



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0015413
(43) 공개일자 2025년02월03일

(51) 국제특허분류(Int. Cl.)
G06T 7/194 (2017.01) G06N 3/0455 (2023.01)
G06N 3/094 (2023.01) G06T 15/08 (2011.01)
G06T 17/00 (2006.01) G06T 19/20 (2011.01)
(52) CPC특허분류
G06T 7/194 (2017.01)
G06N 3/0455 (2023.01)
(21) 출원번호 10-2023-0096803
(22) 출원일자 2023년07월25일
심사청구일자 2023년07월25일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대 학교)
(72) 발명자
송병철
서울특별시 양천구 목동서로 38, 125동 307호(목동, 목동신시가지아파트1단지)
이상혁
인천광역시 미추홀구 한나루로477번길 112-5, 203호(용현동, 신영 빌)
(74) 대리인
양성보

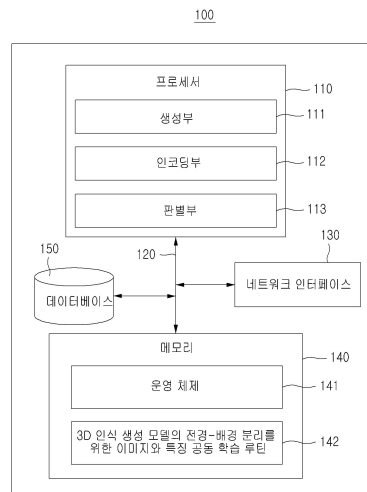
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 방법 및 시스템

(57) 요약

3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 방법 및 시스템이 제시된다. 본 발명에서 제안하는 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 시스템은 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 가상 이미지를 생성하는 생성부, 실제 이미지로부터 전경 벡터 및 배경 벡터를 각각 추출하기 위한 전경 인코더 및 배경 인코더를 포함하는 인코딩부 및 상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별하기 위한 전경 판별기 및 배경 판별기를 포함하는 판별부를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06N 3/094 (2023.01)
G06T 15/08 (2013.01)
G06T 17/00 (2013.01)
G06T 19/20 (2013.01)
G06T 2207/20081 (2013.01)
G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711179347
과제번호 00155915
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 정보통신인재양성(교육운영수익)
연구과제명 [Ezbaro][정부+민간] 인공지능융합혁신인재양성(2차년도)
기 여 율 33/100
과제수행기관명 인하대학교 산학협력단
연구기간 2023.01.01 ~ 2023.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호 1711193592
과제번호 2021-0-02052-003
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 대학ITRC센터육성지원사업(교육운영수익)
연구과제명 [Ezbaro][정부] 스마트 모빌리티를 위한 인공지능 시스템반도체핵심 기술 개발 및

인력 양성(3차년도)

기 여 율 33/100
과제수행기관명 인하대학교 산학협력단
연구기간 2023.01.01 ~ 2023.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호 1711179523
과제번호 2022R1A2C2010095
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 중견연구
연구과제명 [Ezbaro] 적대적 방어 기술 및 이를 적용한 통합 네트워크경량화 프레임워크에 관한

연구

기 여 율 34/100
과제수행기관명 인하대학교
연구기간 2023.03.01 ~ 2024.02.29

명세서

청구범위

청구항 1

전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 가상 이미지를 생성하는 생성부;

실제 이미지로부터 전경 벡터 및 배경 벡터를 각각 추출하기 위한 전경 인코더 및 배경 인코더를 포함하는 인코딩부; 및

상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별하기 위한 전경 판별기 및 배경 판별기를 포함하는 판별부

를 포함하는 이미지 및 특징 공동 학습 시스템.

청구항 2

제1항에 있어서,

상기 생성부는,

실제 이미지에 대한 잠재 벡터, 카메라 포즈, 변환 함수를 포함하는 특징을 이용하여 전경 특징 및 배경 특징을 분리하고, 상기 분리된 전경 특징 및 배경 특징을 이용하여 볼륨 렌더링을 통해 상기 전경 특징 및 배경 특징 각각에 대한 가상 이미지를 생성하는

이미지 및 특징 공동 학습 시스템.

청구항 3

제1항에 있어서,

상기 판별부는,

상기 전경 판별기 및 상기 배경 판별기가 서로 다른 이미지-특징 쌍을 각각 입력 받아 상기 서로 다른 이미지-특징 쌍을 이용하여 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하는

이미지 및 특징 공동 학습 시스템.

