

407.20194v2

로봇 원격 조작을 위한 Radiance Fields

맥시엄 와일더 스미스, 바이샤크 파틸, 마르코 후터

초록— 신경 복사와 같은 복사장 방법

필드(NeRF) 또는 3D 가우시안 스피클레팅(3DGS)은 그래픽과 새로운 뷰 합성에 혁명을 일으켰습니다. 사실적인 품질로 새로운 시점을 합성하고 복잡한 체적 및 반사 장면을 캡처하는 기능은 로봇 원격 조작 설정에 이상적인 시각화를 제공합니다. 직접 카메라 원격 조작은 기동성을 희생하는 대신 충실도가 높은 조작을 제공하는 반면, 재구성 기반 접근 방식은 충실도가 낮지만 제어 가능한 장면을 제공합니다. 이를 염두에 두고, 로봇 원격 조작 파이프라인의 기존 재구성-시각화 구성 요소를 온라인 Radiance Fields로 대체하여 사실적인 품질의 매우 기동성 있는 장면을 제공할 것을 제안합니다. 따라서 최첨단 기술에 대한 세 가지 주요 기여는 다음과 같습니다. (1) 여러 카메라의 라이브 데이터를 사용한 Radiance Fields의 온라인 학습, (2) NeRF 및 3DGS를 포함한 다양한 방사 기법 지원, (3) 가상 현실 장면을 포함한 이러한 기법을 위한 시각화 모음. 기존 설정과의 원활한 통합을 위해, 이러한 구성 요소들을 다양한 구성의 여러 로봇으로 테스트하고 기존 도구와 VR 헤드셋을 사용하여 구현했습니다. 다양한 방법과 로봇에 대한 결과는 메시 재구성 기준과 정량적으로 비교되었으며, 다양한 시각화 방법을 비교하기 위한 사용자 연구가 수행되었습니다. 코드와 추가 샘플은 <https://rfr.leggedrobotics.com/works/teleoperation/>에서 확인할 수 있습니다.



그림 1. 재구성된 Radiance Field 내부에서 VR 인터페이스를 사용하여 로봇을 제어하는 원격 조작자. 재 구성은 센서 구성을 기반으로 모든 로봇에서 온라인으로 생성됩니다. Neural Radiance Field 또는 Gaussian Splatting 렌더링은 몰입형 360° 렌더링 또는 휴대용 뷰어에 표시할 수 있습니다. 이러한 렌더링은 로봇 및 카메라 피드, LiDAR와 같은 센서 데이터와 함께 표시됩니다.

I. 서론

로봇 공학은 빠르고 반복적이며 정확한 동작을 요구하는 일상생활에서 그 역할을 빠르게 확대하고 있습니다. 로봇은 진공청소기부터 일상적인 통근 차량에 이르기까지 고도로 자율화되고 있습니다. 그러나 자율 로봇은 다양한 작업에 대한 인식, 계획 및 제어 능력에 한계가 있습니다. 이러한 로봇의 성공적인 실행을 위해서는 여전히 인간의 개입이 필요합니다. 로봇을 원격으로 제어하는 과정을 로봇 원격조종이라고 합니다. 로봇이 복잡한 환경에 점점 더 통합됨에 따라, 고충실도 재구성과 직관적인 사용자 인터페이스를 제공하는 원격조종 시스템의 필요성이 커지고 있습니다. 로봇을 인간이 원격조종할 수 있도록 시스템은 조작자에게 상황 인식을 제공할 수 있는 몰입형 환경을 조성해야 합니다.

기존의 원격 조작 파이프라인은 로봇의 센서 데이터를 수집하여 운영자에게 표시합니다. 직접 카메라 스트림은 기동성을 희생하는 대신 높은 충실도를 제공하는 반면, 재구성된 환경은 제어가 쉽지만 특히 체적 또는 반사 환경에서 장면의 형상과 질감을 정확하게 포착하는 데 어려움을 겪는 경우가 많습니다. 재구성 품질은 높은 센서 노이즈와 시간적 불일치로 인해 제한되는 경우가 많습니다[1]. 또한, 사용자 인터페이스가 부족할 수 있습니다.

