전 프레임들의 이전 트리거 정보, 가상 환경에 생성된 복수의 복셀 센서들의 센싱 값들을 포함하는 상태 정보를 기초로, 상기 제1 프레임의 다음 프레임인 제2 프레임에서의 상기 가상 손이 취할 모션을 위해 상기 가상 손을 구성하는 관절들 각각의 관절 각도 변화량을 계산한다.
- [0020] 상기 프로세서는, 상기 상태 정보가 입력되면 상기 가상 손을 구성하는 관절들 각각의 관절 각도 변화량이 출력되도록 인공지능 모델을 학습시킬 수 있다.
- [0021] 상기 프로세서는, 상기 가상 손의 포즈, 상기 트리거 정보 및 상기 센싱 값들 각각의 독립적인 특징이 연결된 출력 벡터를 추출하고, 상기 출력 벡터의 차원보다 고차원의 특성 벡터를 생성하며, 상기 특성 벡터로부터 상기 가상 손의 각 관절 각도 변화량에 해당하는 동작 벡터를 생성할 수 있다.
- [0022] 상기 프로세서는, 상기 관절 각도 변화량으로부터 각 관절의 토크를 산출하고, 상기 산출한 토크를 물리 시뮬레이션에 적용할 수 있다.
- [0023] 상기 복수의 복셀 센서들은 글로벌 복셀 센서, 델타 글로벌 복셀 센서, 그리고 복수의 로컬 복셀 센서들을 포함하고, 상기 글로벌 복셀 센서는 상기 가상 손의 손목에 고정되어 상기 가상 손 앞 일정 범위 내에 위치한 상기 물체를 감지하고, 상기 델타 글로벌 복셀 센서는 상기 물체의 움직임과 이동 방향을 유추하며, 상기 복수의 로

컬 복셀 센서들은 상기 물체의 세밀한 굴곡을 감지할 수 있다.

발명의 효과

- [0024] 본 발명의 실시예에 따르면, 별도의 센서 없이도 VR 환경에서 시각적으로, 그리고 물리적으로 현실감 있게 가상 손의 움직임을 표현하고 가상 환경에 구현된 가상 물체를 조절할 수 있다.
- [0025] 또한, VR 컨트롤러만으로도 자유롭게 가상 물체를 잡고 움직일 수 있으며, 더욱 자연스럽고 복잡한 가상현실의 경험을 제공할 수 있다.

도면의 간단한 설명

- [0026] 도 1은 본 발명의 실시예에 따른 가상 손 모션 생성 장치가 구현된 환경의 예시도이다.
- 도 2는 본 발명의 실시예에 따른 가상 손 모션 생성 장치의 구조도이다.
- 도 3은 본 발명의 실시예에 따른 가상 손 모션 생성 장치의 동작에 대한 흐름도이다.
- 도 4는 본 발명의 실시예에 따라 구현된 복셀 센서들의 예시도이다.
- 도 5는 본 발명의 실시예에 따른 인공지능 모델의 예시도이다.
- 도 6은 본 발명의 실시예에 따라 인공지능 모델을 학습시키는 예시도이다.
- 도 7은 본 발명의 실시예에 따른 컴퓨팅 시스템의 구조도이다.

발명을 실시하기 위한 구체적인 내용

- [0027] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [0028] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0029] 이하 도면을 참조로 하여 본 발명의 실시예에 따른 컨트롤러를 기반으로 한 가상 손의 모션 생성 장치 및 이의 동작 방법에 대해 설명한다.
- [0030] 도 1은 본 발명의 실시예에 따른 가상 손 모션 생성 장치가 구현된 환경의 예시도이다.
- [0031] 도 1에 도시된 바와 같이, 가상 손 모션 생성 장치(100)는 VR 컨트롤러(200)와 연동하여, VR 컨트롤러(200)를 가상 환경에 구현된 가상 손으로 매핑한다. 그리고 가상 손 모션 생성 장치(100)는 가상 손과 상호 작용하는 가상 물체도 생성한다.
- [0032] 가상 손 모션 생성 장치(100)는 가상 손의 자세(pose), 사용자가 VR 컨트롤러(200)에 구비된 트리거 버튼의 입력에 따라 수집한 트리거 정보, 가상 손에 구현된 복수의 복셀 센서들이 각각 수집한 센싱 값을 기초로 가상 손의 모션을 결정한다.
- [0033] 디스플레이(300)는 가상 손 모션 생성 장치(100)에서 결정한 가상 손의 모션으로 동작하는 가상 손의 움직임 또는 가상 물체의 움직임을 디스플레이하여, VR 컨트롤러(200)를 동작시키는 사용자가 확인할 수 있도록 한다.
- [0034] 본 발명의 실시예에 따른 가상 손 모션 생성 장치(100)는 가상 손이 가상 물체에 접근하거나 사용자의 트리거 버튼 입력 정도에 따라, 동적으로 가상 손의 손목 또는 손가락의 움직임을 합성하도록 딥러닝 기술 기반의 인공지능 모델(VR-HandNet)을 포함하여 구현되는 것을 예로 하여 설명한다.
- [0035] 인공지능 모델은 매 프레임마다 가상 환경에 대한 가상 손의 포즈, 가상 물체의 상태, 사용자가 VR 컨트롤러(200)의 트리거 버튼을 눌렀는지, 눌렀다면 얼마나 길게 또는 얼마나 세게 눌렀는지에 따른 트리거 정보, 그리고 가상 환경에 구현된 복셀 센서들의 센싱 정보를 입력으로 받는다. 그리고 인공지능 모델은 입력을 기초로, 다음 프레임에서의 가상 손의 손목 또는 손가락 관절의 물리적 움직임을 출력으로 산출한다.