청구항 4

제3항에 있어서,

상기 전경 판별기는,

가상 이미지와 실제 이미지, 가상 이미지의 전경 특징과 실제 이미지의 전경 벡터의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하고,

상기 배경 판별기는,

가상 이미지와 실제 이미지, 가상 이미지의 배경 특징과 실제 이미지의 배경 벡터의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하는

이미지 및 특징 공동 학습 시스템.

청구항 5

생성부를 통해 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하여 가상 이미지를 생성하는 단계;

전경 인코더 및 배경 인코더를 포함하는 인코딩부를 통해 실제 이미지로부터 전경 벡터 및 배경 벡터를 각각 추

출하는 단계; 및

전경 판별기 및 배경 판별기를 포함하는 판별부를 통해 상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별하는 단계

를 포함하는 이미지 및 특징 공동 학습 방법.

청구항 6

제5항에 있어서,

상기 전경 인코더 및 배경 인코더를 포함하는 인코딩부를 통해 실제 이미지로부터 전경 벡터 및 배경 벡터를 각각 추출하는 단계는,

실제 이미지에 대한 잠재 벡터, 카메라 포즈, 변환 함수를 포함하는 특징을 이용하여 전경 특징 및 배경 특징을 분리하고, 상기 분리된 전경 특징 및 배경 특징을 이용하여 볼륨 렌더링을 통해 상기 전경 특징 및 배경 특징 각각에 대한 가상 이미지를 생성하는

이미지 및 특징 공동 학습 방법.

청구항 7

제5항에 있어서,

상기 전경 판별기 및 배경 판별기를 포함하는 판별부를 통해 상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별하는 단계는,

상기 전경 판별기 및 상기 배경 판별기가 서로 다른 이미지-특징 쌍을 각각 입력 받아 상기 서로 다른 이미지-특징 쌍을 이용하여 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하는

이미지 및 특징 공동 학습 방법.

청구항 8

제7항에 있어서,

상기 상기 전경 판별기 및 배경 판별기를 포함하는 판별부를 통해 상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별하는 단계는,

상기 전경 판별기를 통해 가상 이미지와 실제 이미지, 가상 이미지의 전경 특징과 실제 이미지의 전경 벡터의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하고,

상기 배경 판별기를 통해 가상 이미지와 실제 이미지, 가상 이미지의 배경 특징과 실제 이미지의 배경 벡터의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하는

이미지 및 특징 공동 학습 방법.

발명의 설명

기술 분야

[0001] 본 발명은 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 최근에 3차원 그래픽스의 기술을 활용해 영상을 생성하는 연구가 활발하다. NeRF(Neural radiance field)는 2D 학습 데이터셋 만을 사용하여 3D 정보의 특징을 학습하는 방식으로 지속적인 영상 합성(view synthesis)을 가능하게 했고, 이는 3D 인식 생성 모델 연구의 기반이 되었다. 최근에는 2D 생성 모델의 성능이 좋은 생성자 모듈을 활용하여 고해상도, 고퀄리티의 3D 영상을 생성하는 여러 연구들이 보고된 바 있다.

[0003] 하지만, 실험적 관찰에 따르면 종래 3D 인식 생성 모델들은 전경과 배경간 불안정한 분리 성능을 보인다. 즉,

기존의 3D 인식 생성 모델은 영상 속 전경의 모습이 변할 때, 배경의 모습도 변한다는 문제점을 가지고 있다. 이는 영상을 구성하는 특징이 여전히 얽혀있다는 뜻이고, 이는 생성 모델이 품질 높은 영상을 생성하는 것을 방해한다.

[0004] 종래의 대표적인 3D 인식 생성 모델인 GIRAFFE와 GIRAFFE-HD는 영상 내 각 객체(object)를 독립적인 3차원 NFF(Neural Feature Field)로 구성한 뒤 파라메트릭 볼륨 렌더링(Parametric Volume Rendering)을 거쳐 새로운 영상을 생성한다. 여기서, NFF는 한 물체의 밀도와 고차원 특징(high-dimensional feature)의 세트로 구성된 것이며, 볼륨 렌더링의 입력단계에 해당한다. GIRAFFE 모델들은 각 물체가 독립적으로 렌더링되기 때문에 기존 2차원 생성모델들보다 특징 간 분리 및 제어 가능성(controllability) 측면에서 우위에 있다.

[0005] 하지만 이 생성 모델들은 오직 이미지 영역에서만 학습되기 때문에 전경과 배경 간 분리 측면에서는 여전히 불안정하다. 결과적으로 해당 모델들이 생성한 영상의 품질이 저하되는 현상이 관찰된다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) [1] Michael Niemeyer and Andreas Geiger, "Giraffe: Representing scenes as compositional generative neural feature fields," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 11453-11464.

