(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년03월19일
(11) 등록번호 10-2648938
(24) 등록일자 2024년03월14일

(51) 국제특허분류(Int. Cl.)
G06T 17/10 (2006.01) *G06T 15/06* (2011.01)
G06T 15/08 (2011.01) *G06T 15/10* (2006.01)
G06T 3/40 (2024.01) *G06T 7/55* (2017.01)
G06V 10/40 (2022.01)
(52) CPC특허분류
G06T 17/10 (2013.01)
G06T 15/06 (2013.01)
(21) 출원번호 10-2023-0048012
(22) 출원일자 2023년04월12일
심사청구일자 2023년04월12일
(30) 우선권주장
1020230020095 2023년02월15일 대한민국(KR)
(56) 선행기술조사문헌
KR1020220140402 A*
(뒷면에 계속)
전체 청구항 수 : 총 8 항 심사관 : 이병우

(73) 특허권자
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
김승룡
서울특별시 성북구 오패산로 46, 103동 1804호(하월곡동, 월곡두산위브아파트)
곽민섭
서울특별시 서초구 서초중앙로24길 57(서초동, 서초롯데캐슬프레지던트)
(74) 대리인
송인호, 윤형근, 최관락

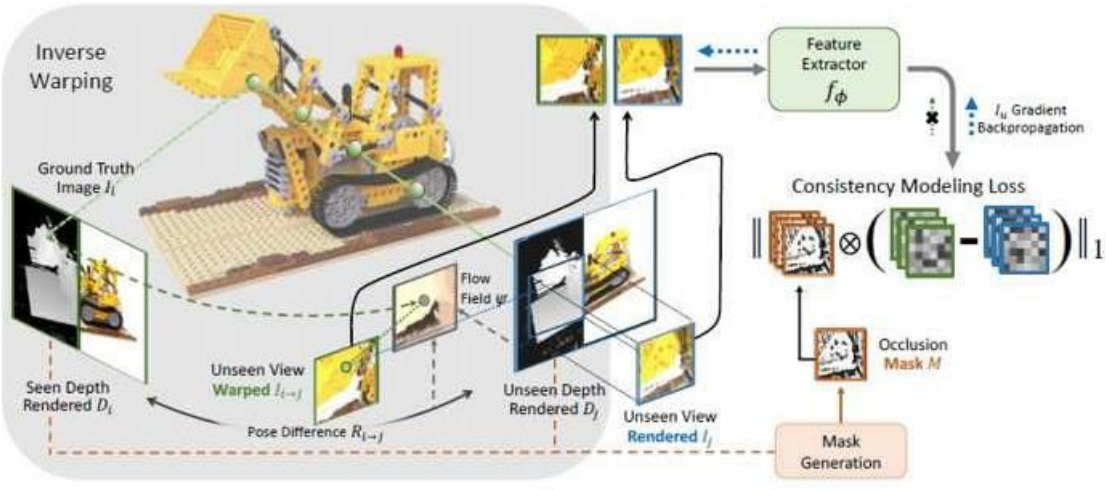
(54) 발명의 명칭 **기하학적 일관성을 이용한 소수 샷 신경 방사장 기반 3차원 이미지 재구성 방법 및 장치**

(57) 요 약

본 발명은 기하학적 일관성을 이용한 소수 샷 신경 방사장 기반 3차원 이미지 재구성 방법 및 장치를 개시한다. 본 발명에 따르면, 프로세서; 및 상기 프로세서에 연결되는 메모리를 포함하되, 상기 메모리는, 소정 장면에 대한 소수 샷에 해당하는 N개 시점에서의 실측 이미지를 입력 받고, 신경 방사장 모델 기반으로 하나 이상의 새로

(뒷면에 계속)

대 표 도



운 시점(unseen viewpoint)에서의 제1 이미지 및 제1 깊이 맵을 획득하고, 상기 제1 깊이 맵을 이용하여 상기 실측 이미지를 상기 새로운 시점으로 워프(warp)하여 제1 워프 이미지를 획득하고, 복수의 레이어를 포함하는 컨볼루션 네트워크를 이용하여 상기 제1 이미지 및 상기 제1 워프 이미지 각각의 특징 맵을 추출하고, 상기 각각의 특징 맵에 서로 다른 시점에서의 기하학적 불일치 영역을 무시하기 위한 일관성 마스크를 적용하고, 상기 일관성 마스크가 적용된 상기 제1 이미지 및 제1 워프 이미지의 특징 맵 사이의 일관성 손실을 이용하여 상기 신경 방사장 모델이 상기 제1 이미지가 상기 제1 워프 이미지를 따르게 학습하여 상기 소정 장면에 상응하는 3차원 이미지를 재구성하도록, 상기 프로세서에 의해 실행되는 프로그램 명령어들을 저장한 3차원 이미지 재구성 장치가 제공된다.

(52) CPC특허분류

G06T 15/08 (2013.01)
G06T 15/10 (2013.01)
G06T 3/4076 (2024.01)
G06T 7/55 (2017.01)
G06V 10/40 (2023.08)
G06T 2207/20084 (2013.01)

(56) 선행기술조사문헌

WO2022182421 A1

Alex Yu et al., pixelNeRF: Neural Radiance Fields from One or Few Images, arXiv:2012.02190v3 [cs.CV] 30, May, 2021*

Clement et al., Unsupervised Monocular Depth Estimation with Left-Right Consistency, arXiv:1609.03677v3 [cs.CV] 12, Apr, 2017*

Mohammad et al., GeoNeRF: Generalizing NeRF with Geometry Priors, arXiv:2111.13539v2 [cs.CV] 21, Mar, 2022*

*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호	1711159611
과제번호	2020-0-01819-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	ICT명품인재양성(고려대학교)
기 여 율	1/1
과제수행기관명	고려대학교산학협력단
연구기간	2022.01.01 ~ 2022.12.31