- [0036] 여기서, 가상 손 모션 생성 장치(100)는 가상 손의 포즈를 이용하여 VR 컨트롤러(200)를 쥐고 있는 사용자 손의 움직임에 파악할 수 있다. 그리고 가상 손 모션 생성 장치(100)는 트리거 정보를 기초로 가상 물체를 쥐는 세기를 파악할 수 있다.
- [0037] 이를 위해, VR 컨트롤러(200)는 사용자의 움직임에 따라 3차원 위치, 움직임 속도, 각속도 등의 손의 정보를 이미 알려진 기술로 수집할 수 있으며, 본 발명의 실시예에서는 어느 하나의 방법으로 한정하지 않는다. 또한, VR 컨트롤러(200)는 사용자가 트리거 버튼을 어느 시점에 누르기 시작하였는지, 얼마나 길게 누르고 있는지 등을 트리거 정보로 수집하는 방법 역시 이미 알려진 기술을 활용할 수 있으므로, 본 발명의 실시예에서는 어느 하나의 방법으로 한정하지 않는다.
- [0038] 본 발명의 실시예에 따른 가상 손 모션 생성 장치(100)는 사용자가 손가락을 얼마나 움직일지 사용자 의도를 파악하기 위해 트리거 정보를 사용한다. 인공지능 모델은 복셀 센서 값을 기초로 가상 환경의 현재 상황을 파악하고, 트리거 정보를 이용하여 사용자의 의도를 알아낼 수 있다.
- [0039] 가상 손 모션 생성 장치(100)는 사용자가 VR 컨트롤러(200)를 쥐고 물리적으로 움직임이 있는 움직임을 가상 손을 제어하는 물리 시뮬레이션에 적용하여, 다음 프레임에서 가상 손의 최종 모션을 결정한다. 즉, 인공지능 모델은 복셀 센서 값과 트리거 정보를 기초로 매 프레임마다 반복하여 손목 또는 손가락의 물리적 움직임 즉, 가상 손을 구성하는 복수의 관절들 각각의 관절 움직임 각도를 산출하고, 이를 기초로 가상 손의 모션을 결정함으로써 가상 손의 움직임을 실시간으로 생성한다.
- [0040] 이를 위해, 인공지능 모델은 훈련 데이터를 이용하여, 가상현실 환경에서 가상 손과 가상 물체 간 상호작용이 어떻게 이루어지는지를 강화학습 기반으로 학습한다. 그리고 인공지능 모델은 현실의 데이터 즉, 손 모션 캡처 데이터의 움직임을 모사하도록 학습되어 있다. 이를 통해, 가상 손 모션 생성 장치(100)는 현실과 같이 시각적, 물리적으로 자연스러운 가상 손의 모션을 생성할 수 있다. 본 발명의 실시예에서는 손 모션 캡처 데이터를 핸드 트래킹 방식으로 수집하는 것을 예로 하여 설명하나, 광학식 모션 캡처나 트래킹 글로브(glove) 등 다양한 손 모션 캡처 방식을 통해 손 모션 캡처 데이터를 수집할 수 있으며, 어느 하나의 방법으로 한정하지 않는다.
- [0041] 가상 손 모션 생성 장치(100)는 가상 손을 구성하는 복수의 관절들(예를 들어, 손목 관절, 손가락 관절 등)의 방향과 관절 각도 변화량을 결정하기 위해, VR 환경에서 세 가지 복셀(Voxel) 센서를 이용한다. 여기서, 복셀은 2차원 이미지에서 2차원 이미지를 구성하는 최소 단위(picture element)인 픽셀(pixel)을 3차원으로 확장한 것을 의미하며, 2차원 이미지에 비하여 깊이 정보를 더 포함하도록 결정되는 것을 의미한다.
- [0042] 본 발명의 실시예에서는 가상 손 모션 생성 장치(100)가 가상 손에 설정한 세 가지 복셀 센서로 글로벌 복셀 센서(GVS: Global Voxel Sensor), 델타 글로벌 복셀 센서(DGVS: Delta Global Voxel Sensor), 그리고 로컬 복셀 센서(LVS: Local Voxel Sensor)인 것을 예로 하여 설명한다. 그러나, 반드시 이와 같이 한정되는 것은 아니며, 복셀 센서의 종류 또한 이와 같이 한정되는 것은 아니다.
- [0043] 가상 손 모션 생성 장치(100)는 글로벌 복셀 센서를 가상 손의 손목에 고정되도록 구현한다. 이를 통해 글로벌 복셀 센서는 가상 손 앞에 미리 설정된 범위(예를 들어, 18cm 등)의 물체를 감지할 수 있다. 글로벌 복셀 센서는 물체의 대략적인 실루엣은 감지할 수 있으나, 물체의 세밀한 굴곡이나 물체의 움직임은 감지할 수 없다.
- [0044] 따라서, 가상 손 모션 생성 장치(100)는 연속된 프레임에서 공간 관계의 일차 도함수를 센싱하도록 델타 글로벌 복셀 센서를 구현한다. 델타 글로벌 복셀 센서는 물체의 물리적 속성, 예를 들어, 물체의 이동 속도나 각속도를 추정하는 것이다.
- [0045] 이를 위해, 가상 손 모션 생성 장치(100)는 글로벌 복셀 센서의 위치를 제1 시점(t 시점)의 손목 위치로 설정한다. 그러면, 델타 글로벌 복셀 센서는 제1 시점에서의 글로벌 복셀 센서 값과 제2 시점($t+1$ 시점)에서의 글로벌 복셀 센서 값의 차를 계산함으로써, 물체의 움직임을 포착한다.
- [0046] 델타 글로벌 복셀 센서의 센서 값은 제1 시점과 제2 시점의 글로벌 복셀 센서의 센서 값의 차이로부터 구할 수 있다. 즉, 델타 글로벌 복셀 센서의 센서 값은 글로벌 복셀 센서가 센싱한 값의 프레임 간 변화량을 의미한다.
- [0047] 따라서, 델타 글로벌 복셀 센서는 두 프레임 사이에 물체가 빠져나가는 복셀은 음수로, 물체가 더 겹쳐지게 된 복셀은 양수로 표현한다. 양수 또는 음수로 표현한 정보를 기초로, 가상 손 모션 생성 장치(100)는 물체의 방향을 포함한 움직임을 유추할 수 있다.
- [0048] 여기서, 제2 시점의 글로벌 복셀 센서 값은 제1 시점의 글로벌 복셀 센서의 위치 즉, 제1 시점의 손목의 위치에서 센싱된다. 그 이유는, 제1 시점에서 제2 시점으로 손목 위치가 이동하면 손목에 고정된 글로벌 복셀 센서의

위치 또한 이동하기 때문에, 물체의 움직임만을 포착하지 못하고 글로벌 복셀 센서 값에 손목의 움직임이 포함되기 때문이다.