(특허문헌 0002) [2] Yang Xue, Yuheng Li, Krishna Kumar Singh, and Yong Jae Lee, "Giraffe hd: A high-resolution 3d-aware generative model," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 18440-18449.

발명의 내용

해결하려는 과제

[0007] 본 발명이 이루고자 하는 기술적 과제는 생성 모델을 이미지 레벨(image-level) 및 특징 레벨(feature-level)에서 공동 적대적 학습(co-adversarial learning)함으로써 종래기술의 GIRAFFE 대비 전경-배경 분리 성능을 개선하기 위해 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 방법 및 시스템을 제공하는데 있다.

과제의 해결 수단

[0008] 일 측면에 있어서, 본 발명에서 제안하는 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 시스템은 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 가상 이미지를 생성하는 생성부, 실제 이미지로부터 전경 벡터 및 배경 벡터를 각각 추출하기 위한 전경 인코더 및 배경 인코더를 포함하는 인코딩부 및 상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별하기 위한 전경 판별기 및 배경 판별기를 포함하는 판별부를 포함한다.

[0009] 상기 생성부는 실제 이미지에 대한 잠재 벡터, 카메라 포즈, 변환 함수를 포함하는 특징을 이용하여 전경 특징 및 배경 특징을 분리하고, 상기 분리된 전경 특징 및 배경 특징을 이용하여 볼륨 렌더링을 통해 상기 전경 특징 및 배경 특징 각각에 대한 가상 이미지를 생성한다.

[0010] 상기 판별부는 상기 전경 판별기 및 상기 배경 판별기가 서로 다른 이미지-특징 쌍을 각각 입력 받아 상기 서로 다른 이미지-특징 쌍을 이용하여 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.

[0011] 상기 전경 판별기는 가상 이미지와 실제 이미지, 가상 이미지의 전경 특징과 실제 이미지의 전경 벡터의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.

[0012] 상기 배경 판별기는 가상 이미지와 실제 이미지, 가상 이미지의 배경 특징과 실제 이미지의 배경 벡터의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.

[0013] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징

공동 학습 방법은 생성부를 통해 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하여 가상 이미지를 생성하는 단계, 전경 인코더 및 배경 인코더를 포함하는 인코딩부를 통해 실제 이미지로부터 전경 벡터 및 배경 벡터를 각각 추출하는 단계 및 전경 판별기 및 배경 판별기를 포함하는 판별부를 통해 상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별하는 단계를 포함한다.

발명의 효과

[0014] 본 발명의 실시예들에 따르면 생성 모델을 이미지 레벨(image-level) 및 특징 레벨(feature-level)에서 공동 적대적 학습(co-adversarial learning)함으로써 종래기술의 GIRAFFE 대비 전경-배경 분리 성능을 개선할 수 있다. 본 발명의 실시예에 따른 생성 모델은 우수한 전경-배경 분리 성능과 함께, 종래기술에 따른 3D 인식 생성 모델보다 작은 수의 파라미터만으로도 사실적인 영상을 생성할 수 있다.

도면의 간단한 설명

[0015] 도 1은 본 발명의 일 실시예에 따른 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지-특징 공동 학습 시스템의 구성을 나타내는 도면이다.

도 2는 본 발명의 일 실시예에 따른 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 방법을 설명하기 위한 흐름도이다.

도 3은 본 발명의 일 실시예에 따른 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 프레임워크를 설명하기 위한 도면이다.

도 4는 본 발명의 일 실시예에 따른 생성 모델과 종래기술 간의 정성적 성능 비교 결과를 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0016] 최근 3D 정보를 이용한 생성 모델에 대한 연구가 활발하다. 최신 3D 인식 생성 모델 중 하나인 GIRAFFE는 개별적으로 형성된 3D NFF(Neural Feature Field)를 볼륨 렌더링(volume rendering)하여 이미지를 생성하기 때문에 기존 생성 모델보다 나은 특징(feature) 간 분리(disentanglement)를 보여준다. 그러나 GIRAFFE는 여전히 전경과 배경의 분리가 안 된다는 문제를 겪고 있다.