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

소수 샷이 제공되는 환경에서 기하학적 일관성을 이용한 신경 방사장 기반 3차원 이미지 재구성 장치로서,

프로세서; 및

상기 프로세서에 연결되는 메모리를 포함하되,

상기 메모리는,

소정 장면에 대한 소수 샷에 해당하는 N개 시점에서의 실측 이미지를 입력 받고,

신경 방사장 모델 기반으로 하나 이상의 새로운 시점(unseen viewpoint)에서의 제1 이미지 및 제1 깊이 맵을 획득하고,

상기 제1 깊이 맵을 이용하여 상기 실측 이미지를 상기 새로운 시점으로 워프(warp)하여 제1 워프 이미지를 획득하고,

복수의 레이어를 포함하는 컨볼루션 네트워크를 이용하여 상기 제1 이미지 및 상기 제1 워프 이미지 각각의 특징 맵을 추출하고,

상기 각각의 특징 맵에 서로 다른 시점에서의 기하학적 불일치 영역을 무시하기 위한 일관성 마스크를 적용하고,

상기 일관성 마스크가 적용된 상기 제1 이미지 및 제1 워프 이미지의 특징 맵 사이의 일관성 손실을 이용하여 상기 신경 방사장 모델이 상기 제1 이미지가 상기 제1 워프 이미지를 따르게 학습하여 상기 소정 장면에 상응하는 3차원 이미지를 재구성하도록,

상기 프로세서에 의해 실행되는 프로그램 명령어들을 저장하되,

상기 프로그램 명령어들은,

상기 실측 이미지, 상기 제1 이미지 및 상기 워프 이미지의 패치를 이용하여 상기 일관성 마스크를 생성하고,

상기 일관성 마스크를 상기 컨볼루션 네트워크의 복수의 레이어 각각이 출력하는 특징 맵에 적용하고,

픽셀 단위로 소정 지점에 대한 상기 실측 이미지에서의 깊이 및 상기 제1 이미지에서의 깊이 사이의 유클리드 거리가 미리 설정된 임계치를 초과하는 경우 해당 지점을 상기 일관성 마스크에 포함되는 지점으로 결정하고,

상기 실측 이미지에 이미지 그래디언트를 사용하는 엣지 인식 불일치 정규화를 도입하며,

상기 제1 워프 이미지를 의사 실측 데이터(pseudo ground truth)로 하여 그래디언트 역전파가 상기 제1 이미지를 통해서만 통과하고 상기 제1 워프 이미지에 대해서는 차단되도록 하는 3차원 이미지 재구성 장치.

청구항 2

제1항에 있어서,

상기 N은 3 또는 5인 3차원 이미지 재구성 장치.

청구항 3

제1항에 있어서,

상기 프로그램 명령어들은,

랜덤하게 카메라 시점을 샘플링하고, 상기 카메라 시점에서의 광선 벡터를 패치별로 생성하여 상기 제1 이미지를 획득하는 3차원 이미지 재구성 장치.

청구항 4

제1항에 있어서,

상기 프로그램 명령어들은,

상기 신경 방사장 모델이 상기 새로운 시점에서 출력하는 밀도 값을 이용하여 저해상도 깊이 맵을 획득하는 3차원 이미지 재구성 장치.

청구항 5

제4항에 있어서,

상기 프로그램 명령어들은,

상기 신경 방사장 모델을 통한 체적 렌더링 후, 쌍선형 보간으로 상기 저해상도 깊이 맵을 고해상도의 제1 깊이 맵으로 업샘플링하는 3차원 이미지 재구성 장치.

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

프로세서 및 메모리를 포함하는 장치에서 소수 샷이 제공되는 환경에서 기하학적 일관성을 이용한 신경 방사장 기반으로 3차원 이미지를 재구성하는 방법으로서,

소정 장면에 대한 소수 샷에 해당하는 N개 시점에서의 실측 이미지를 입력 받는 단계;

신경 방사장 모델 기반으로 하나 이상의 새로운 시점(unseen viewpoint)에서의 제1 이미지 및 제1 깊이 맵을 획득하는 단계;

상기 제1 깊이 맵을 이용하여 상기 실측 이미지를 상기 새로운 시점으로 워프(warp)하여 제1 워프 이미지를 획득하는 단계;

복수의 레이어를 포함하는 컨볼루션 네트워크를 이용하여 상기 제1 이미지 및 상기 제1 워프 이미지 각각의 특징 맵을 추출하는 단계;

상기 각각의 특징 맵에 서로 다른 시점에서의 기하학적 불일치 영역을 무시하기 위한 일관성 마스크를 적용하는 단계; 및

상기 일관성 마스크가 적용된 상기 제1 이미지 및 제1 워프 이미지의 특징 맵 사이의 일관성 손실을 이용하여 상기 신경 방사장 모델이 상기 제1 이미지가 상기 제1 워프 이미지를 따르게 학습하여 상기 소정 장면에 상응하는 3차원 이미지를 재구성하는 단계를 포함하되,

상기 일관성 마스크를 적용하는 단계는,

상기 실측 이미지, 상기 제1 이미지 및 상기 워프 이미지의 패치를 이용하여 상기 일관성 마스크를 생성하는 단계; 및

상기 일관성 마스크를 상기 컨볼루션 네트워크의 복수의 레이어 각각이 출력하는 특징 맵에 적용하는 단계를 포

함하되,

상기 일관성 마스크를 생성하는 단계는,

픽셀 단위로 소정 지점에 대한 상기 실측 이미지에서의 깊이 및 상기 제1 이미지에서의 깊이 사이의 유클리드 거리가 미리 설정된 임계치를 초과하는 경우 해당 지점을 상기 일관성 마스크에 포함되는 지점으로 결정하는 단계를 포함하고,

상기 3차원 이미지를 재구성하는 단계는, 상기 실측 이미지에 이미지 그래디언트를 사용하는 엣지 인식 불일치 정규화를 도입하는 단계; 및

상기 제1 워프 이미지를 의사 실측 데이터(pseudo ground truth)로 하여 그래디언트 역전파가 상기 제1 이미지를 통해서만 통과하고 상기 제1 워프 이미지에 대해서는 차단되도록 하는 단계를 포함하는 3차원 이미지 재구성 방법.

청구항 11

제10항에 있어서,

상기 제1 깊이 맵을 획득하는 단계는,

상기 신경 방사장 모델이 상기 새로운 시점에서 출력하는 밀도 값을 이용하여 저해상도 깊이 맵을 획득하는 단계; 및

상기 신경 방사장 모델을 통한 체적 렌더링 후, 쌍선형 보간으로 상기 저해상도 깊이 맵을 고해상도의 제1 깊이 맵으로 업샘플링하는 단계를 포함하는 3차원 이미지 재구성 방법.