저자는 취리히 연방공과대학교(ETH Zurich) 로봇 시스템 연구실 소속입니다. 이메일: {mwilder, patilv, mahutter}@ethz.ch 본 연구는 ETH 모빌리티 이니셔티브를 통해 NCCR Automation과 스위스 연방 철도(SBB)의 자금 지원을 받았습다.

몰입감과 직관적 상호작용이 부족하여 작업자의 효과적인 작업 수행 능력이 제한됩니다. 제안된 파이프라인은 Radiance Fields와 몰입형 시각화 기술의 발전을 활용하여 충실도가 높고 조작이 가능한 장면을 구현함으로써 이러한 문제를 해결합니다.

Neural Radiance Fields(NeRF) [2]의 기반은 희소한 포즈 이미지로부터 새롭고 사실적인 뷰를 렌더링할 수 있는 능력 덕분에 3D 재구성에 있어 상당한 도약을 이루었습니다. NeRF [3], [4], [5]의 최근 발전은 학습 및 렌더링 시간을 크게 단축시켜 원격 조작 시스템의 시각화 및 몰입도를 향상시킬 수 있는 훌륭한 후보가 되었습니다.

그러나 NeRF는 원격 조작에 필수적인 계산 효율성과 장면 제어성[6] 측면에서 여전히 어려움을 겪고 있습니다. 이러한 과제를 해결하기 위해 효율적인 렌더링 방식을 제공하는 3D 가우시안 스피클레팅(3DGS)[6]이라는 또 다른 기술이 등장했습니다. NeRF와 달리 3DGS는 명시적인 방사 필드 표현과 고도로 병렬화된 처리를 결합하여 효율적인 계산 및 렌더링 시간을 달성합니다. 이는 특히 기하 구조의 정밀한 조작과 조명 조건에 대한 적응이 필수적인 복잡한 환경에서 유용합니다.

이를 염두에 두고, 본 연구는 Radiance Fields와 실시간 로봇을 통합하는 새로운 동적 시각화 파이프라인을 제시함으로써 로봇 원격조작 분야의 최첨단 기술을 발전시키는 것을 목표로 합니다. 무엇보다도, 본 시스템은 Radiance Field Node 형태의 실시간 로봇 데이터로부터 고품질 Radiance Fields를 학습시킬 수 있습니다. 이 시스템은 간단한 로봇 팔부터 고도로 이동 가능한 네 발 로봇까지 모든 ROS 기반 로봇의 데이터를 처리할 수 있을 만큼 동적입니다. 둘째,

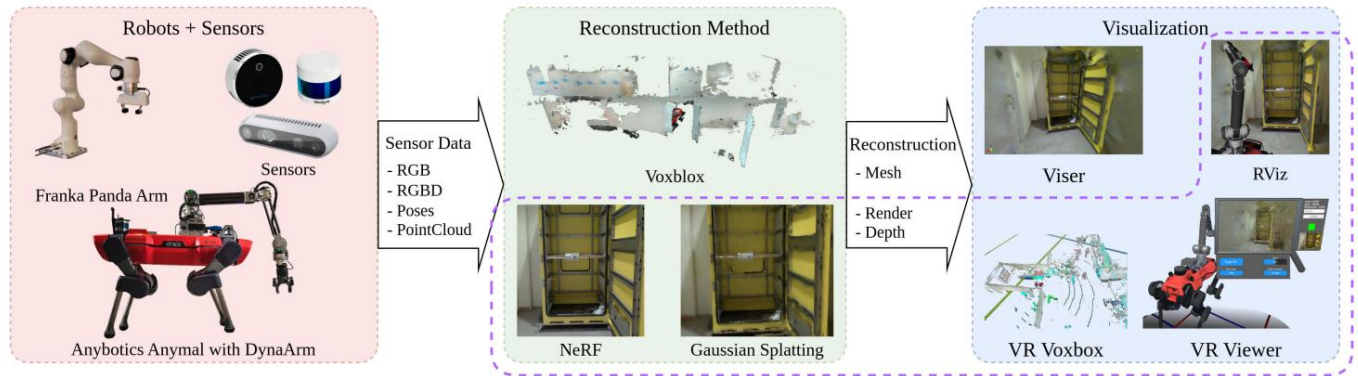


그림 2. 로봇, 재구성 방법, 시각화의 세 부분으로 구분된 일반적인 원격 조작 시각화 파이프라인. 센서 및 자세 데이터는 다양한 로봇 구성 요소(빨간색)에서 재구성 방법(녹색)으로 이동하여 시각화 도구(파란색)에서 사용자에게 표시되는 장면 표현을 생성합니다. 본 연구에서는 NeRF(Neural Radiance Fields) 및 가우시안 스플래팅과 같은 복사장 재구성 지원과 이러한 방법을 위한 RViz 및 VR 시각화 도구를 제시합니다. 기준선 비교를 위해 Voxblox 메시 뷰어도 VR로 이식되었습니다. 이러한 기여는 보라색 점선으로 강조되어 있습니다.