[0017] 따라서, 본 발명에서는 생성 모델을 이미지 레벨(image-level) 및 특징 레벨(feature-level)에서 공동 적대적 학습(co-adversarial learning)함으로써 GIRAFFE 대비 전경-배경 분리 성능을 개선하고자 하였다. 풍부한 모의 실험 결과, 본 발명의 실시예에 따른 생성 모델은 우수한 전경-배경 분리 성능과 함께, 종래기술에 따른 3D 인식 생성 모델보다 작은 수의 파라미터만으로도 사실적인 영상을 생성할 수 있다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.

[0019] 도 1은 본 발명의 일 실시예에 따른 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지-특징 공동 학습 시스템의 구성을 나타내는 도면이다.

[0020] 본 발명은 3D 인식 생성 모델을 이미지 영역뿐만 아니라 특징 영역에서도 적대적 학습을 시키는 방법을 제안한다. 본 발명의 실시예에 따르면 정성적, 정량적 실험결과를 통해 이러한 학습 방식이 기존의 3D 인식 생성 모델의 전경-배경의 분리 성능 및 영상 생성 성능을 향상시켰다는 것을 확인했다.

[0021] 본 실시예에 따른 이미지-특징 공동 학습 시스템(100)은 프로세서(110), 버스(120), 네트워크 인터페이스(130), 메모리(140) 및 데이터베이스(150)를 포함할 수 있다. 메모리(140)는 운영체제(141) 및 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지-특징 공동 학습 루틴(142)을 포함할 수 있다. 프로세서(110)는 생성부(111), 인코딩부(112) 및 판별부(113)를 포함할 수 있다. 다른 실시예들에서 이미지-특징 공동 학습 시스템(100)은 도 1의 구성요소들보다 더 많은 구성요소들을 포함할 수도 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다. 예를 들어, 이미지-특징 공동 학습 시스템(100)은 디스플레이나 트랜시버(transceiver)와 같은 다른 구성요소들을 포함할 수도 있다.

[0022] 메모리(140)는 컴퓨터에서 판독 가능한 기록 매체로서, RAM(random access memory), ROM(read only memory) 및 디스크 드라이브와 같은 비소멸성 대용량 기록장치(permanent mass storage device)를 포함할 수 있다. 또한,

메모리(140)에는 운영체제(141)와 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지-특징 공동 학습 루틴(142)을 위한 프로그램 코드가 저장될 수 있다. 이러한 소프트웨어 구성요소들은 드라이브 메커니즘(drive mechanism, 미도시)을 이용하여 메모리(140)와는 별도의 컴퓨터에서 판독 가능한 기록 매체로부터 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록 매체는 플로피 드라이브, 디스크, 테이프, DVD/CD-ROM 드라이브, 메모리 카드 등의 컴퓨터에서 판독 가능한 기록 매체(미도시)를 포함할 수 있다. 다른 실시예에서 소프트웨어 구성요소들은 컴퓨터에서 판독 가능한 기록 매체가 아닌 네트워크 인터페이스(130)를 통해 메모리(140)에 로딩될 수도 있다.

[0023] 버스(120)는 이미지-특징 공동 학습 시스템(100)의 구성요소들간의 통신 및 데이터 전송을 가능하게 할 수 있다. 버스(120)는 고속 시리얼 버스(high-speed serial bus), 병렬 버스(parallel bus), SAN(Storage Area Network) 및/또는 다른 적절한 통신 기술을 이용하여 구성될 수 있다.

[0024] 네트워크 인터페이스(130)는 이미지-특징 공동 학습 시스템(100)을 컴퓨터 네트워크에 연결하기 위한 컴퓨터 하드웨어 구성요소일 수 있다. 네트워크 인터페이스(130)는 이미지-특징 공동 학습 시스템(100)을 무선 또는 유선 커넥션을 통해 컴퓨터 네트워크에 연결시킬 수 있다.

[0025] 데이터베이스(150)는 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지-특징 공동 학습을 위해 필요한 모든 정보를 저장 및 유지하는 역할을 할 수 있다. 도 1에서는 이미지-특징 공동 학습 시스템(100)의 내부에 데이터베이스(150)를 구축하여 포함하는 것으로 도시하고 있으나, 이에 한정되는 것은 아니며 시스템 구현 방식이나 환경 등에 따라 생략될 수 있고 혹은 전체 또는 일부의 데이터베이스가 별개의 다른 시스템 상에 구축된 외부 데이터베이스로서 존재하는 것 또한 가능하다.