청구항 12

제10항에 따른 방법을 수행하는 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

발명의 설명

기 술 분 야

[0001] 본 발명은 기하학적 일관성을 이용한 소수 샷 신경 방사장 기반 3차원 이미지 재구성 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 3차원 물체를 신경 방사장(Neural Radiance Fields, NeRF)으로 표현하는 것은 최근 새로운 시점 합성(novel view synthesis) 및 3차원 재구성을 위한 강력한 접근법으로 입증되었다.

[0003] NeRF는 3차원 이미지 재구성 작업을 위해 간단한 다층 퍼셉트론(Multi-Layer Perceptron, MLP) 아키텍처에 체적 렌더링(volumetric rendering)을 사용해 신경 방사장에 3차원 공간을 과적합(overfitting) 시킨 후 이를 새로운 시점에서 물체의 이미지를 형성하는 새로운 시점 합성에 활용한다.

[0004] 이와 같은 neural implicit representation의 탁월한 성능을 기반으로 한 다양한 후속 연구가 촉진되었는데, 여기에는 Nerfies와 같은 동적이고 변형 가능한 장면을 모델링 한 네트워크와, Plenoctree와 KiloNeRF와 같은 실시간 고속 렌더링 모델이 존재한다. 생성 모델링 또한 GRAF와 GIRAFFE와 같은 생성 네트워크에서 NeRF로 모델링 되었다.

[0005] NeRF는 인상적인 성능에도 불구하고 최적화를 위해 다수의 밀도가 높고 잘 분산된 보정된 이미지를 필요로 하므로 적용 가능성이 제한된다. 반면 이미지의 개수가 제한된 소수 샷(few-shot) 환경으로 제한될 때, NeRF는 입력 이미지에 쉽게 과적합되고 3차원 재구성을 안정적으로 수행할 수 없다.

[0006] Few-shot novel view synthesis이라고 하는 이 문제를 직접 해결하는 작업은 이러한 희소 관측 시나리오에서 충실도가 높은 NeRF를 최적화하는 것을 목표로 한다.

[0007] 이전 연구들은 의미론적 특징, 엔트로피 최소화 또는 normalizing flow를 활용하여 이 작업을 해결하는 것을 시

도했지만, 인위적인 fine-tuning 방법의 필요성과 물체의 미세 구조를 재구성할 수 없는 것이 지금까지 적용 가능성을 제한했다.

선행기술문헌

비특허문헌

[0008] (비특허문헌 0001) Alex Yu Vickie Ye Matthew Tancik Angjoo Kanazawa, pixelNeRF: Neural Radiance Fields from One or Few Images, https://arxiv.org/pdf/2012.02190.pdf,arXiv:2012.02190v3 [cs.CV] 30, May, 2021

발명의 내용

해결하려는 과제

[0009] 상기한 종래기술의 문제점을 해결하기 위해, 본 발명은 희소 관측 설정에도 불구하고 이전 모델에 비해 부드러운 표면을 렌더링하고 배경에서 artifacts를 줄일 수 있는 기하학적 일관성을 이용한 소수 샷 신경 방사장 기반 3차원 이미지 재구성 방법 및 장치를 제안하고자 한다.

과제의 해결 수단

[0010] 상기한 바와 같은 목적을 달성하기 위하여, 본 발명의 일 실시예에 따르면, 소수 샷이 제공되는 환경에서 기하학적 일관성을 이용한 신경 방사장 기반 3차원 이미지 재구성 장치로서, 프로세서; 및 상기 프로세서에 연결되는 메모리를 포함하되, 상기 메모리는, 소정 장면에 대한 소수 샷에 해당하는 N개 시점에서의 실측 이미지를 입력 받고, 신경 방사장 모델 기반으로 하나 이상의 새로운 시점(unseen viewpoint)에서의 제1 이미지 및 제1 깊이 맵을 획득하고, 상기 제1 깊이 맵을 이용하여 상기 실측 이미지를 상기 새로운 시점으로 워프(warp)하여 제1 워프 이미지를 획득하고, 복수의 레이어를 포함하는 컨볼루션 네트워크를 이용하여 상기 제1 이미지 및 상기 제1 워프 이미지 각각의 특징 맵을 추출하고, 상기 각각의 특징 맵에 서로 다른 시점에서의 기하학적 불일치 영역을 무시하기 위한 일관성 마스크를 적용하고, 상기 일관성 마스크가 적용된 상기 제1 이미지 및 제1 워프 이미지의 특징 맵 사이의 일관성 손실을 이용하여 상기 신경 방사장 모델이 상기 제1 이미지가 상기 제1 워프 이미지를 따르게 학습하여 상기 소정 장면에 상응하는 3차원 이미지를 재구성하도록, 상기 프로세서에 의해 실행되는 프로그램 명령어들을 저장한 3차원 이미지 재구성 장치가 제공된다.

[0011] 상기 N은 3 또는 5일 수 있다.

[0012] 상기 프로그램 명령어들은, 랜덤하게 카메라 시점을 샘플링하고, 상기 카메라 시점에서의 광선 벡터를 패치별로 생성하여 상기 제1 이미지를 획득할 수 있다.

[0013] 상기 프로그램 명령어들은, 상기 신경 방사장 모델이 상기 새로운 시점에서 출력하는 밀도 값을 이용하여 저해상도 깊이 맵을 획득할 수 있다.

[0014] 상기 프로그램 명령어들은, 상기 신경 방사장 모델을 통한 체적 렌더링 후, 쌍선형 보간으로 상기 저해상도 깊이 맵을 고해상도의 제1 깊이 맵으로 업샘플링할 수 있다.

[0015] 상기 프로그램 명령어들은, 상기 실측 이미지, 상기 제1 이미지 및 상기 워프 이미지의 패치를 이용하여 마스크를 생성하고, 상기 생성된 마스크를 상기 컨볼루션 네트워크의 복수의 레이어 각각이 출력하는 특징 맵에 적용할 수 있다.

[0016] 상기 프로그램 명령어들은, 픽셀 단위로 소정 지점에 대한 상기 실측 이미지에서의 깊이 및 상기 제1 이미지에서의 깊이 사이의 유클리드 거리가 미리 설정된 임계치를 초과하는 경우 해당 지점을 마스크에 포함되는 지점으로 결정할 수 있다.

[0017] 상기 프로그램 명령어들은, 상기 실측 이미지에 이미지 그래디언트를 사용하는 엣지 인식 불일치 정규화를 도입할 수 있다.