- [0049] 따라서, 델타 글로벌 복셀 센서는 제1 시점의 위치에 고정된 글로벌 복셀 센서의 제1 시점 센싱 값과 제2 시점 센싱 값의 차를 계산함으로써, 제1 시점에서 출력한 움직임으로 제2 시점에 물체가 어떻게 움직였는지를 포착할 수 있다.
- [0050] 또한, 가상 손 모션 생성 장치(100)는 복수의 로컬 복셀 센서들을 가상 손에서 미리 설정한 접촉 위치에 구현한다. 여기서 접촉 위치나 로컬 복셀 센서의 수를 어느 하나로 한정하지 않는다. 이를 통해, 가상 손 모션 생성 장치(100)는 글로벌 복셀 센서가 감지하지 못하는 물체의 세밀한 굴곡을 로컬 복셀 센서를 이용하여 감지할 수 있다.
- [0051] 이와 같이 가상 손의 위치나 가상 손의 움직임을 제공하는 가상 손 모션 생성 장치(100)의 구조에 대해 도 2를 참조로 설명한다.
- [0052] 도 2는 본 발명의 실시예에 따른 아바타 손 상호작용 모션 생성 장치의 구조도이다.
- [0053] 도 2에 도시된 바와 같이, 가상 손 모션 생성 장치(100)는 인공지능 모델(110)과 비례 미분 제어기(Proportional Derivative Controller)(120) 그리고 물리 시뮬레이터(130)를 포함한다. 또한, 가상 손 모션 생성 장치(100)는 상술한 구조 이외에 가상 손의 상호작용 모션을 생성하기 위한 기타의 구성 요소들을 추가로 포함할 수 있다.
- [0054] 인공지능 모델(110)은 상태 정보(state)를 동작(action)으로 매핑하는 함수를 통해 훈련 데이터를 이용하여 학습되어 있다. 인공지능 모델(110)은 인코더 블록(111), 텐스 블록(112), 그리고 디코더 블록(113)으로 구현되어 있는 것을 예로 하여 설명한다.
- [0055] 인코더 블록(111)은 손의 포즈, VR 컨트롤러(200)의 트리거 버튼 눌림에 따른 트리거 정보, VR 환경에 구현된 복셀 센서들이 수집한 센싱 정보를 포함하는 상태 정보를 그룹화하여, 가상 손과 물체의 상호 작용에 대한 유사한 의미를 가지는 특징들을 함께 그룹화시킨다. 여기서, 유사한 의미를 가지는 특징이라 함은, 인코더 블록(111)이 처리하는 입력들의 정보에 따라 입력들을 그룹화하기 위한 기준을 의미한다.
- [0056] 즉, 인코더 블록(111)이 처리하는 그룹화 된 입력들은 손의 포즈, 트리거 버튼 눌림 정도, 그리고 복셀 센서들이 센싱한 센싱 정보로 구분된다. 여기서 손의 포즈, 트리거 버튼 눌림 정도, 센싱 정보가 유사한 의미를 가지는 특징을 나타낸다.
- [0057] 손의 포즈는 손의 현재 모습과 움직임을 표현하는 정보이고, 트리거 버튼 눌림 정도는 사용자의 의도를 나타내는 정보이다. 그리고 복셀 센서들이 센싱한 센싱 정보는 공간 상의 현재 모습과 움직임을 표현하는 정보를 의미한다.
- [0058] 본 발명의 실시예에 따른 인공지능 모델은 전체 입력의 모든 관계를 처음부터 추론하지 않고, 단계별로 먼저 서로 관련 있는 입력들끼리 그룹을 만든 후 처리하게 구현한다. 이와 같이 그룹화 된 입력들을 유사한 의미를 가지는 특징이라 볼 수 있다.
- [0059] 그리고, 인코더 블록(111)은 상태 정보를 구성하는 각 그룹을 특징 벡터의 세트로 인코딩하여 복수의 정보들을 하나의 특성(feature)으로 합쳐(concat), 출력 벡터를 생성한다.
- [0060] 여기서, 인코더 블록(111)으로 입력되는 상태 정보는 $\{H_t, B_t, V_t\}$ 으로 표현할 수 있으며, H_t 는 손의 포즈에 대한 정보(이하, '손 포즈 정보'라 지칭함), B_t 는 트리거 정보, V_t 는 복셀 센서들의 센싱 값(이하, '복셀 센서 센싱 값'이라 지칭함)을 의미한다. 상태 정보는 매 프레임마다 수집되어 신경망인 인공지능 모델의 인코더 블록(111)으로 입력된다.
- [0061] 손 포즈 정보 H_t 는 사용자가 사용하는 VR 컨트롤러(200)의 움직임, 가상 손이 현재 취하고 있는 포즈와 움직임으로, VR 컨트롤러(200)의 속도와 각속도, 그리고 가상 손의 3차원 위치, 속도, 각속도 등의 정보를 포함한다. 가상 손 모션 생성 장치(100)가 손 포즈 정보를 수집하는 방법은 다양하므로, 본 발명의 실시예에서는 어느 하나의 방법으로 한정하지 않는다.
- [0062] 트리거 정보 B_t 는 현재 프레임인 제1 프레임에 들어온 트리거 버튼 입력 정보와 제1 프레임을 기준으로 미리 설정된 이전 프레임(예를 들어, 과거 10 프레임 등)의 트리거 버튼 입력 정보를 포함한다. 트리거 정보 B_t 는 사용

자의 상호작용 의도, 예를 들어, VR 컨트롤러(200)를 쥐는 정도, VR 컨트롤러(200)에 가해지는 힘의 정도 등을 대변하는 것으로, 가상 손이 가상 물체를 쥐는 힘이라 볼 수 있다.