[0026] 프로세서(110)는 기본적인 산술, 로직 및 이미지-특징 공동 학습 시스템(100)의 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(140) 또는 네트워크 인터페이스(130)에 의해, 그리고 버스(120)를 통해 프로세서(110)로 제공될 수 있다. 프로세서(110)는 생성부(111), 인코딩부(112) 및 판별부(113)를 위한 프로그램 코드를 실행하도록 구성될 수 있다. 이러한 프로그램 코드는 메모리(140)와 같은 기록 장치에 저장될 수 있다.

[0027] 생성부(111), 인코딩부(112) 및 판별부(113)는 도 2의 단계들(210~230)을 수행하기 위해 구성될 수 있다.

[0028] 이미지-특징 공동 학습 시스템(100)은 생성부(111), 인코딩부(112) 및 판별부(113)를 포함할 수 있다.

[0029] 본 발명의 실시예에 따른 생성부(111)는 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 가상 이미지를 생성한다.

[0030] 본 발명의 실시예에 따른 생성부(111)는 실제 이미지에 대한 잠재 벡터(Latent vector), 카메라 포즈(Camera pose), 변환 함수(Transformation)를 포함하는 특징을 이용하여 전경 특징 및 배경 특징을 분리하고, 상기 분리된 전경 특징 및 배경 특징을 이용하여 볼륨 렌더링을 통해 상기 전경 특징 및 배경 특징 각각에 대한 가상 이미지를 생성한다.

[0031] 종래기술에 따르면 생성 모델을 이미지와 특징 쌍으로 적대적 학습을 수행하는 경우 실제 이미지와 가상 이미지 간의 분포(distribution)뿐만 아니라 이들의 특징 분포(feature distribution)까지 매칭시킨다. 결과적으로 생성 모델은 실제 이미지가 가진 잠재 특징(latent feature)의 구조(structure)와 패턴(pattern)까지 파악함으로써 사실적(photo-realistic)이면서도 특징 간 분리(disentanglement)가 잘된 영상을 생성할 수 있다.

[0032] 본 발명의 실시예에 따른 인코딩부(112)는 실제 이미지로부터 전경 벡터 및 배경 벡터를 각각 추출하기 위한 전경 인코더 및 배경 인코더를 포함한다.

[0033] 본 발명의 실시예에 따른 인코딩부(112)는 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 구현하기 위해 전경 인코더 및 배경 인코더를 적용하여, 각각 전경 벡터와 배경 벡터를 추출한다.

[0034] 본 발명의 실시예에 따른 판별부(113)는 상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별하기 위한 전경 판별기 및 배경 판별기를 포함한다.

[0035] 본 발명의 실시예에 따른 판별부(113)는 전경 판별기 및 배경 판별기가 서로 다른 이미지-특징 쌍을 각각 입력받아 상기 서로 다른 이미지-특징 쌍을 이용하여 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.

- [0036] 더욱 상세하게는, 본 발명의 실시예에 따른 전경 판별기는 가상 이미지와 실제 이미지, 가상 이미지의 전경 특징과 실제 이미지의 전경 벡터의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.
- [0037] 본 발명의 실시예에 따른 배경 판별기는 가상 이미지와 실제 이미지, 가상 이미지의 배경 특징과 실제 이미지의 배경 벡터의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.
- [0039] 도 2는 본 발명의 일 실시예에 따른 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 방법을 설명하기 위한 흐름도이다.
- [0040] 본 발명의 실시예에 따른 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 방법은 생성부를 통해 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하여 가상 이미지를 생성하는 단계(210), 전경 인코더 및 배경 인코더를 포함하는 인코딩부를 통해 실제 이미지로부터 전경 벡터 및 배경 벡터를 각각 추출하는 단계(220) 및 전경 판별기 및 배경 판별기를 포함하는 판별부를 통해 상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별하는 단계(230)를 포함한다.
- [0041] 단계(210)에서, 생성부를 통해 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하여 가상 이미지를 생성한다.
- [0042] 본 발명의 실시예에 따르면, 실제 이미지에 대한 잠재 벡터(Latent vector), 카메라 포즈(Camera pose), 변환 함수(Transformation)를 포함하는 특징을 이용하여 전경 특징 및 배경 특징을 분리하고, 상기 분리된 전경 특징 및 배경 특징을 이용하여 볼륨 렌더링을 통해 상기 전경 특징 및 배경 특징 각각에 대한 가상 이미지를 생성한다.
- [0043] 종래기술에 따르면 생성 모델을 이미지와 특징 쌍으로 적대적 학습을 수행하는 경우 실제 이미지와 가상 이미지 간의 분포(distribution)뿐만 아니라 이들의 특징 분포(feature distribution)까지 매칭시킨다. 결과적으로 생성 모델은 실제 이미지가 가진 잠재 특징(latent feature)의 구조(structure)와 패턴(pattern)까지 파악함으로써 사실적(photo-realistic)이면서도 특징 간 분리(disentanglement)가 잘된 영상을 생성할 수 있다.
- [0044] 단계(220)에서, 전경 인코더 및 배경 인코더를 포함하는 인코딩부를 통해 실제 이미지로부터 전경 벡터 및 배경 벡터를 각각 추출한다.
- [0045] 본 발명의 실시예에 따르면, 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 구현하기 위해 전경 인코더 및 배경 인코더를 적용하여, 각각 전경 벡터와 배경 벡터를 추출한다.
- [0046] 단계(230)에서, 전경 판별기 및 배경 판별기를 포함하는 판별부를 통해 상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별한다.
- [0047] 본 발명의 실시예에 따르면, 전경 판별기 및 배경 판별기가 서로 다른 이미지-특징 쌍을 각각 입력 받아 상기 서로 다른 이미지-특징 쌍을 이용하여 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.
- [0048] 더욱 상세하게는, 본 발명의 실시예에 따른 전경 판별기는 가상 이미지와 실제 이미지, 가상 이미지의 전경 특징과 실제 이미지의 전경 벡터의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.
- [0049] 본 발명의 실시예에 따른 배경 판별기는 가상 이미지와 실제 이미지, 가상 이미지의 배경 특징과 실제 이미지의 배경 벡터의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.
- [0051] 도 3은 본 발명의 일 실시예에 따른 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습 프레임워크를 설명하기 위한 도면이다.
- [0052] 본 발명의 일 실시예에 따른 3D 인식 생성 모델의 전경-배경 분리를 위한 이미지와 특징 공동 학습은, 먼저 생성부(Generator)(310)를 통해 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행하여 가상 이미지를 생성한다.

