

Non-Glare NeRF: 빛의 반사가 강한 표면에서 깊이 왜곡 완화

최성준^o, 황효석

경희대학교 소프트웨어융합대학 소프트웨어융합학과

csj000714@khu.ac.kr, hyoseok@khu.ac.kr

Non-Glare NeRF: To prevent unwanted glare

Seongjun Choi^o, Hyoseok Hwang

Software Convergence, Kyung Hee University

csj000714@khu.ac.kr, hyoseok@khu.ac.kr

요약

본 논문은 Neural Radiance Fields(NeRF) 기반 3D 재구성 과정에서 Glare한 표면에서 발생하는 깊이 왜곡 문제를 해결하기 위한 두 가지 접근 방식을 제안한다. 첫째, Inpainting 기법을 활용한 전처리 과정을 도입하여 Glare로 인한 텍스처와 깊이 정보의 왜곡을 완화한다. 둘째, 아이폰의 LiDAR 센서와 RGB 카메라를 활용하여 고정밀 깊이 정보와 풍부한 색상 정보를 결합한 데이터셋을 구축한다. 제안된 방법은 Glare 현상이 존재하는 장면에서도 NeRF의 3D 재구성 성능을 향상시킴을 실험을 통해 확인하였다. 더 많은 정보는 다음의 링크에서 확인할 수 있다: <https://drawingprocess.github.io/Non-Glare-NeRF/>

1. 서론¹

1.1 연구 배경

3D 재구성 기술은 가상현실(VR), 증강현실(AR), 자율주행, 로봇 공학 등 다양한 분야에서 핵심적인 역할을 하고 있다. 특히, NeRF[1]는 신경망을 활용하여 복잡한 장면의 3D 구조를 고해상도로 재구성할 수 있는 강력한 기술로 주목받고 있다. 그러나 NeRF는 빛의 강한 반사로 인한 Glare 현상에 취약하며, 이는 그림 1과 같이 깊이 정보의 손실과 재구성 결과의 왜곡을 초래한다.

Glare 현상은 유리, 금속, 물과 같이 빛의 반사가 강한 표면에서 두드러지며, 이는 NeRF의 정확한 3D 재구성을 방해하는 주요 요인이다. 따라서 Glare 문제를 해결하지 않고서는 고품질의 3D 재구성을 기대하기 어렵다.

1.2 연구 목표

본 연구의 목표는 NeRF 기반 3D 재구성 시 발생하는 Glare 문제를 해결하여 보다 신뢰성 있는 3D 모델을 생성하는 것이다. 이러한 Glare 문제를 해결하기 위해 두 가지 접근 방식을 제안한다.



그림 1. Glare한 표면에서 발생하는 깊이 왜곡 문제

첫째, Inpainting 기법을 활용하여 Glare로 인한 이미지의 왜곡을 전처리 단계에서 완화한다. 둘째, 아이폰의 LiDAR와 RGB 카메라를 활용하여 정확한 깊이 정보를 포함한 데이터셋을 구축한다. 이를 통해 Glare 현상이 존재하는 장면에서도 NeRF의 깊이 추정 성능을 개선하여 평가하고자 한다.

2. 관련 연구

2.1 Neural Radiance Fields (NeRF)

NeRF[1]는 신경망을 이용하여 3D 장면을 복원하는 기술로, 위치와 방향을 입력받아 해당 위치에서의 색상과 밀도를 예측한다. 이는 3D 재구성이 가능하지만 렌더링 성능과 학습 속도가 느리다는 단점이 있다.

¹ "본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2024년도 SW중심대학사업의 결과로 수행되었음"(2023-0-00042)

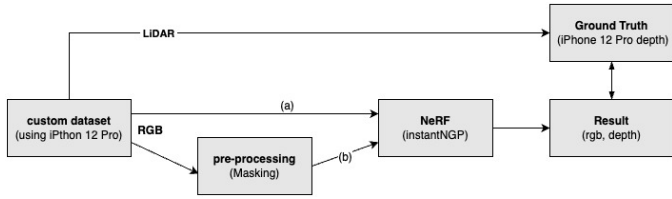


그림 2. 제안한 워크플로우

이러한 한계를 극복하기 위해 Instant Neural Graphics Primitives(iNGP)[2]가 개발되었으며, 이는 NeRF보다 훨씬 빠른 학습 및 추론 속도를 제공한다. 본 연구에서는 제안된 알고리즘의 성능을 평가하기 위해 iNGP 모델을 사용하여 3D 장면을 재구성하였다. 이를 통해 더 효율적이고 신속한 성능 평가를 진행할 수 있었다.