신경 렌더링의 급속한 발전에 발맞추기 위해 이 시스템은 NerfStudio[7]와 긴밀하게 통합되어 가우시안 스플래팅을 포함한 다양한 방법을 지원할 수 있습니다. 마지막으로, 이 시스템은 원격 조작자가 Radiance Fields를 통해 주변 환경을 이해할 수 있도록 하는 시각화 옵션 세트를 제공합니다. 여러 렌더링 모드, 장면 자르기, 깊이 정보 클리핑을 갖춘 대부분의 기존 ROS 원격 조작 시스템과 완벽하게 통합되는 RViz 호환 Radiance Field 플러그인 시스템을 제공합니다. 더욱 몰입감 있는 사용을 위해 VR 시각화 세트가 설계되어 사용자가 장면의 가상 복제본에 서거나 조감도 시스템을 통해 명령을 보낼 수 있습니다. 궁극적으로 이 연구는 원격 조작과 Radiance Fields를 결합하는 것뿐만 아니라 로봇 공학 전반에 온라인 Radiance Fields를 사용하는 새로운 가능성을 열어줍니다.

진보로 인해 고주파 세부 정보[3], 제한 없는 뷰 구성[4]이 가능해졌고 작은 장면에서 작업 속도가 거의 실시간으로 빨라졌습니다[5].

NeRF보다 빠른 대안으로, 가우시안 스플래팅(Gaussian Splatting, 3DGS) [6]은 실시간 렌더링을 위해 더욱 명확한 장면 표현을 활용합니다. 3D 가우시안으로 장면을 표현하면 시점 의존적 색상, 입체적 요소, 그리고 무한한 장면을 더 빠르게 표현할 수 있습니다. 하지만 NeRF의 장면 충실도는 부족하며, 특히 반사 물체와 날카로운 모서리가 있는 복잡한 반사에서 그렇습니다.

최근 NeRF 기법의 폭발적인 증가는 NerfStudio[7]의 등장으로 더욱 가속화되었습니다. NerfStudio는 Radiance Field 기법을 구축, 학습 및 적용하기 위한 포괄적인 시스템을 제공합니다. 여기에는 Nerfacto(NeRF 구현) 및 Splatfacto(Gaussian Splatting 구현)와 같은 기법뿐만 아니라 Instant-NGP의 변형 기법, 깊이 지도 학습 기법 등 다양한 기법이 포함됩니다.

II. 관련 저작

이 섹션에서는 원격 조작 및 재구성 방법과 Radiance Fields의 최근 작업에 대한 간략한 개요를 제공합니다.

1) 원격 조작 및 재구성: 로봇이 산업 현장[8]이나 동굴 및 기타 자연 환경[9]과 같이 더욱 복잡한 용도로 배치됨에 따라 원격 조작에 대한 필요성이 증가하고 있습니다. 이러한 환경에서는 센서 데이터가 직접 카메라 피드[10] 또는 장면 재구성[11]의 형태로 운영자에게 스트리밍됩니다. 최첨단 장비는 직접 카메라 시스템의 높은 충실도를 선호합니다[10]. 그러나 고정된 센서 피드로 인해 기동성이 제한되는 단점이 있습니다.

한편, 재구성 기반 원격조종은 조작하기 쉬운 장면 표현을 구축하지만[11], [12], 종종 충실도가 낮거나 성능이 떨어지는 문제가 발생합니다[13]. 최상의 환경에서도 이러한 시스템은 연기와 같은 체적을 표현하지 못하고 반사 효과와 같은 시점 의존적인 데이터를 무시합니다. Radiance Fields는 충실도가 높고 조작이 용이한 장면 표현으로 연구되고 있습니다.