- [0053] 본 발명의 실시예에 따른 생성부(310)는 실제 이미지에 대한 잠재 벡터(Latent vector), 카메라 포즈(Camera pose), 변환 함수(Transformation)를 포함하는 특징을 이용하여 전경 특징(f^{fg}) 및 배경 특징(f^{bg})을 분리하고, 상기 분리된 전경 특징 및 배경 특징을 이용하여 볼륨 렌더링(volume rendering)을 통해 상기 전경 특징 및 배경 특징 각각에 대한 가상 이미지(I^F)를 생성한다.
- [0054] 본 발명의 실시예에 따른 인코딩부(Encoder) (320)는 실제 이미지(Real image)(I)로부터 전경 벡터(v^{fg}) 및 배경 벡터(v^{bg})를 각각 추출하기 위한 전경 인코더(321) 및 배경 인코더(322)를 포함한다.
- [0055] 본 발명의 실시예에 따른 인코딩부(320)는 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 구현하기 위해 전경 인코더(321) 및 배경 인코더(322)를 적용하여, 각각 전경 벡터(v^{fg})와 배경 벡터(v^{bg})를 추출한다.
- [0056] 본 발명의 실시예에 따른 판별부(Discriminator)(330)는 상기 전경 특징 및 배경 특징을 분리하고, 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 통해 상기 가상 이미지와 상기 실제 이미지를 판별하기 위한 전경 판별기(Foreground Discriminator)(331) 및 배경 판별기(Background Discriminator)(332)를 포함한다.
- [0057] 본 발명의 실시예에 따른 판별부는 전경 판별기(331) 및 배경 판별기(332)가 서로 다른 이미지-특징 쌍을 각각 입력 받아 상기 서로 다른 이미지-특징 쌍을 이용하여 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.
- [0058] 더욱 상세하게는, 본 발명의 실시예에 따른 전경 판별기(331)는 가상 이미지(I^F)와 실제 이미지(I), 가상 이미지의 전경 특징(f^{fg})과 실제 이미지의 전경 벡터(v^{fg})의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.
- [0059] 본 발명의 실시예에 따른 배경 판별기(332)는 가상 이미지(I^F)와 실제 이미지(I), 가상 이미지의 배경 특징(f^{bg})과 실제 이미지의 배경 벡터(v^{bg})의 이미지-특징 쌍을 입력 받아 이미지 레벨 및 특징 레벨에서 공동 적대적 학습을 수행한다.
- [0060] 이와 같이 본 발명의 실시예에 따르면, 또한, 전경과 배경을 위한 두 개의 판별기를 도입하여 전경 및 배경 간 특징의 분리(disentanglement) 효과를 높이고자 하였다. 두 판별기는 서로 다른 이미지-특징 쌍을 입력 받아 독립적으로 적대적 학습을 수행한다. 도 3과 같이 두 개의 인코더들과 두 개의 판별기들로 구성된 제안하는 모듈이 GIRAFFE에 적용되어 이미지-특징 레벨에서 적대적 학습을 수행하도록 한다.
- [0062] 도 4는 본 발명의 일 실시예에 따른 생성 모델과 종래기술 간의 정성적 성능 비교 결과를 위한 도면이다.
- [0063] 도 4(a)는 종래기술에 따른 GIRAFFE, 도 4(b)는 종래기술에 따른 GIRAFFE-HD, 도 4(c)는 본 발명의 일 실시예에 따른 이미지-특징 공동 학습을 통한 생성 모델의 결과를 나타낸다.
- [0064] 도 4를 참조하여, 영상의 배경(background)이 전경(foreground)의 변화에 따라 어떻게 변하는지 확인한다. 본 발명의 일 실시예에 따른 이미지-특징 공동 학습을 통한 생성 모델은 전경의 특징(예를 들어, shape, appearance, and pose)이 변해도 배경의 변화가 종래기술들 모델보다 작다.
- [0065] 표1은 3D인식 생성 모델의 성능을 정량적으로 보여준다.