- [0063] 복셀 센서 센싱 값 V_i 는 현재 프레임인 제1 프레임에 복셀 센서들이 각각 센싱한 센싱 값을 의미한다.
- [0064] 텐스 블록(112)은 인코더 블록(111)에서 생성한 출력 벡터를 입력으로 받아, 출력 벡터의 차원보다 고차원의 특성 벡터를 생성한다. 이를 위해, 텐스 블록(112)은 ELU(Exponential Linear Units) 활성화 기능을 갖춘 4개의 완전 연결(Fully Connected) 레이어들로 구성된다.
- [0065] 텐스 블록(112)은 과적합을 완화하기 위해 4개의 완전 연결 레이어들 앞에 일정 수치의 유지 확률을 갖는 드롭아웃을 적용한다. 텐스 블록(112)의 기능은 이미 알려진 것으로 본 발명의 실시예에서는 상세한 설명을 생략한다.
- [0066] 디코더 블록(113)은 텐스 블록(112)에서 생성된 특성 벡터를 입력으로 받아, 손 관절 각도의 목표 변화량(a_t)에 해당하는 동작 벡터를 생성한다. 동작 벡터는 손가락 관절마다 복수의 목표 변화량들을 포함한다. 이를 위해, 디코더 블록(113)은 단일 완전 연결 계층으로 구성된다.
- [0067] 현재 프레임의 가상 손의 관절 각도에 디코더 블록(113)에서 생성한 목표 변화량(a_t) 만큼 더하면, 다음 프레임의 가상 손의 관절 각도가 된다. 이렇게 획득한 다음 프레임의 관절 각도를 바로 가상 손에 반영하는 것 보다, 비례 미분 제어기(120)에 입력하여 목표 변화량(a_t)을 토크의 형태로 변환하여 적용하는 것이 보다 안정적이다.
- [0068] 즉, 비례 미분 제어기(120)는 현재 프레임인 제1 프레임(frame t)에서 가상 손의 각 관절의 토크를 산출하고, 산출한 각 관절의 토크를 물리 시뮬레이터(130)에 적용하면 제2 프레임(frame t+1)의 가상 손 포즈가 결정된다. 이렇게 물리 시뮬레이터(130)가 물리 시뮬레이션 과정을 수행하도록 함으로써, 가상 손의 자연스러운 애니메이션을 출력할 뿐만 아니라, 물리적인 상호 작용도 가능하다.
- [0069] 다음은 가상 손 모션 생성 장치(100)의 동작에 대해 도 3을 참조로 설명한다.
- [0070] 도 3은 본 발명의 실시예에 따른 가상 손 모션 생성 장치의 동작에 대한 흐름도이다.
- [0071] 도 3에 도시된 바와 같이, 가상 손 모션 생성 장치(100)는 VR 컨트롤러(200)가 매핑된 가상 손을 생성한다(S100). 가상 손 모션 생성 장치(100)가 가상 손을 생성하는 방법은 다양하므로, 본 발명의 실시예에서는 어느 하나의 방법으로 한정하지 않는다.
- [0072] 이때, 가상 손 모션 생성 장치(100)는 글로벌 복셀 센서를 가상 손의 손목에 고정되도록 설정하여, 가상 손 앞에 미리 설정된 범위의 가상 물체를 감지하도록 한다. 또한, 가상 손 모션 생성 장치(100)는 가상 물체의 움직임으로 물체의 방향과 물체의 이동 속도, 각속도 등을 유추하도록 델타 글로벌 복셀 센서를 설정한다. 그리고, 가상 손 모션 생성 장치(100)는 가상 손의 여러 접촉 위치에 복수의 로컬 복셀 센서들이 부착되도록 생성한다.
- [0073] 가상 손 모션 생성 장치(100)는 글로벌 복셀 센서와 복수의 로컬 복셀 센서들이 부착된 가상 손이 임의의 가상 물체를 쥐거나 움직이는 등 가상 손과 상호 작용할 가상 물체를 가상 환경에 생성한다(S110). 본 발명의 실시예에서는 다양한 상호 작용 중 가상 손이 가상 물체를 쥐는 것을 예로 하여 설명한다. 여기서, 가상 물체의 모양이나 크기, 그리고 가상 손이 가상 물체를 쥐고 있도록 하거나 가상 물체 근처로 이동하는 등의 다양한 모션들을 생성하는 것은 다양한 방법으로 생성할 수 있으므로, 본 발명의 실시예에서는 어느 하나의 방법으로 한정하지 않는다.
- [0074] 가상 손 모션 생성 장치(100)는 가상 물체를 쥐는 가상 손의 포즈, 복셀 센서들이 수집한 센싱 정보를 수집하여, 상태 정보 즉, 훈련 데이터의 입력 정보로 생성한다(S120). 가상 손의 포즈는 3차원 위치, 손목 방향과 손목 회전 정도 등을 포함할 수 있으며, VR 컨트롤러(200)를 통해 수집할 수 있다.
- [0075] 이때, 가상 손 모션 생성 장치(100)는 손의 모션을 캡처한 손 모션 캡처 데이터로부터 트리거 정보를 계산한다. 본 발명의 실시예에서는 가상 손 모션 생성 장치(100)가 계산한 트리거 정보도 입력 정보에 포함시키는 것을 예로 하여 설명하나, 반드시 이와 같이 한정되는 것은 아니다.
- [0076] 트리거 정보는 가상 손이 가상 물체를 쥐는 세기에 해당하는 정보로 사용될 수 있다. 또한, 가상 손 모션 생성 장치(100)는 복셀 센서들이 각각 센싱한 값들도 수집하여 훈련 데이터의 입력 정보에 포함시킬 수 있다.
- [0077] 여기서, 인공지능 모델을 훈련시키는 훈련 데이터를 수집하기 위해, 가상 손 모션 생성 장치(100)는 손 모션 캡

처 데이터로부터 트리거 정보를 계산하여 손 모션 캡처 데이터와 사용자 의도인 트리거 정보를 매핑하여 훈련 데이터 셋으로 만든다. 그리고 가상 손 모션 생성 장치(100)는 트리거 정보가 매핑된 손 모션 캡처 데이터를 인공지능 모델을 훈련시키는 데 사용한다.