2) Radiance Fields: 원래는 사실적인 화질로 새로운 뷰를 렌더링하는 데 사용되었던 Radiance Fields는 최근 대규모 3D 재구성에 적용되었습니다. NeRFs(Neural Radiance Fields)[2]는 지오메트리, 색상 및 밀도를 MLP(Multiple Point Representation)로 인코딩하여 작은 장면을 캡처할 수 있으며, 이를 통해 뷰 종속 색상 및 체적 렌더링이 가능합니다. 빠른

최근까지 로봇공학과 Radiance Fields의 융합은 긴 학습 시간과 모바일 하드웨어로의 이식 어려움으로 인해 제한적이었습니다. 그 대신, Radiance Fields는 주로 오프라인 데이터를 이용한 위치 추정[14]이나 오프라인 경로 계획[15]에 사용되었습니다.

NerfBridge[16]에서는 NeRF의 온라인 통합이 가능하지만, 다중 카메라 및 새로운 기법에 대한 지원이 부족합니다. 본 연구에서는 3DGS와 같은 새로운 기법을 쉽게 지원하는 온라인 다중 카메라 Radiance Field 재구성을 통해 이러한 문제를 해결합니다.

III. 방법론

이상적인 원격 조작 시스템은 조작자에게 현장에 대한 충실도 높고, 기동성이 뛰어나며, 충실한 표현을 제공해야 합니다[17]. 이 시스템은 최소한의 재구성으로 로봇에 구매받지 않아야 하며, 작업에 가장 적합한 환경 데이터(센서 스트림 또는 재구성)를 제공할 수 있어야 합니다[1]. 그림 2에 표시된 확장된 3D 재구성 기반 파이프라인은 ROS를 통해 로봇의 센서 데이터를 가져와 재구성 방법에 입력하는 것을 포함합니다[11], [7]. 이 방법은 고정된 위치를 기준으로 환경의 표현을 생성하여 시각화 장치를 통해 조작자에게 제공합니다. 표현을 구성함으로써