[0066] <표1>

	GRAF	GIRAFFE	GIRAFFE-HD	Ours
FFHQ	59	35	34	29
CompCars	95	28	23	26

- [0067]
- [0068] FFHQ에서 본 발명의 일 실시예에 따른 이미지-특징 공동 학습을 통한 생성 모델은 GIRAFFE-HD보다 5만큼 낮아진 FID 성능을 보여준다. 이는 GIRAFFE와 GIRAFFE-HD의 수치적 차이보다 무려 5배 크며 우리의 공동 적대적 학습 방식의 유효성을 보여준다.

[0069] 표2는 GIRAFFE-HD와 본 발명의 일 실시예에 따른 이미지-특징 공동 학습을 통한 생성 모델의 생성기 파라미터 수를 보여준다.

[0070] <표 2>

Method	GIRAFFE-HD	Ours
Number of Parameters (M)	12.66	0.74

[0071]

[0072] GIRAFFE-HD는 2개의 무거운 StyleGAN2 생성기를 바이어스 생성기(basis generator)로 채택하였기 때문에 GIRAFFE의 것과 유사한 메커니즘으로 설계된 본 발명의 일 실시예에 따른 생성기는 이들보다 약 17배 적은 파라미터 수를 보인다.

[0074] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0075] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0076] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

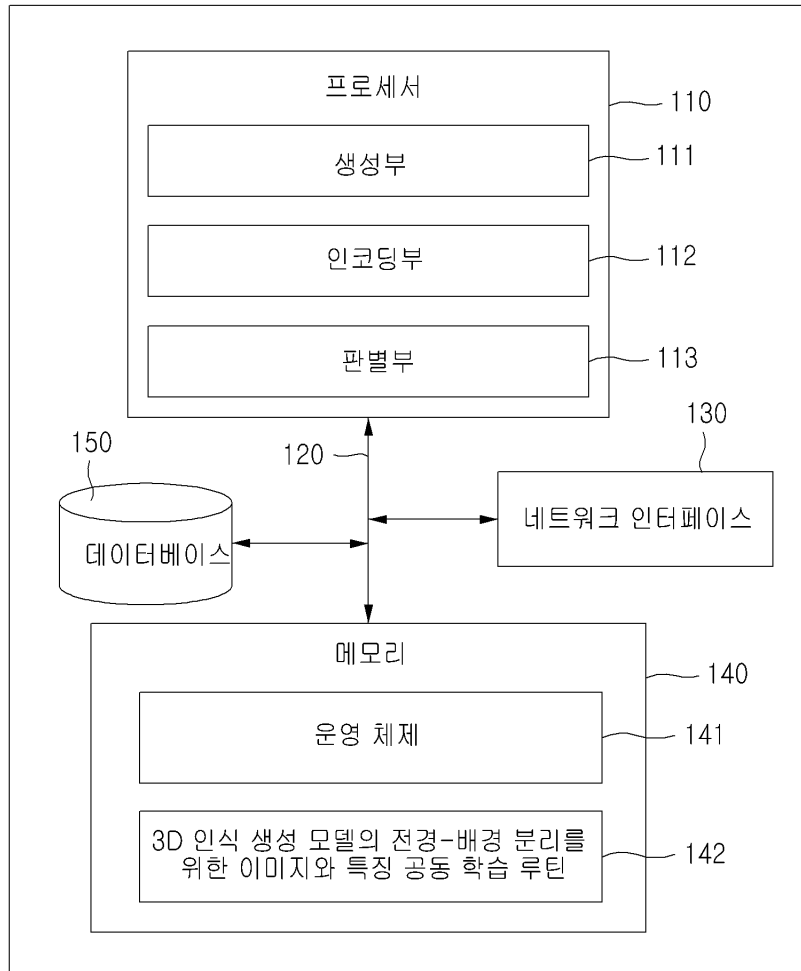
[0077] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0078] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