[0018] 상기 프로그램 명령어들은, 상기 제1 워프 이미지를 의사 실측 데이터(pseudo ground truth)로 하여 그래디언트 역전파가 상기 제1 이미지를 통해서만 통과하고 상기 제1 워프 이미지에 대해서는 차단되도록 할 수 있다.

[0019] 본 발명의 다른 측면에 따르면, 프로세서 및 메모리를 포함하는 장치에서 소수 샷이 제공되는 환경에서 기하학적 일관성을 이용한 신경 방사장 기반으로 3차원 이미지를 재구성하는 방법으로서, 소정 장면에 대한 소수 샷에 해당하는 N개 시점에서의 실측 이미지를 입력 받는 단계; 신경 방사장 모델 기반으로 하나 이상의 새로운 시점(unseen viewpoint)에서의 제1 이미지 및 제1 깊이 맵을 획득하는 단계; 상기 제1 깊이 맵을 이용하여 상기 실측 이미지를 상기 새로운 시점으로 워프(warp)하여 제1 워프 이미지를 획득하는 단계; 복수의 레이어를 포함하는 컨볼루션 네트워크를 이용하여 상기 제1 이미지 및 상기 제1 워프 이미지 각각의 특징 맵을 추출하는 단계; 상기 각각의 특징 맵에 서로 다른 시점에서의 기하학적 불일치 영역을 무시하기 위한 일관성 마스크를 적용하는 단계; 및 상기 일관성 마스크가 적용된 상기 제1 이미지 및 제1 워프 이미지의 특징 맵 사이의 일관성 손실을 이용하여 상기 신경 방사장 모델이 상기 제1 이미지가 상기 제1 워프 이미지를 따르게 학습하여 상기 소정 장면에 상응하는 3차원 이미지를 재구성하는 단계를 포함하는 3차원 이미지 재구성 장치가 제공된다.

[0020] 본 발명의 또 다른 측면에 따르면, 상기한 방법을 수행하는 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램이 제공된다.

발명의 효과

[0021] 본 발명에 따르면, 희소 관측 설정에서 미세한 세부 정보를 캡처할 수 있고, 희소 관측 설정에도 불구하고 이전 모델에 비해 부드럽게 표면을 렌더링하고 배경에서 artifacts를 줄이는데 더 높은 안정성을 보여주는 장점이 있다.

도면의 간단한 설명

[0022] 도 1은 본 발명의 바람직한 일 실시예에 따른 소수 샷이 제공되는 환경에서 기하학적 일관성을 이용한 신경 방사장 기반 3차원 이미지 재구성 과정을 도시한 도면이다.

도 2는 본 발명의 바람직한 일 실시예에 따른 기하학적 일관성을 고려하는 NeRF 프레임워크를 도시한 도면이다.

도 3은 NeRF가 기하학적 불일치가 있는 영역을 무시하기 위한 일관성 마스크를 나타낸 도면이다.

도 4는 본 실시예에 따른 마스크 생성 과정을 설명하기 위한 도면이다.

도 5는 본 실시예에 따른 3차원 재구성 이미지의 성능을 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

[0023] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0024] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0025] 또한, 각 도면을 참조하여 설명하는 실시예의 구성 요소가 해당 실시예에만 제한적으로 적용되는 것은 아니며, 본 발명의 기술적 사상이 유지되는 범위 내에서 다른 실시예에 포함되도록 구현될 수 있으며, 또한 별도의 설명이 생략될지라도 복수의 실시예가 통합된 하나의 실시예로 다시 구현될 수도 있음은 당연하다.

[0026] 또한, 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일하거나 관련된 참조 부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

[0028] 본 발명은 NeRF를 통해 생성된 깊이를 이용하여 희소 관측 설정에서 입력 이미지를 새로운 시점으로 워핑(warping)시킨 후 실측 데이터(ground truth)를 이용하여 세밀한 3차원 구조와 고주파 텍스처를 학습하도록 하는 모델을 제안한다.

[0029] 이하에서는, NeRF를 먼저 설명한 후, 본 실시예에 따른 희소 관측 환경에서 3차원 이미지 재구성 과정을 상세하

게 설명한다.

[0031] NeRF는 장면(scene)을 파라미터 Θ가 있는 신경망으로 표현되는 연속 함수 f_θ로 표현하며, 여기서 점들은 신경망에서 평가하기 위해 r로 표현되는 광선(ray)을 따라 샘플링된다.

[0032] 일반적으로 샘플링된 좌표 $\mathbf{x} \in \mathbb{R}^3$ 및 시점 방향 $\mathbf{d} \in \mathbb{R}^2$는 위치 인코딩 γ에 의해 푸리에 특징으로 변환되어 고주파 세부 사항의 학습을 용이하게 한다.

[0033] 신경망 f_θ는 아래와 같이, 변환된 좌표 γ(x)와 시점 방향 γ(d)를 입력으로 하여 시점 불변 밀도 값 $\sigma \in \mathbb{R}$과 시점 의존 색상 값 $\mathbf{c} \in \mathbb{R}^3$을 출력한다.

수학식 1

[0034] $$\{\mathbf{c}, \sigma\} = f_\theta(\gamma(\mathbf{x}), \gamma(\mathbf{d})).$$

[0035] $\mathbf{r}_p(t) = \mathbf{o} + t\mathbf{d}_p$로 파라미터화된 광선을 사용하여 카메라 중심 o에서 방향 $\mathbf{d}_p$를 따르는 픽셀 p를 통한 색상은 다음과 같이 렌더링된다.

수학식 2

[0036] $$C(\mathbf{r}_p) = \int_{t_n}^{t_f} T(t)\sigma(\mathbf{r}_p(t))\mathbf{c}(\mathbf{r}_p(t), \mathbf{d}_p)dt,$$
$$\text{where } T(t) = \exp\left(-\int_{t_n}^{t} \sigma(\mathbf{r}_p(s))ds\right),$$

[0037] 여기서 $C(\mathbf{r}_p)$는 t_n에서 t_f까지의 광선 $\mathbf{r}_p(t)$를 따라 픽셀 p에서 예측된 색상 값이고, $T(t)$는 t_n에서 t까지의 광선에 따른 누적 투과율을 나타낸다. 신경망 f_θ를 최적화하기 위해 관측 손실 $\mathcal{L}_{\text{obs}}$는 렌더링된 색상 값이 실측 색상 값 $C'(\mathbf{r})$과 일치하도록 강제한다.