- [0078] 손 모션 캡처 상황에선 VR 컨트롤러(200)를 사용하는 것이 아니라, 실제 손의 움직임을 트래킹하는 것이기 때문에, 손 모션 캡처 데이터를 수집하는 상황에서 트리거 정보를 같이 수집할 수 없다. 따라서, 가상 손 모션 생성 장치(100)는 수집한 손 모션이 어떤 트리거 정보를 가져야 하는지를 따로 결정해주어야 하며, 이를 위해 트리거 매핑 식이 사용된다. 여기서 트리거 매핑 식은 어느 하나의 수식으로 한정하지 않는다.
- [0079] 가상 손 모션 생성 장치(100)가 손 모션 캡처 데이터의 움직임만을 보고 VR 컨트롤러(200)를 사용하는 사용자가 손 모션 캡처 데이터의 움직임과 동일한 모션을 생성하려면 어떤 트리거 눌림 정도가 가장 자연스럽게 적절할지 파악하기 위해, 가상 손 모션 생성 장치(100)는 트리거 매핑 식을 통해 매 프레임 계산하여 손 모션 캡처 데이터의 매 프레임마다 라벨링하는 것을 예로 하여 설명한다.
- [0080] 그리고, 가상 손 모션 생성 장치(100)는 가상 손의 모양을 훈련 데이터의 출력 정보로 생성한다(S130). 본 발명의 실시예에서는 가상 손 모션 생성 장치(100)가 가상 손의 모양으로써 가상 손을 구성하는 관절들 각각의 관절 각도 변화량을 출력하는 것을 예로 하여 설명한다.
- [0081] 가상 손 모션 생성 장치(100)는 S120 및 S130 단계를 통해 생성한 훈련 데이터로 가상 손 모션 생성 장치(100)의 인공지능 모델을 학습시키거나 훈련시킨다(S140). 즉, 가상 손 모션 생성 장치(100)는 손의 포즈, 손 모션 캡처 데이터로부터 결정된 트리거 정보, VR 환경에 구현된 복셀 센서들이 수집한 센싱 정보를 포함하는 상태 정보가 입력되면, 가상 손이 가상 물체를 쥐는 모양인 모션을 취하도록 관절 각도 변화량을 출력하도록 인공지능 모델을 학습시킨다.
- [0082] 이와 같이 인공지능 모델이 학습되면, 가상 손 모션 생성 장치(100)는 학습시킨 인공지능 모델을 이용하여 프레임마다 가상 손과 VR 컨트롤러(200)의 상태 정보가 입력되면, 입력된 상태 정보에 대응하는 가상 손의 모션을 생성하여 디스플레이(300)를 통해 제공한다(S150).
- [0083] 다음은, VR 환경에서 사용되는 복셀 센서들에 대해 도 4를 참조로 설명한다.
- [0084] 도 4는 본 발명의 실시예에 따라 구현된 복셀 센서들의 예시도이다.
- [0085] 도 4의 (a)에 도시된 바와 같이, 글로벌 복셀 센서는 일정 크기의 정육면체가 복수의 복셀 개수로 나누어진 형태의 센서다. 글로벌 복셀 센서는 가상 손의 손목에 고정되어 있으므로, 가상 손 앞의 일정 범위 내의 물체를 감지한다. 이때, 글로벌 복셀 센서는 가상 물체의 대략적인 실루엣을 알 수 있지만, 가상 손의 손가락의 정밀한 움직임을 제공할 수 있을 정도의 가상 물체의 세밀한 굴곡이나 가상 물체의 움직임은 알 수 없다.
- [0086] 따라서, 델타 글로벌 복셀 센서는 도 4의 (b)에 도시된 바와 같이 물체의 움직임을 감지한다.
- [0087] 즉, 글로벌 복셀 센서는 가상 물체의 어떤 한 시점의 고정된 모습을 포착하기 때문에, 가상 물체 움직임의 방향성을 추측할 수는 없다. 그러므로, 델타 글로벌 복셀 센서는 글로벌 복셀 센서의 위치를 제1 시점의 손목 위치로 고정하고, 제1 시점에서의 글로벌 복셀 센서 값(이하, '제1 센서 값'이라 지칭함)에서 제2 시점에서 글로벌 복셀 센서 값(이하, '제2 센서 값'이라 지칭함)의 차이(제1 센서 값 - 제2 센서 값)을 계산하여 가상 물체의 움직임과 방향성을 추측하는데 사용한다.
- [0088] 즉, 델타 글로벌 복셀 센서는 글로벌 복셀 센서가 고정된 위치에서 가상 물체의 움직임을 포착한다. 이를 통해, 델타 글로벌 복셀 센서는 두 프레임 사이에 가상 물체가 빠져나가는 복셀은 음수로, 가상 물체가 더 겹쳐지게 된 복셀은 양수로 표현한다. 이 정보로 가상 물체의 방향을 포함한 움직임 즉, 가상 물체 움직임 속도나 각속도 등을 유추할 수 있다.
- [0089] 한편, 도 4의 (c)에 도시된 바와 같이, 로컬 복셀 센서는 글로벌 복셀 센서가 감지하지 못하는 가상 물체의 세밀한 굴곡을 감지한다. 본 발명의 실시예에서는 정육면체 형태로 총 19개의 로컬 복셀 센서들이 가상 손의 각 접촉 위치에서 센싱하는 것을 예로 하여 설명한다.
- [0090] 그리고, 로컬 복셀 센서들에 의해 형성되는 복셀들의 수는 $5*5*5$ 복셀 개수로, 로컬 복셀 센서에 의한 하나의 복셀 크기는 글로벌 복셀 센서의 복셀 크기보다 작은 크기로 구현된다. 따라서, 글로벌 복셀 센서의 복셀이 감지하지 못한 물체의 더 세밀한 굴곡을 감지할 수 있다.
- [0091] 글로벌 복셀 센서가 너무 작은 크기의 복셀을 가진다면 신경망이 특정 물체의 굴곡 자체를 기억해버려 과적합

(Overfitting)이 발생할 수 있다. 따라서, 글로벌 복셀 센서는 가상 물체의 전체 실루엣을 파악하게 하고, 로컬 복셀 센서가 지역적인 정보만을 감지하도록 하여 가상 물체의 세밀한 굴곡을 파악하게 한다. 여기에, 델타 글로벌 복셀 센서를 추가해 가상 물체의 전반적인 움직임을 포착할 수 있도록 한다.