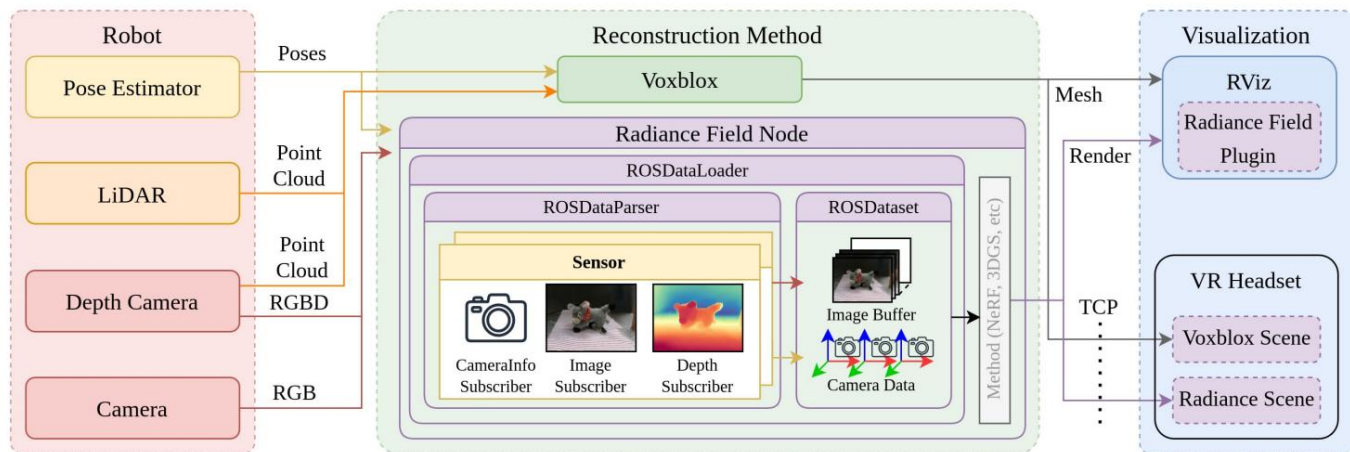


그림 3. 로봇 시스템(빨간색과 주황색), 재구성 방법(녹색), 그리고 시각화 장치(파란색) 간의 연결을 보여주는 구성 요소 다이어그램. 로봇 시스템에서 생성된 데이터는 재구성 방법 노드로 이동하며, 여기서 Radiance Fields의 렌더링된 뷰 또는 Voxelbox의 메시로 병합됩니다. Radiance Field 노드는 사용자 지정 DataLoader, DataSet, 그리고 Dataset으로 구성되며, Sensor 객체들을 사용하여 ROS 구독자를 관리합니다. DataLoader와 DataSet은 학습 과정에서 모든 표준 방법으로 데이터를 전송하며, 업데이트된 이미지와 카메라만 반환합니다. 이 데이터는 RViz에 직접 표시되거나 TCP를 통해 VR 헤드셋으로 전송됩니다.

고정된 지점을 기준으로 로봇이 공간을 이동하는 동안에도 추가 데이터를 표시하고 정렬할 수 있습니다.

전반적으로 이 파이프라인은 로봇과 센서, 재구성 방법, 시각화 시스템의 세 가지 구성 요소로 단순화할 수 있습니다.

본 연구에서는 기존 원격 운영 파이프라인을 기반으로 Radiance Fields를 포함한 가능한 재구성 방법을 확장하고, 이를 화면과 VR 헤드셋 모두에서 시각화하는 시스템을 제시합니다. 이 시스템은 그림 2에서 보라색으로 강조 표시되어 있습니다. 이 시스템은 간단한 정적 구성, 더 넓은 환경을 탐색할 수 있는 이동식 사족보행 로봇, 그리고 팔이 부착된 사족보행 로봇 등 여러 로봇을 사용하여 테스트되었습니다. 이 로봇 데이터는 기존 메시 재구성을 기준으로, 그리고 NeRF 및 3DGS를 지원하는 Radiance Fields용 새로운 ROS 노드로 전송됩니다. 이러한 재구성 방법은 화면의 2D RViz 창부터 그림 7에 표시된 VR 헤드셋의 2.5D 및 3D 뷰까지 다양한 시각화 도구에서 사용할 수 있습니다. 이 파이프라인의 각 구성 요소는 배포 및 작업에 따라 교체하거나 재구성할 수 있습니다.

A. 로봇

이 시스템에서 로봇의 목적은 사용자가 장면을 이해할 수 있도록 환경을 캡처하는 것입니다. 이 파이프라인에 사용되는 주요 데이터는 ROS TF 시스템의 포즈, 온보드 카메라의 컬러 이미지, 그리고 깊이 데이터와 색상 정보가 포함된 RGBD 이미지입니다.

시스템이 로봇에 구매하지 않도록 이러한 데이터 소스를 모든 배포 환경에 맞게 쉽게 구성할 수 있습니다. 그림 2와 그림 3에서 수신 데이터 소스는 "로봇" 상자 왼쪽에 표시되어 있으며, 빨간색 선은 이미지 데이터, 주황색 선은 포인트 클라우드 데이터, 노란색 선은 TF 시스템의 포즈를 나타냅니다.

가장 간단한 로봇 구성은 테이블에 부착된 로봇 팔과 같은 접지형 로봇입니다. 이 로봇은 이동성이 제한적이며, 특히 대상 물체나 대규모 장면을 스캔할 때 그렇습니다. 하지만 로봇의 기본 프레임이 고정된 전역 프레임이기 때문에 매우 정확한 자세를 제공합니다.

프레임입니다. 이는 장면의 크기와 시야각을 희생하더라도 재구성 시 높은 포즈 신뢰도를 제공합니다.

더 크고 복잡한 장면을 포착하려면 모바일 로봇이 필요합니다. 이러한 설정에서 로봇은 환경 내에서 이동하며 어떤 각도에서든 가상 데이터를 수집할 수 있습니다. 그러나 로봇과 고정된 전역 프레임 간의 관계는 불확실하기 때문에 모바일 로봇은 환경 내에서 위치를 파악해야 합니다.

이는 CompSLAM[9]과 같은 SLAM(동시 위치 측정 및 매핑) 시스템을 사용하여 달성할 수 있습니다.

부정확한 내장 센서에 의존하기 때문에 이동 중 자세가 흐트러질 수 있습니다. 이동식 베이스는 더 넓은 범위를 커버할 수 있지만, 자세가 항상 신뢰할 수 있는 것은 아닙니다.