2.2 아이폰 LiDAR 센서와 깊이 정보

LiDAR(Light Detection and Ranging)는 레이저를 사용하여 물체까지의 거리를 측정하는 기술로, 정확한 깊이 정보를 제공한다. 아이폰의 LiDAR 센서는 실시간으로 주변 환경의 깊이 정보를 수집할 수 있으며, RGB 카메라와 결합하여 풍부한 데이터셋을 구축하는 데 활용될 수 있다[3]. 또한 별도의 추가적인 캘리브레이션 없이 iOS 라이브러리에서 자동으로 포즈 추출(extrinsic parameters)이 가능하다. 따라서 전문적인 3D 스캔 장비나 고가의 LiDAR 센서에 비해 저렴한 비용으로 고품질의 데이터를 수집할 수 있다는 장점이 있다.

2.3 Glare 문제와 Inpainting 기법

Glare 현상은 빛의 강한 반사로 인해 이미지의 일부 영역이 과도하게 밝아지는 현상으로, 해당 영역의 텍스처와 깊이 정보가 손실된다. 이러한 문제를 해결하기 위해 이미지 복원 기법 중 하나인 Inpainting이 활용될 수 있다. Inpainting은 손상된 이미지의 영역을 주변 정보로부터 복원하는 기술로, 최근에는 딥러닝 기반의 방법들이 활발히 연구되고 있다[3].

3. 본론

본 연구에서는 Glare 문제를 해결하기 위해 두 가지 접근 방식을 제안하였다. 첫째, Inpainting 기법을 활용하여 Glare 영역을 복원하는 전처리 과정을 도입하였다. 둘째, 아이폰의 LiDAR 센서와 RGB 카메라를 활용하여 깊이 정보가 포함된 데이터셋을 구축하였다. 이를 통해 NeRF 모델의 학습 및 평가를 진행하였다.

3.1 전체 시스템 구성

본 연구에서는 방법론의 효과를 입증하고자 그림 2와 같은 워크플로우를 제시한다. 제안한 워크플로우는 (a) iNGP 모델 결과 추출과 (b) pre-processing을 진행한 iNGP 모델 결과 추출의 두 가지 단계로 구성된다. 이후 두 결과를 비교하여 평가를 진행하고자 한다.

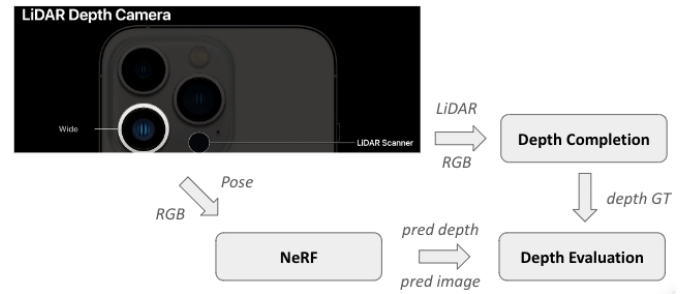


그림 3. 아이폰 12 프로를 활용한 데이터 엔진 구축

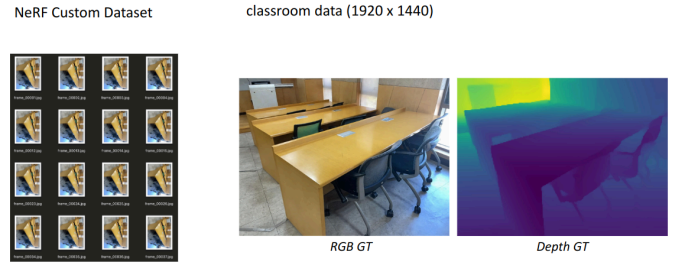


그림 4. 데이터 엔진을 활용한 데이터셋 구성



그림 5. Threshold 기반의 Glare 영역 추출

3.2 LiDAR와 RGB를 활용한 데이터셋 구축

아이폰의 LiDAR 센서와 RGB 카메라를 활용하여 정확한 깊이 정보와 색상 정보를 포함한 데이터셋을 구축하였다.

1) 데이터 수집: 아이폰을 사용하여 특정 장면을 촬영하고, 각 프레임의 RGB 이미지와 카메라 포즈를 추출하였다.

2) 깊이 정보 획득: 그림 3과 같이 LiDAR 센서로부터 얻은 깊이 정보를 RGB 이미지와 결합하여 Depth Ground Truth를 생성하였다. 이때, Depth Completion 기법을 적용하여 LiDAR 데이터의 빈틈을 보완하였다.

3) 데이터셋 구성: 생성된 RGB 이미지와 깊이 정보를 활용하여 NeRF 모델의 학습 및 평가에 사용될 데이터셋을 그림 4와 같이 구축하였다.