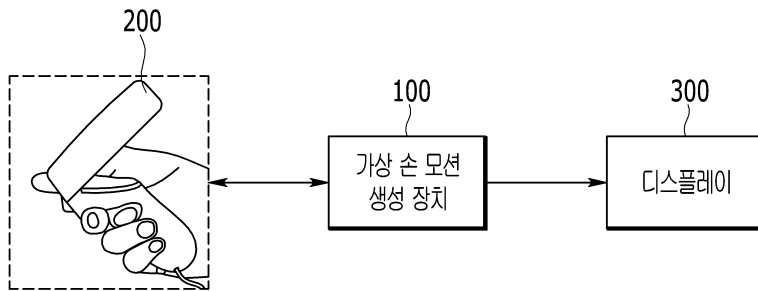
- [0092] 이를 위해, 인공지능 모델은 훈련 데이터와 똑같은 출력을 내도록 학습된다. 즉, 훈련 데이터인 모션 데이터를 수집할 때 85~250 프레임 정도의 시퀀스 데이터가 수집된다. 모션 데이터를 사용해 제1 시점의 프레임에서 어떤 출력을 내야 모션 데이터의 제2 시점의 프레임과 동일한 상황을 만들어 낼 수 있는지를 인공지능 모델을 학습시킨다.
- [0093] 여기서, 얼마나 동일한 상황인지는 보상 함수로 수치화되며, 보상 함수로의 수치화는 이미 알려진 기술로 본 발명의 실시예에서는 상세한 설명을 생략한다. 보상 함수를 최대화할수록 주어진 훈련 데이터와 동일한 손, 물체 움직임을 보이도록 인공지능 모델이 학습된다. 보상 함수를 최대화하려면 인공지능 모델이 변화해야 하기 때문에, 최적화 과정에서 인공지능 모델이 보상 함수를 최대화하도록 최적화된다.
- [0094] 다음은 가상 손 모션 생성 장치(100)에 구현된 인공지능 모델에 대해 도 5를 참조로 설명한다.
- [0095] 도 5는 본 발명의 실시예에 따른 인공지능 모델의 예시도이다.
- [0096] 도 5에 도시된 바와 같이, 본 발명의 실시예에서는 상태 정보를 목표 변화량으로 매핑하는 함수를 이용하여, 인공 신경망 인공지능 모델을 학습시킨다. 도 5에 도시된 인공지능 모델의 구조와 같이, 인공지능 모델은 우선 입력 특징을 구분하여 처리한다.
- [0097] 여기서, 상태 정보는 219차원의 손 포즈(H_t), 11차원의 트리거 정보(B_t), 1,000차원의 글로벌 복셀 센서의 센싱 값과 1,000 차원의 델타 글로벌 복셀 센서의 센싱 값, 그리고 2,375차원의 로컬 복셀 센서의 센싱 값을 포함하는 센싱 값(V_t)을 포함한다. 여기서, V_t 에 포함된 각 복셀 센서들의 센싱 값은 구분되어 처리된다. 본 발명의 실시예에서는 상태 정보에 포함된 정보들의 차원을 219차원, 1,000차원, 2,375차원 등으로 상술한 바와 같이 나타내었으나, 반드시 이와 같이 한정되는 것은 아니다.
- [0098] 이렇게 하나의 입력 특징에 세 가지 정보를 포함하는 이유는, 서로 연관된 의미를 가지는 정보끼리 묶어서 1차적인 특징 즉, 손의 포즈 해석, 사용자 의도, 공간 해석을 1차적인 특징으로 기준하여 입력들을 그룹핑한 후 출력 벡터를 추출하기 위함이다. 인코더 블록(111)의 역할이 상태 정보에서 1차적인 특징을 출력 벡터로 추출하는 것으로, 세 가지 정보를 각각 구분하여 독립적인 신경망 레이어들에서 각각 처리한 뒤, 하나의 특성(feature)으로 이어 붙이는 과정을 수행한다.
- [0099] 인코더 블록(111)을 통과하면서 2,059 차원으로 이어진 출력 벡터는 연속된 텐스 레이어(dense layer)들이 연결된 텐스 블록(112)으로 입력된다.
- [0100] 인코더 블록(111)의 출력 벡터는 H_t , B_t , V_t 의 정보를 모두 고려하여 특징을 추출한 것이 아니라 각각의 독립적인 특징을 추출한 것이다. 따라서, 각각의 특징을 하나로 이어 생성된 출력 벡터에서 최종적인 특성 벡터를 산출해야 한다. 텐스 블록(112)은 출력 벡터에서 특성 벡터를 생성하며, 텐스 블록(112)을 통과해 출력된 2,059 차원의 특성 벡터는 디코더 블록(113)을 통해 2,059차원에서 최종 출력인 25차원의 목표 변화량 즉, 동작 벡터로 출력된다.
- [0101] 여기서, 25차원은 가상 손 관절의 DoF(Degrees of Freedom) 개수를 나타낸다. 인공지능 모델의 출력인 목표 변화량은 비례 미분 제어기(120)로 입력되고 토크로 변환되어, 가상 손의 각 관절의 DoF에 적용된 후 물리 시뮬레이터(130)에서 물리 시뮬레이션이 수행된다.
- [0102] 여기서, 인공지능 모델을 학습시키는 예에 대해 도 6을 참조로 설명한다
- [0103] 도 6은 본 발명의 실시예에 따라 인공지능 모델을 학습시키는 예시도이다.
- [0104] 도 6의 (a)에 도시된 바와 같이, 본 발명의 실시예에서는 가상 손이 5개의 DoF와 함께 15개의 관절을 가지고 있고, 각 관절은 가상 손과 가상 물체 사이의 충돌을 감지하기 위해 캡슐 콜라이더 또는 박스 콜라이더(㉠)를 가지고 있다.
- [0105] 또한, 도 6의 (b)와 같이 기본 물체와 복합 물체로 분류할 수 있는 총 8가지 유형의 물체를 이용하여 데이터를 수집하는 것을 예로 하여 설명한다. 모든 물체는 밀도가 2.66g/m³인 균일한 강체로 구현되어 있으며, 밀도를 결정하기 위해 원시 물체와 동일한 모양을 가진 여러 개의 실제 플라스틱 물체를 얻었고 평균 밀도를 측정하여 학

습에 사용한다.