높은 자세 정확도와 이동성을 갖춘 더 넓은 장면을 촬영할 수 있도록 돕는 한 가지 방법은 로봇 팔을 모바일 플랫폼에 부착하는 것입니다. 이 설정에서 로봇은 새로운 시각각으로 장면을 이동할 수 있으며, 주로 베이스가 고정된 상태에서 팔의 움직임에서 데이터를 수집합니다. 로봇 베이스가 고정되어 있으므로 엔드 이펙터의 움직임은 정밀한 운동학 관절 센서에서만 측정할 수 있습니다. 이를 통해 시각적 또는 LiDAR 위치 측정 오류를 방지할 수 있습니다.

B. 재건 방법

로봇이 데이터를 캡처하면 재구성 방법으로 전달됩니다. 이 단계의 목적은 수많은 데이터 스트림을 작업자가 로봇을 제어하는 데 사용할 수 있는 단일 스트림으로 변환하는 것입니다. 여기에는 기하학과 텍스처 총실도 측면에서 모두 정확한 표현을 만드는 것이 포함됩니다. 이 작업에서는 이 문제에 대한 두 가지 주요 접근 방식인 Voxelbox[11]를 통한 메시 재구성과 NerfStudio[7]를 통한 Radiance Field 재구성을 비교합니다. 이 두 ROS 노드는 그림 2와 그림 3의 "재구성 방법" 섹션에서 녹색으로 표시됩니다. Voxelbox 노드는 들어오는 포즈 데이터와 포인트 클라우드를 사용하여 출력 메시를 생성하는 반면 NerfStudio 노드는 포즈와 이미지 데이터에서 렌더링된 Radiance Field를 생성합니다. 메시 재구성은 복잡한 체적 장면을 캡처하는 데 어려움을 겪고 빠른 재구성과 높은 충실도 간의 균형을 맞추기 때문에 로봇 원격 조작에 Radiance Fields를 사용할 것을 제한합니다.

Radiance Fields는 가벼운 머신 러닝 알고리즘을 사용하여 장면의 표현을 학습하려고 합니다.

Radiance Field 노드: 새로운 메서드와의 상호 운용성을 유지하기 위해 파이프라인의 Radiance Field 학습 노드는 NerfStudio에서 지원하는 거의 모든 메서드를 사용하도록 설정할 수 있습니다[7]. NerfStudio 파이프라인의 사용자 지정 구성 요소 수는 사용자 지정 데이터셋(예: 의미 모델)이나 사용자 지정 DataLoader를 필요로 하지 않는 한 모든 메서드와의 호환성을 보장하기 위해 최소화되었습니다.

이는 NerfStudio의 3DGS 방식인 Splatfacto를 최소한의 오버헤드로 지원했습니다. 이 맞춤형 데이터셋은 처음에 학습에 필요한 모든 이미지를 저장할 미리 구성된 크기의 버퍼를 할당합니다.

그런 다음 노드는 각 카메라의 주제 집합을 구축하여 이미지(RGB 또는 RGBD), 내장 함수 및 포즈를 캡처합니다.

모든 카메라에 깊이 토크이 있는 경우, 깊이 이미지를 저장하기 위한 특수 버퍼가 할당되어 깊이 지도 모델과의 통합이 가능합니다. 각 입력 이미지는 흐릿하지 않고 TF 프레임 간 평균 거리를 기반으로 충분히 다른 포즈를 갖도록 필터링됩니다. 카메라 해상도가 다른 경우, 모든 이미지는 가장 큰 카메라의 해상도로 조정되고, 배치로 통합될 때 다시 축소됩니다. 실험에서는 이미지를 1440p로 저장하고, 720p 이미지는 훈련 전에 크기를 조정했습니다.

각 이미지에 대해 현재 포즈 데이터와 마지막 CameraInfo 메시지를 기반으로 연관된 카메라별 내재적 및 외재적 정보가 저장됩니다. 이미지가 이미 보정되었고 토크 이름에 "rect"가 포함된 경우, 보정을 방지하기 위해 연관된 카메라의 왜곡 매개변수가 0으로 설정됩니다.