수학식 3

[0038] $$\mathcal{L}_{\text{obs}} = \sum_{\mathbf{r}_p \in \mathcal{R}} \|C'(\mathbf{r}_p) - C(\mathbf{r}_p)\|_2^2,$$

[0039] 여기서 R은 훈련 광선의 배치를 나타낸다.

[0040] i번째 시점의 이미지를 I_i라 하고, 몇 번의 새로운 시점 합성에서 NeRF에는 작은 N(예: N = 3 또는 N = 5)을 갖는 $i \in \{1, \ldots, N\}$에 대한 희소 이미지(소수 샷) $\{I_i\}$만 제공된다.

[0041] 새로운 시점 합성의 목적은 j번째 보이지 않는 또는 새로운 시점에서 이미지 I_j를 복구하는데 사용할 수 있는

매핑 함수 f_θ를 훈련하는 것이다.

[0042] 위에서 설명한 것처럼 $\{I_i\}$가 주어진 소수 샷 설정에서 픽셀 단위 재구성 손실 $\mathcal{L}_{\text{obs}}$만으로 f_θ를 직접 최적화하는 것은 시점 의존 효과를 모델링할 수 없기 때문에 네트워크를 개선하기 위한 추가 정규화가 요구된다.

[0043] 도 1은 본 발명의 바람직한 일 실시예에 따른 소수 샷이 제공되는 환경에서 기하학적 일관성을 이용한 신경 방사장 기반 3차원 이미지 재구성 과정을 도시한 도면이다.

[0044] 도 1에 전체적인 과정은 프로세서 및 메모리를 포함하는 장치에서 수행될 수 있다.

[0045] 여기서, 프로세서는 컴퓨터 프로그램을 실행할 수 있는 CPU(central processing unit)나 그 밖에 가상 머신 등을 포함할 수 있다.

[0046] 메모리는 고정식 하드 드라이브나 착탈식 저장 장치와 같은 불휘발성 저장 장치를 포함할 수 있다. 착탈식 저장 장치는 컴팩트 플래시 유닛, USB 메모리 스틱 등을 포함할 수 있다. 메모리는 각종 랜덤 액세스 메모리와 같은 휘발성 메모리도 포함할 수 있다.

[0047] 본 실시예에 메모리에는 프로그램 명령어들이 저장되며, 프로그램 명령어들은, 소정 장면에 대한 소수 샷에 해당하는 N개 시점에서의 실측 이미지를 입력 받고, 신경 방사장 모델 기반으로 하나 이상의 새로운 시점(unseen viewpoint)에서의 제1 이미지 및 제1 깊이 맵을 획득하고, 상기 제1 깊이 맵을 이용하여 상기 실측 이미지를 상기 새로운 시점으로 워프(warp)하여 제1 워프 이미지를 획득하고, 복수의 레이어를 포함하는 컨볼루션 네트워크를 이용하여 상기 제1 이미지 및 상기 제1 워프 이미지 각각의 특징 맵을 추출하고, 상기 각각의 특징 맵에 서로 다른 시점에서의 기하학적 불일치 영역을 무시하기 위한 일관성 마스크를 적용하고, 상기 일관성 마스크가 적용된 상기 제1 이미지 및 제1 워프 이미지의 특징 맵 사이의 일관성 손실을 이용하여 상기 신경 방사장 모델이 상기 제1 이미지가 상기 제1 워프 이미지를 따르게 학습하여 상기 소정 장면에 상응하는 3차원 이미지를 재구성한다.

[0048] 여기서, 실측 이미지는 소스 시점 이미지, 제1 이미지는 타겟 시점 이미지로 정의될 수 있고, 본 실시예에 따른 프로그램 명령어들은, 랜덤하게 카메라 시점을 샘플링하고, 카메라 시점에서의 광선 벡터를 패치별로 생성하여 상기 제1 이미지를 획득한다.

[0049] 본 실시예에 따르면 일관된 모양과 형상을 생성하기 위한 f_θ를 위해 기하학적 일관성을 강화하는 새로운 정규화 기술을 제안한다.

[0050] 깊이 기반 워핑(warping) 및 일관성 모델링을 통해 다양한 시점에서 본 실시예에 따른 NeRF 프레임워크는 본질적으로 컬러 이미지뿐만 아니라 깊이 이미지(깊이 맵)도 렌더링한다.

[0051] 알려진 시점과의 차이와 결합된 렌더링된 깊이 맵은 임의의 두 시점 간에 기하학적 일치 관계를 정의하는데 사용될 수 있다.

[0052] 본 실시예는 보이지 않는 시점 j에서 NeRF 모델에 의해 렌더링된 깊이 맵 D_j를 고려한다.

[0053] 깊이 맵 D_j와 시점 차이 $R_{i\to j}$에 따라 이미지 I_i를 워프하는 워핑 함수 $\psi(I_i; D_j, R_{i\to j})$를 공식화함으로써 워프 이미지 $I_{i\to j} = \psi(I_i; D_j, R_{i\to j})$ 및 j번째 보이지 않는 시점에서 렌더링된 이미지 I_j 간의 일관성을 촉진하여 소수 샷 환경에서 새로운 시점 합성 성능을 향상시킬 수 있다.

[0054] 이하에서는, 본 실시예에 따른 프레임워크에서 입력 이미지가 보이지 않는 시점으로 어떻게 워핑되는지를 설명한다. 그런 다음 정규화를 위해 워프 이미지와 렌더링된 이미지 쌍에 일관성을 부여하는 방법을 설명하고 폐색(occlusion) 처리 방법과 소수 샷 환경에서 NeRF 최적화의 안정성에 중요한 것으로 입증된 몇 가지 훈련 전략에 대해 설명한다.

[0055] 새로운 시점에서 이미지를 렌더링하기 위해, 먼저 랜덤하게 카메라 시점을 샘플링하고 해당 광선 벡터를 패치별로 생성한다.

[0056] NeRF가 새로운 시점에서의 광선을 따라 샘플링된 지점의 밀도 및 색상 값을 출력하므로 복구된 밀도 값을 사용

하여 일관된 깊이 맵을 렌더링한다.

[0057] 원점에서 이동한 거리의 가중 구성으로 광선당 깊이 값을 공식화한다.