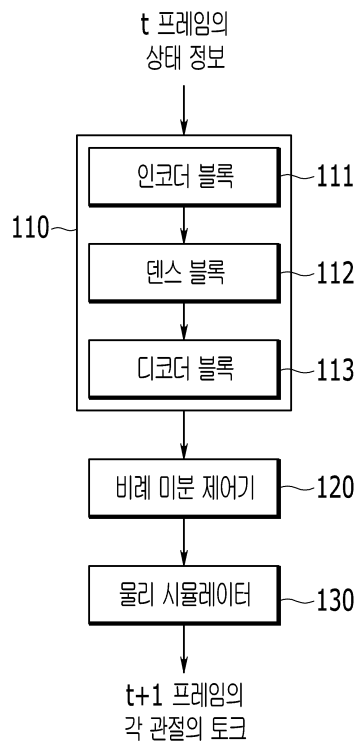
- [0106] 본 발명의 실시예에서는 인공지능 모델은 강화학습의 PPO(Proximal Policy Optimization) 알고리즘을 이용하여 학습되는 것을 예로 하여 설명한다. 강화학습에는 각 상태(state)-행동(action) 마다 얼마나 행동을 잘 산출했는지 평가하는 리워드 함수(reward function)가 존재한다. 이 리워드 값을 PPO에 입력하면, PPO 알고리즘은 리워드 값을 최대화하는 최적화(optimization) 문제를 풀어, 리워드 함수의 구성이 신경망의 학습 방향을 결정한다.
- [0107] 본 발명의 실시예에서는 리워드 함수의 구성을 위해 모조 학습(imitation learning) 방법론을 적용하였다. 이는 사전에 준비된 참조 데이터를 목표로 지정하고, 그와 얼마나 비슷한 결과를 만들어내는지에 따라 높은 리워드를 책정하는 방법론이다.
- [0108] 따라서, 레퍼런스 데이터와 그것을 평가하는 리워드 함수 모두에 신경망이 학습하고자 하는 목적이 충분히 반영되어야 한다. 본 발명의 실시예에서 인공지능 모델을 학습시킬 때 준비한 참조 데이터는 임의의 프레임(frame t)의 입력이 되는 상태와 그에 대한 최종 결과인 물체와 손의 포즈를 포함한다. 즉, 인공지능 모델은 참조 데이터와 같은 상태를 입력으로 하여, 참조 데이터와 물체 및 손의 포즈를 산출한다.
- [0109] 본 발명의 실시예에서는 물체와 손의 포즈 결과를 얼마나 잘 따라하였는지를 반영하도록 리워드 함수를 이용한다. 이때, 리워드 함수는 참조 데이터의 물체 포즈를 얼마나 잘 따라하였는지는 대상 가상 물체의 위치와 각도 차이를 기준으로 점수화하여 산출한다.
- [0110] 그리고, 참조 데이터의 손 포즈를 얼마나 잘 따라하였는지는 손 관절의 위치, 각속도, 그리고 각 다섯 손가락 끝의 위치 차이를 기준으로 점수화하여 산출한다.
- [0111] 가상 손 모션 생성 장치(100)의 인공지능 모델은 물체 포즈 기준 차이와 가중치를 합하고 손 포즈 기준 차이와 가중치를 합하여, 각각의 합이 1이 되도록 가중치를 적용한다. 최종적으로, 1이 된 물체 포즈 기준 차이 합과 손 포즈 기준 차이를 곱하여 최종 리워드로 사용한다.
- [0112] 이렇게 구성된 리워드 함수의 값을 PPO 알고리즘에 입력하여 최적화하면, 해당 리워드 함수를 최대화하도록 인공지능 모델의 신경망 파라미터들이 학습된다. 실제 사람의 참조 데이터처럼 가상 물체와 가상 손 포즈가 얼마나 잘 따라하는지의 의미를 리워드 함수에 담기 때문에, 리워드 함수를 최대화하도록 학습된 인공지능 모델은 사람의 상호작용 방식처럼 물체와 자연스럽게 상호작용하게 된다.
- [0113] 도 7은 본 발명의 실시예에 따른 컴퓨팅 시스템의 구조도이다.
- [0114] 도 7을 참고하면, 적어도 하나의 프로세서에 의해 동작하는 아바타 손 상호작용 가상 손 모션 생성 장치(100)는 컴퓨팅 시스템(400)으로 구현될 수 있으며, 컴퓨팅 시스템(400)에서 본 발명의 동작을 실행하도록 기술된 명령들(instructions)이 포함된 프로그램을 실행한다. 프로그램은 컴퓨터 판독 가능한 저장매체에 저장될 수 있고, 유통될 수 있다.
- [0115] 컴퓨팅 시스템(400)의 하드웨어는 적어도 하나의 프로세서(410), 메모리(420), 스토리지(430), 통신 인터페이스(440)를 포함할 수 있고, 버스를 통해 연결될 수 있다. 이외에도 입력 장치 및 출력 장치 등의 하드웨어가 포함될 수 있다. 컴퓨팅 시스템(400)은 프로그램을 구동할 수 있는 운영 체제를 비롯한 각종 소프트웨어가 탑재될 수 있다.
- [0116] 프로세서(410)는 컴퓨팅 시스템(400)의 동작을 제어하는 장치로서, 프로그램에 포함된 명령들을 처리하는 다양한 형태의 프로세서일 수 있고, 예를 들면, CPU(Central Processing Unit), MPU(Micro Processor Unit), MCU(Micro Controller Unit), GPU(Graphic Processing Unit) 등 일 수 있다.
- [0117] 메모리(420)는 본 발명의 동작을 실행하도록 기술된 명령들이 프로세서(410)에 의해 처리되도록 해당 프로그램을 로드한다. 메모리(420)는 예를 들면, ROM(read only memory), RAM(random access memory) 등 일 수 있다. 스토리지(430)는 본 발명의 동작을 실행하는데 요구되는 각종 데이터, 프로그램 등을 저장한다. 통신 인터페이스(440)는 유/무선 통신 모듈일 수 있다.
- [0118] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