다른 메서드와의 호환성을 보장하기 위해 이 패키지는 DataLoader를 사용자 지정 ROS 변형으로 재정의합니다. 이 변형은 일반적으로 수행되는 캐싱 및 데이터 정렬을 건너뛰고 대신 ROS Datasets 함수를 사용하여 업데이트된 이미지만 학습에 가져오도록 합니다. 또한, 이미지 시퀀스 ID 목록을 받아 여러 실행에서 동일한 이미지를 사용할 수 있도록 하는 특수 평가 DataLoader도 있습니다. 이는 특히 미리 기록된 ROS 백 데이터 간의 비교를 실행할 때 유용합니다. 또한, 사용자 지정 DataParser를 사용하여 ROS Dataset과 ROS 구독자를 설정하고, 이는 도우미 Sensor 클래스를 통해 관리됩니다. 이러한 Sensor 객체는 이미지 및 CameraInfo 메시지 구독자를 관리하고, 각 캡처에서 동기화된 포즈를 조화합니다.

이 도우미 클래스를 사용하면 여러 대의 카메라를 등록하여 각 카메라가 서로 다른 속도로 데이터를 수집할 수 있습니다. 그러면 센서 객체는 적절한 데이터 항목만 업데이트하고, 관련 카메라 매개변수를 시스템의 나머지 부분으로 전달합니다. 그림 3은 데이터 흐름을 보여주며, 모든 사용자 지정 구성 요소는 보라색으로, 센서 도우미 클래스는 노란색으로 강조 표시되어 있습니다. 모든 설정이 완료되면 노드는 온라인 학습 모드로 전환하기 전에 여러 이미지를 캡처하여 초기 배치를 생성합니다.

노드가 온라인 훈련 모드에 진입하면 지정된 버퍼 크기까지 새 이미지를 계속 수신하고, 모든 ROS 프로세스가 렌더링 데이터를 요청할 수 있도록 하는 ActionServer를 시작합니다. 뷰 포즈가 포함된 렌더링 요청이 서버로 전송되고, 렌더링된 이미지와 추정된 장면 깊이가 반환됩니다. 이러한 렌더링은 이후 전송됩니다.

원격 조작자가 상호작용할 수 있는 시각화 도구 중 하나에 연결합니다.

또한, 요청은 고유한 클라이언트 ID와 연결되어 여러 시각화 프로그램이 ROS 네트워크의 단일 Radiance Field 노드를 사용하여 서로 다른 렌더링을 요청할 수 있습니다.

C. 시각화

파이프라인의 마지막 단계는 시각화 프로그램입니다. 이 프로그램은 재구성 방법을 통해 수신된 데이터를 기반으로 운영자에게 정확한 세계 표현을 제공합니다. 그림 2와 그림 3에서 파란색으로 표시된 부분은 저희가 직접 제작한 Radiance Field RViz 플러그인이 2D 시각화 도구로, VR 장면이 2.5D 및 3D 시각화 도구로 사용되는 모습입니다. 기준선 비교를 위해 Voxblox RViz 플러그인을 사용하여 메시 재구성을 보여줍니다.



그림 4. 장면 깊이에 따라 로봇을 가리는 RViz 플러그인 샘플(왼쪽)과 로봇을 맨 위에 렌더링하는 RViz 플러그인 샘플(오른쪽).

1) RViz: RViz는 ROS를 위한 사실상의 시각화 제품군으로, 포인트 클라우드, 로봇 모델, 이미지 등 다양한 센서 스트림을 시각화할 수 있습니다. 또한, 포즈 목표와 같은 명령을 시스템을 통해 직접 전송할 수 있도록 지원하는 도구도 포함되어 있습니다. Voxblox와 같은 재구성 방법을 통해 사용자 정의 뷰어를 구현할 수 있도록 타사 플러그인을 지원합니다. 이를 통해 새로운 환경 데이터를 기존 로봇 및 원격 조작 스택과 완벽하게 통합할 수 있습니다.