[0058] 픽셀 p에 상응하는 광선 $\mathbf{r}_p$ 는 $\mathbf{r}_p(t) = \mathbf{o} + t\mathbf{d}_p$ 로 파라미터화되므로 깊이 렌더링은 색상 렌더링과 유사하게 정의된다.

수학식 4

[0059] $$D(\mathbf{r}_p) = \int_{t_n}^{t_f} T(t)\sigma(\mathbf{r}_p(t))t dt,$$

[0060] 여기서 $D(\mathbf{r}_p)$ 는 광선 $\mathbf{r}_p$ 를 따라 예측된 깊이이다.

[0061] 도 1에 도시된 바와 같이, 렌더링된 깊이 맵 D_j 를 사용하여 입력 실측 이미지 I_i 를 j번째 보이지 않는 시점으로 역 워핑(inverse Warping)하여 워프 이미지 $I_{i\to j}$ 를 획득하고, 이는 $I_{i\to j} = \psi(I_i; D_j, R_{i\to j})$ 로 정의된다.

[0062] 보다 구체적으로, 보이지 않은 새로운 시점 이미지(타겟 시점 이미지)의 픽셀 위치 p_j 는 시점 차이 $R_{j\to i}$ 및 카메라 고유 파라미터 K에 의해 실측 이미지(소스 시점 이미지)에서 $p_{j\to i}$ 로 변환되어 다음과 같이 표현된다.

수학식 5

[0063] $$p_{j\to i} \sim K R_{j\to i} D_j(p_j) K^{-1} p_j,$$

[0064] 여기서 ~는 대략적인 동등성을 나타내고 투영 좌표 $p_{j\to i}$ 는 연속 값이다. 미분 가능한 샘플러를 사용하여 I_i 에서 $p_{j\to i}$ 의 색상 값을 추출한다. 보다 공식적으로 변환 컴포넌트 프로세스는 다음과 같이 표현할 수 있다.

수학식 6

[0065] $$I_{i\to j}(p_j) = \text{sampler}(I_i; p_{j\to i}),$$

[0066] 여기서 sampler(·)는 쌍선형(bilinear) 샘플링 연산자이다.

[0067] NeRF 체적 렌더링으로 전체 이미지를 렌더링하는 것은 계산 복잡도가 높고 시간이 오래 걸린다.

[0068] 전체 이미지 렌더링 및 워핑의 계산 병목 현상을 극복하기 위해 광선을 스트라이드 그리드에서 샘플링하여 스트

라이드 s를 갖는 패치를 생성하며, 본 실시예에서는 s를 2로 설정한다.

[0069] 체적 렌더링을 거친 후 쌍선형 보간으로 저해상도 깊이 맵을 다시 원래 해상도의 깊이 맵으로 업샘플링한다.

[0070] 전체 해상도 깊이 맵은 역 워핑에 사용된다. 이렇게 하면 원래 크기의 광선 배치를 렌더링할 때 필요한 계산 비용의 일부만으로 전체 해상도의 세부적으로 워프 패치를 생성할 수 있다.

[0071] 이하에서는 일괄성 모델링을 설명한다.

[0072] j번째 시점에서 렌더링된 패치 I_j와 깊이 맵 D_j 및 시점 차이 $R_{i\to j}$에 상응하는 워프 패치 $I_{i\to j}$가 주어지면, 전역적으로 일관된 렌더링을 위한 추가 정규화를 위해 이들 사이의 일관성을 정의한다.

[0073] 실행 가능한 옵션 중 하나는 다음과 같이 간단하게 픽셀별 이미지 재구성 손실(pixel-wise image reconstruction loss) $\mathcal{L}_{\text{pix}}$를 적용하는 것이다.

수학식 7

[0074] $$\mathcal{L}_{\text{pix}} = \|I_{i\to j} - I_j\|.$$

[0075] 그러나 이러한 간단한 적용은 시점에 따라 모양이 크게 변하는 반사 비 램버시안 표면(reflectant non-Lambertian surfaces)에서 실패를 일으키는 경향이 있다.

[0076] 또한 폐색 및 아티팩트와 같은 기하학 관련 문제는 보이지 않는 시점에서 정규화를 위해 픽셀 단위 이미지 재구성 손실을 상기와 같이 적용하는 것이 바람직하다.

[0077] 이러한 문제를 극복하기 위해 본 실시예에서는 시점 의존 방사 효과는 무시하면서 구조적 일관성을 장려하는 마스킹된 특징 수준 정규화 손실을 제안한다.

[0078] 도 2는 본 발명의 바람직한 일 실시예에 따른 기하학적 일관성을 고려하는 NeRF 프레임워크를 도시한 도면이다.

[0079] 도 2를 참조하면, 이미지 I가 입력으로 주어졌을 때, l번째 레이어가 채널 깊이 C_l를 갖는 $f_{\phi,l}(I) \in \mathbb{R}^{H_l \times W_l \times C_l}$과 같은 다단계 특징 맵을 추출하기 위해 컨벌루션 네트워크를 사용한다.

[0080] 워프 이미지 $I_{i\to j}$와 렌더링된 이미지 I_j 사이의 특징 수준 일관성을 측정하기 위해 l번째 레이어에서 특징 맵을 추출하고 동일한 레이어에서 추출된 각 특징 맵 사이의 차이를 계산한다.

[0081] 워프 이미지 $I_{i\to j}$를 의사 실측 데이터(pseudo ground truth)로 사용한다는 아이디어에 따라 그래디언트 역전파가 렌더링된 이미지를 통해서만 통과하도록 허용하고 워프 이미지에 대해 차단되도록 한다.

[0082] 특징 맵의 여러 수준에서 일관성 손실을 적용하여 I_j가 의미론적 및 구조적 수준 모두에서 $I_{i\to j}$를 따르게 모델링하도록 한다.

[0083] 이에 따른 일관성 손실 $\mathcal{L}_{\text{cons}}$는 다음과 같이 정의된다.

수학식 8

$$\mathcal{L}_{\text{cons}} = \sum_{l=1}^{L} \frac{1}{C_l} \left\| f_\phi^l(I_{j\to i}) - f_\phi^l(I_j) \right\|.$$

[0084]

[0085] 손실 함수 $\mathcal{L}_{\text{cons}}$ 에 대해 작업에 가장 적합한 l-1 거리 함수를 찾고 이를 사용하여 특징 맵 차이에서 일관성을 계산한다.