3.3 Inpainting 기반 전처리 과정

Glare 영역을 복원하기 위해 다음을 수행하였다.

1) Glare 영역 추출: Local Thresholding 기법을 사용하여 이미지에서 Glare로 인한 과도한 밝기 영역을 추출하였다. 이때, Morphological Operations를 통해 노이즈를 제거하고 Glare 영역의 윤곽을 그림 5와 같이 명확히 하였다.

RGB GT data (1920 x 1440)

Glare inpainting SOTA: co-mod-gan¹

그림 6. Glare 영역의 inpainting 결과 시각화

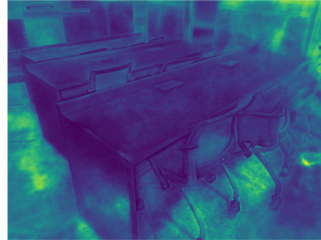
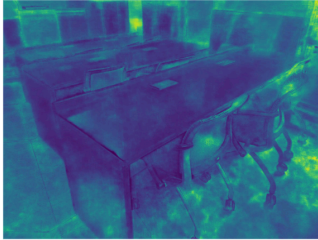


그림 7. 깊이 생성: 정성 평가 결과 (좌: iNGP, 우: Ours)

2) Inpainting 적용: 추출된 Glare 영역에 대해 다양한 Inpainting 알고리즘을 적용하였다. 여러 기법 중에서 co-mod-gan[4] 알고리즘이 가장 우수한 성능을 보여, 이를 전처리에 활용하였다.

3) 전처리된 이미지 생성: Inpainting 기술의 SOTA 모델인 co-mod-gan[2] 알고리즘을 통해 그림 6과 같은 Glare 영역이 복원된 이미지를 생성하고, 이를 NeRF 모델의 학습 데이터로 사용하였다.

4. 실험

4.1. 실험 환경, 모델, 성능 평가 지표

본 연구는 Intel Xeon E-2334, GeForce RTX 3090 환경에서 실험을 진행하였다. NeRF 모델은 기존 방식과 제안된 방법을 비교하기 위해 동일한 데이터셋인 아이폰으로 수집한 데이터셋을 기반으로 진행되었다.

성능 평가는 Absolute Relative Error(AbsRel), Squared Relative Error(SqRel), Root Mean Square Error(RMSE), Root Mean Square Error Log(RMSELog), Accuracy under a threshold (δ) 를 사용하였다.

4.2 정성 평가

NeRF 모델의 성능을 시각적으로 평가하기 위해 Glare가 포함된 장면에서의 3D 재구성 결과를 비교하였다. 제안된 방법론을 통해 그림 7에 보이듯이 Glare 영역이 효과적으로 복원되었으며, 기존 방식과 비교하여 훨씬 자연스러운 텍스처와 깊이 정보를 얻을 수 있었다.

표 1. 깊이 생성: 정량 평가 결과

	AbsRel↓	SqRel↓	RMSE↓	RMSELog↓	δ ↑
iNGP	0.40	0.26	0.62	0.62	0.20
Ours	0.41	0.22	0.60	0.59	0.22

4.3 정량 평가

정량적으로 NeRF와 제안한 방법의 성능을 비교한 결과, 제안된 방법이 iNGP보다 전반적으로 더 우수한 성능을 보였다. 표 1은 실험 결과를 요약한 것으로, SqRel 및 RMSELog 지표에서 성능 향상이 특히 두드러졌다. 이를 통해 제안된 방법이 스케일을 반영한 깊이 정보 획득에 있어 더 뛰어난 성능을 발휘했음을 확인할 수 있었다.

5. 결론

본 논문에서는 NeRF 기반 3D 재구성에서 발생하는 Glare 문제를 해결하기 위해 Inpainting 기반 전처리와 iPhone의 RGB와 LiDAR 센서를 활용한 데이터셋 구축을 제안하였다. 실험 결과, 제안한 방법이 Glare 현상이 존재하는 장면에서도 NeRF의 성능을 향상시킴을 확인하였다. 이러한 연구는 다양한 조명 조건에서 신뢰성 있는 3D 재구성 기술을 구현하는 데 기여할 것으로 기대된다.

참고 문헌

- [1] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2020). NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. European Conference on Computer Vision (ECCV).
- [2] Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. ACM SIGGRAPH Conference on Computer Graphics and Interactive Techniques.
- [3] Apple Inc. (2021). iPhone 12 Pro and iPhone 12 Pro Max: The ultimate smartphone experience.
- [4] Zhao, Y., Dai, B., & Zhu, J. (2021). Large Scale Image Completion via Co-Modulated Generative Adversarial Networks. International Conference on Learning Representations (ICLR).