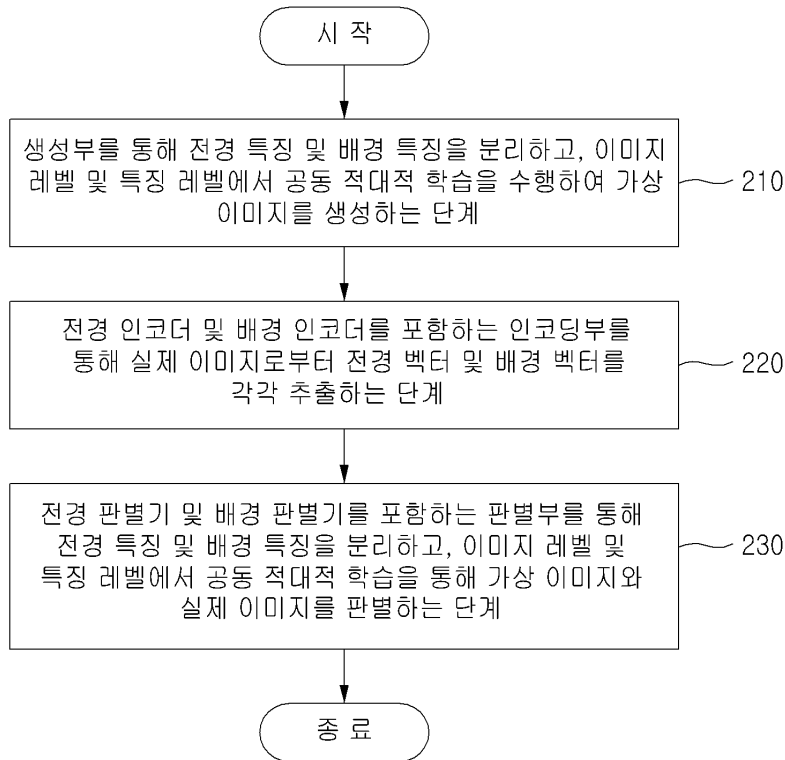
도면

도면1

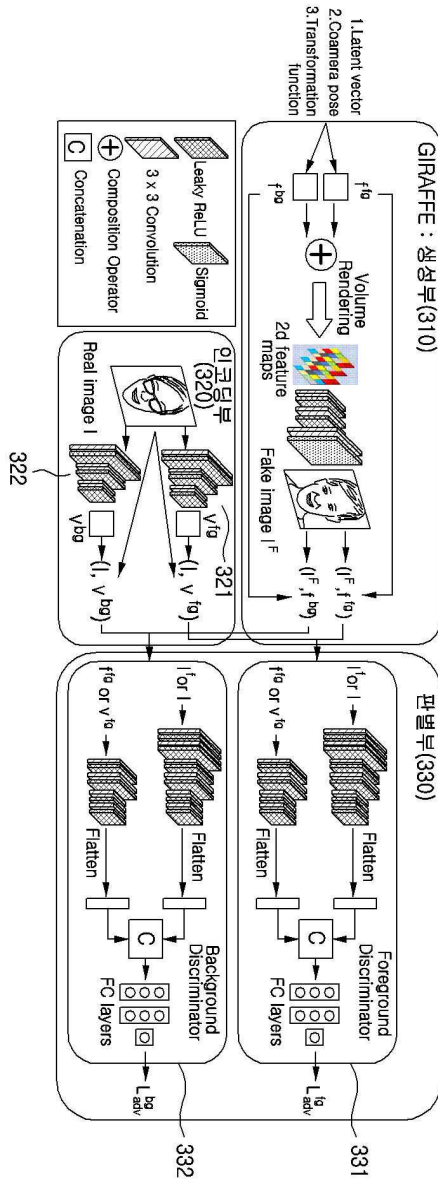
100



도면2



도면3



도면4