[0086] 경험적으로, 특징 차이의 절대값을 축소하는 정규화 레이어의 부재로 인해 VGG-19 네트워크가 모델링 일관성에서 최상의 성능을 제공한다는 것을 발견했다.

[0087] 따라서 본 실시예에서는 모든 모델에서 VGG19 네트워크를 특징 추출기 네트워크 f_ϕ로 사용한다.

[0088] 본 실시예에 따른 손실 함수는 네트워크에 로컬 구조적 수준에서도 매우 풍부한 정보를 제공할 수 있다. 즉, 높은 수준의 기능 일관성만 제공하는 기존 방식과 달리 특징 차이 계산을 위해 여러 수준의 컨볼루션 네트워크를 사용하는 방법은 높은 수준의 의미적 일관성에서 낮은 수준의 구조적 일관성에 이르기까지 모든 수준의 혼합할 수 있다.

[0089] 잘못된 기하학적 정보로 인해 발생하는 불완전하고 왜곡된 워핑이 모델에 영향을 미쳐 전체 재구성 품질을 저하시키는 것을 방지하기 위해 본 실시예에 따르면, 도 3에서와 같이 NeRF가 기하학적 불일치가 있는 영역을 무시하도록 일관성 마스크 M_l 을 구성한다.

[0090] 이미지에 마스크를 적용하는 대신 특징 추출기 네트워크에 입력하기 전에 가장 가까운 이웃 다운 샘플링을 사용하여 l번째 레이어 출력의 차원과 일치하도록 만든 후 크기 조정된 마스크 M_l을 특징 맵에 직접 적용한다.

[0091] 타겟 시점(시점 j)과 소스 시점(시점 i)에서 렌더링된 깊이 값 간의 일관성을 측정하여 M을 생성한다.

수학식 9

$$M(p_j) = \left[\| D_j(p_j) - D_i(p_{j\to i}) \| < \tau \right].$$

[0092]

[0093] 여기서 [·]는 아이버슨 브래킷(Iverson bracket)이고, $p_{j\to i}$ 는 j번째 시점의 재투영된 타겟 픽셀 p_j 에 대한 소스 시점 i의 대응 픽셀을 가리킨다.

[0094] 여기에서 임계값 마스킹의 기준으로 타겟 및 소스 시점에서 렌더링된 깊이 지점 사이, 즉 픽셀 단위로 소정 지점에 대한 실측 이미지(소스 시점 이미지)에서의 깊이 및 제1 이미지(타겟 시점 이미지)에서의 깊이 사이의 유클리드 거리의 유클리드 거리를 측정한다.

[0095] 본 실시예는 상기 실측 이미지, 제1 이미지 및 워프 이미지의 패치를 이용하여 마스크를 생성하고, 상기와 같이 생성된 마스크를 컨볼루션 네트워크의 복수의 레이어 각각이 출력하는 특징 맵에 적용한다.

[0096] 또한, 픽셀 단위로 소정 지점에 대한 상기 실측 이미지에서의 깊이 및 상기 제1 이미지에서의 깊이 사이의 유클리드 거리가 미리 설정된 임계치를 초과하는 경우 해당 지점을 마스크에 포함되는 지점으로 결정하는 방식으로 마스크를 생성한다.

[0097] 도 4에서와 같이 두 지점 사이의 거리가 주어진 임계값 τ보다 크면 두 개의 광선을 개별 표면의 렌더링 깊이로

결정하고 시점 I_j 에서 해당 픽셀을 마스크 처리한다.

[0098] 상기한 과정은 렌더링된 픽셀과 동일한 크기의 마스크 M 을 생성하기 위해 시점 I_j 의 모든 픽셀에서 일어난다. 이를 통해 특징 수준에서 문제가 있는 부분을 필터링하고 신뢰도가 높은 이미지 특징만으로 NeRF를 정규화한다.

[0099] 이를 기반으로 일관성 손실 $\mathcal{L}_{\text{cons}}$ 는 다음과 같이 확장된다.

수학식 10

$$\mathcal{L}_{\text{cons}}^{M} = \sum_{l=1}^{L} \frac{1}{C_l m_l} \left\| M_l \odot (f_\phi^l(I_{i\to j}) - f_\phi^l(I_j)) \right\|,$$

[0100]

[0101] 여기서 m_l 은 0이 아닌 값의 합이다.

[0102] 본 실시예는 NeRF에 의해 렌더링된 깊이의 품질에 의존하기 때문에 최적화를 용이하게 하기 위해 렌더링된 깊이에 추가 정규화를 도입한다. 또한 임의로 샘플링된 입력 시점 패치 내에서 차이 구배에 l-1 페널티를 도입하여 렌더링된 장면의 로컬 깊이 평활도를 개선한다.

[0103] 또한, 본 실시예에서는 깊이 맵의 깊이 불연속성이 컬러 이미지의 그래디언트에 정렬될 가능성이 있다는 사실을 고려하여 불일치 값에 가중치를 부여하기 위해 이미지 그래디언트 ∂I 를 사용하여 엣지 인식 불일치 정규화(Edge-aware disparity regularization)를 도입하며, 이는 다음과 같다.

수학식 11

$$\mathcal{L}_{\text{reg}} = |\partial_x D_i^*| e^{-|\partial_x I_i|} + |\partial_y D_i^*| e^{-|\partial_y I_i|},$$

[0104]

[0105] 여기서 $D_i^* = D_i / \overline{D_i}$ 는 평균 정규화 역 깊이(mean-normalized inverse depth)이며, 이는 추정 깊이의 축소를 방지한다.

[0106] 도 5는 본 실시예에 따른 3차원 재구성 이미지의 성능을 나타낸 도면이다.

[0107] 도 5를 참조하면, 3차원 이미지에서 본 실시예에 따른 방법이 이전 방법에 비해 미세한 세부 사항(예: 마이크 장면의 와이어)을 더 강건하게 캡처하고 더 적은 아티팩트(재료 장면의 배경)를 생성하는 것을 확인할 수 있다.

[0108] 본 실시예에 따른 3차원 이미지 재구성 방법은, 컴퓨터에 의해 실행되는 애플리케이션이나 프로그램 모듈과 같은 컴퓨터에 의해 실행가능한 명령어를 포함하는 기록매체의 형태로도 구현될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두 포함한다. 또한, 컴퓨터 판독 가능 매체는 컴퓨터 저장 매체를 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독 가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함한다.