도면

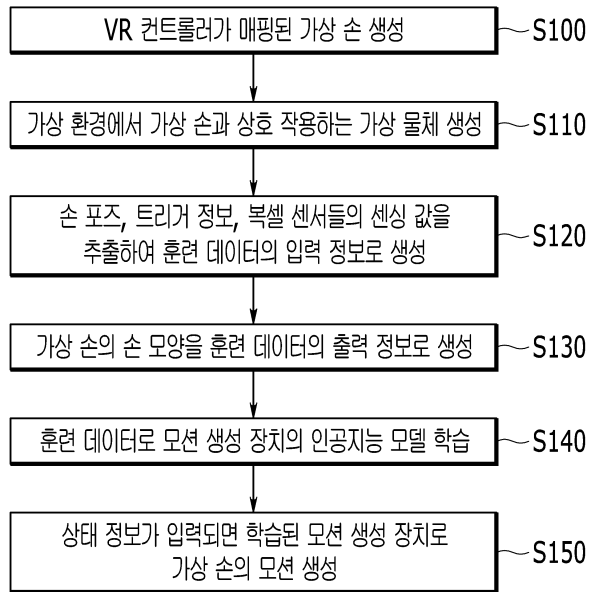
도면1



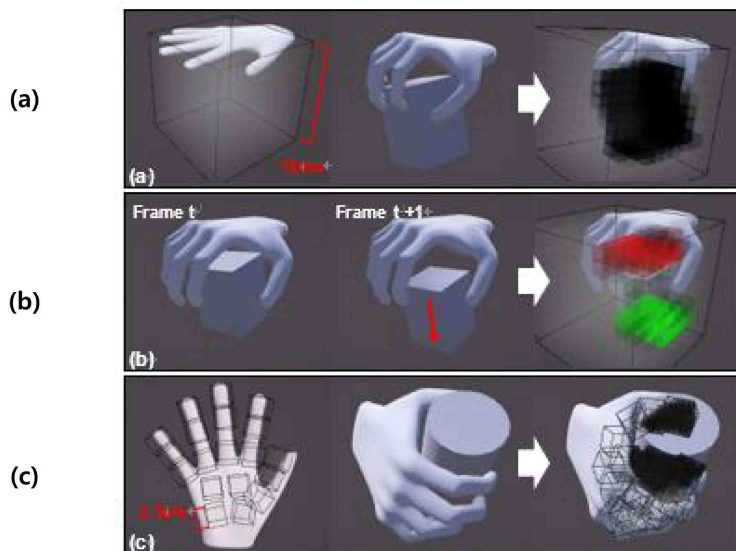
도면2



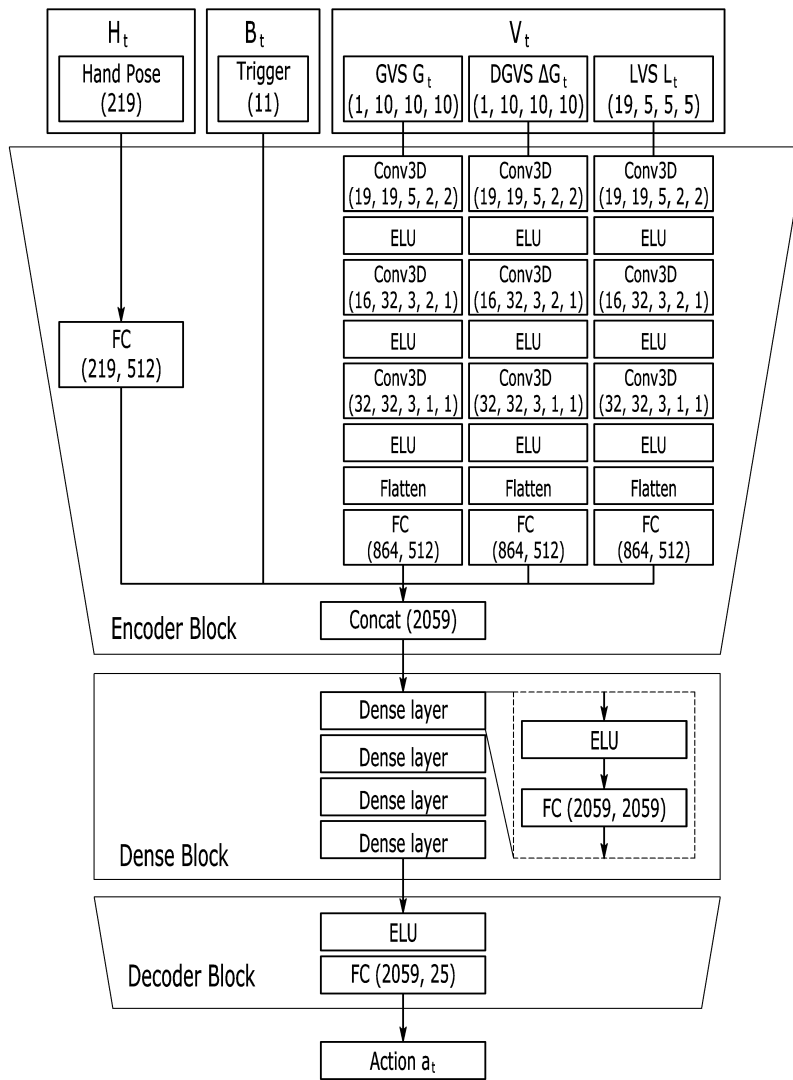
도면3



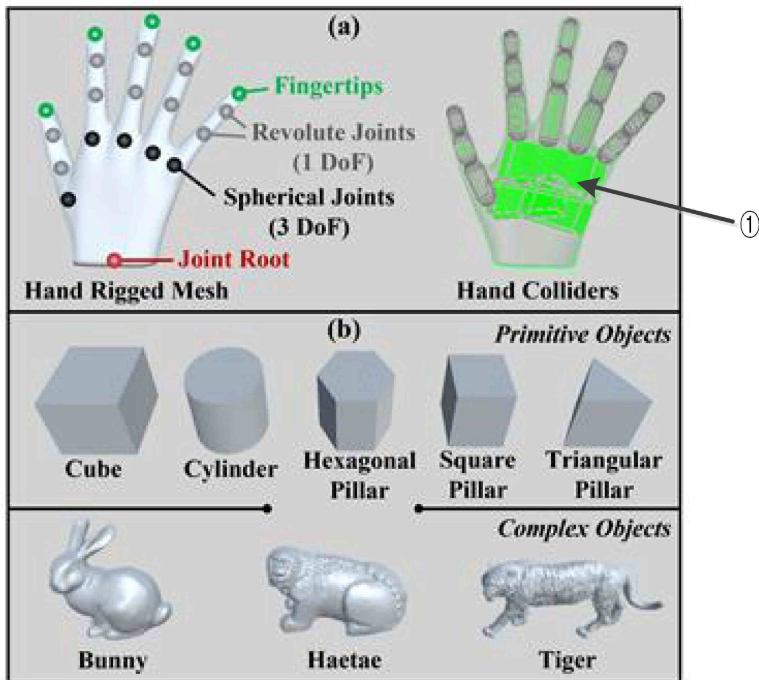
도면4



도면5



도면6



도면7