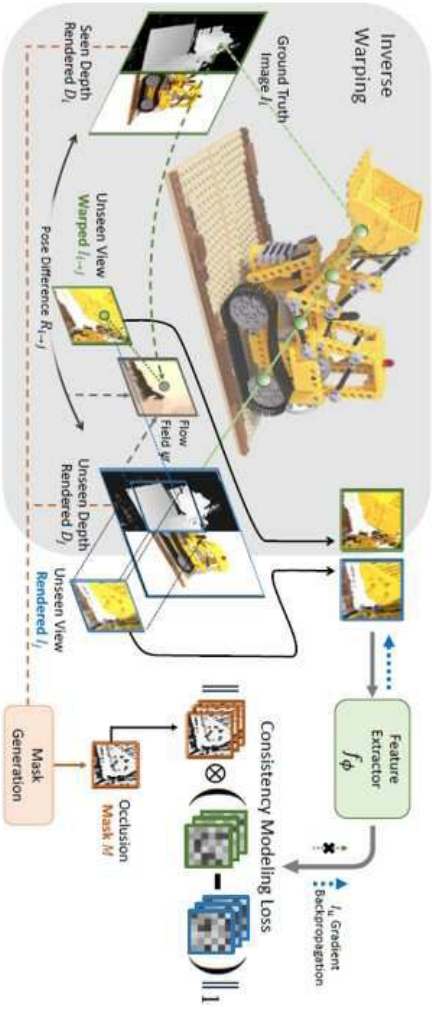
[0109] 전술한 3차원 이미지 재구성 방법은, 단말기에 기본적으로 설치된 애플리케이션(이는 단말기에 기본적으로 탑재된 플랫폼이나 운영체제 등에 포함된 프로그램을 포함할 수 있음)에 의해 실행될 수 있고, 사용자가 애플리케이션 스토어 서버, 애플리케이션 또는 해당 서비스와 관련된 웹 서버 등의 애플리케이션 제공 서버를 통해 마스터 단말기에 직접 설치한 애플리케이션(즉, 프로그램)에 의해 실행될 수도 있다. 이러한 의미에서, 전술한 3차원 이미지 재구성 방법은 단말기에 기본적으로 설치되거나 사용자에 의해 직접 설치된 애플리케이션(즉, 프로그

램)으로 구현되고 단말기 등의 컴퓨터로 읽을 수 있는 기록매체에 기록될 수 있다.

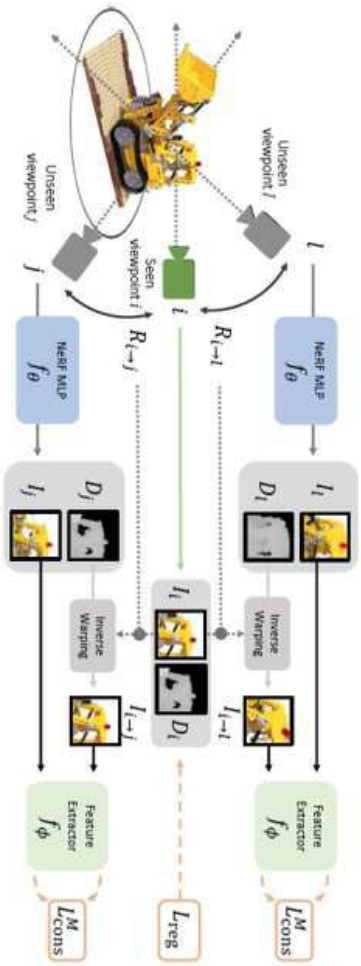
[0110] 상기한 본 발명의 실시예는 예시의 목적을 위해 개시된 것이고, 본 발명에 대한 통상의 지식을 가지는 당업자라면 본 발명의 사상과 범위 안에서 다양한 수정, 변경, 부가가 가능할 것이며, 이러한 수정, 변경 및 부가는 하기의 특허청구범위에 속하는 것으로 보아야 할 것이다.

도면

도면1



도면2

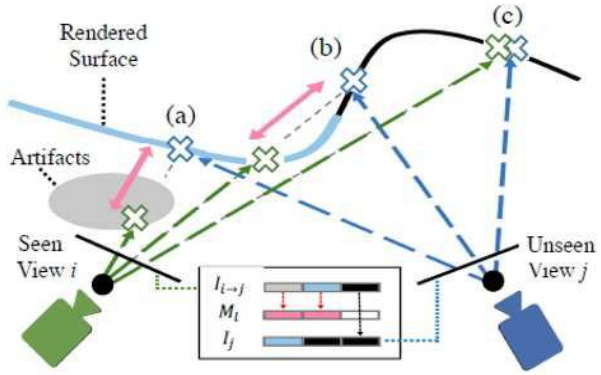


도면3

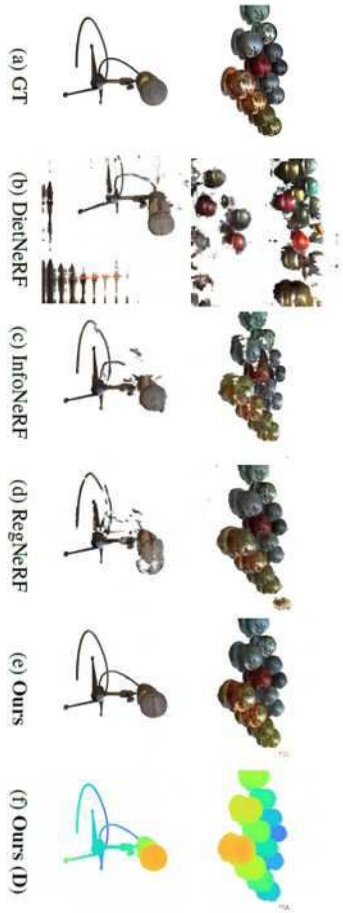


(a) GT patch (b) Rendered patch (c) Warped patch (d) Occlusion mask (e) Masked patch

도면4


Rendered
Surface
(a)
(b)
(c)
Artifacts
Seen
View i
Unseen
View j
$I_{i\to j}$
M_i
I_j

도면5


(a) GT
(b) DietNeRF
(c) InfoNeRF
(d) RegNeRF
(e) Ours
(f) Ours (D)