Radiance Field 플러그인: 온라인 Radiance Field 생성 기능을 ROS와 통합하기 위해 Radiance Field RViz 플러그인이 제공됩니다. 이 플러그인은 RViz OpenGL 장면의 사용자 지정 카메라 역할을 하며, 카메라가 움직일 때마다(동적 모드) Radiance Field 노드로 렌더링 요청을 전송하거나, 움직이는 프레임에 기반으로 렌더링 스트림을 전송합니다(연속 모드). 동적 모드는 탐색을 위한 오버헤드 뷰나 검사를 위한 클로즈업 뷰처럼 정적인 관점에서 장면을 볼 때 가장 적합합니다. 반면 연속 모드는 3인칭 조종과 같이 움직이는 기본 프레임이나 새로운 수신 데이터를 관찰해야 할 때 유용합니다. 두 모드 모두 렌더링 속도는 출력 이미지의 해상도에 따라 크게 달라지며, 출력 이미지의 해상도는 뷰 패널의 설정을 통해 조정할 수 있습니다. 동적 모드에서 빠른 데이터를 제공하기 위해 렌더링은 최종 해상도의 10%로 먼저 전송된 후, 전체 이미지를 전송하기 전에 50%로 전송됩니다. 이를 통해 사용자는 렌더링된 환경에서 더욱 빠르게 반응하며 장면을 탐색할 수 있습니다.

RViz와 Radiance Field의 깊이 데이터는 렌더링을 현실적으로 가리는 것을 표시하는 데 사용됩니다.

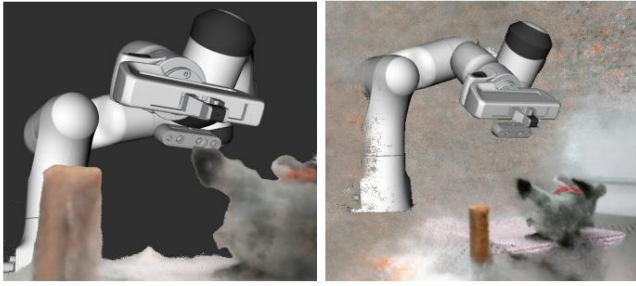


그림 5. 축 정렬 경계 상자를 사용하면 배경을 잘라내거나(왼쪽) 벽을 제거하여(오른쪽) 새로운 보기가 더 명확한 조정이 가능합니다.

그림 4와 같이 장면 요소 또는 컷아웃으로 표현됩니다. 이는 렌더링에서 캡처된 실제 깊이를 0에서 1까지의 비선형 스케일인 OpenGL의 z-값으로 변환하여 구현됩니다. 이렇게 변환된 깊이는 특정 카메라 시점에 대해 RViz의 Ogre 엔진(역시 OpenGL 기반)에서 사용하는 z-값과 일치될 수 있습니다. 결과적으로 로봇, 센서 시각화, 그리드, 포즈 등 RViz 내의 모든 객체가 정확하게 그려져 더욱 사실적인 시각화를 제공합니다. 이는 사용자가 환경 내 깊이를 더 잘 이해하는 동시에 장면에서 RViz 구성 요소를 더 쉽게 찾을 수 있도록 하는 것을 목표로 합니다.

대규모 실내 장면에서는 렌더링된 이미지가 항상 RViz 요소 뒤에 표시되어 쉽게 찾을 수 있는 컷아웃 모드도 있습니다.

또한, 3DGS와 NeRF 모두 체적 렌더링을 수행하기 때문에 폐색이나 미탐색 노이즈로 인해 자체 위치 파악이 어려울 수 있습니다. 이 문제를 해결하기 위해 이 플러그인은 광선 적분을 제한하여 장면을 자르는 축 정렬 바운딩 박스를 생성할 수 있도록 합니다. 이는 그림 5에서 확인할 수 있으며, 바운딩 박스는 노이즈가 많은 배경과 그림 8에서 볼 수 있듯이 조작자의 시야를 가리는 뒷벽을 제거하는 데 사용됩니다. 이 두 렌더링은 로봇이 벽이나 팔의 이동성 제한으로 인해 포착할 수 없었던 완전히 새로운 장면입니다.



그림 6. 데이터에 Voxblox 메시가 있고 왼쪽(왼쪽)에 Radiance Field Viewer가 있는 VR 시스템을 사용하는 원격 조작자, 메시 재구성에서 실시간 LiDAR 데이터가 중첩되어 있음.

2) VR: 테스트 결과, 오클루전이 있더라도 작업자가 2D 화면을 볼 때 깊이감을 느끼기 어렵다는 점이 확인되었습니다. 이 문제를 해결하기 위해, 로봇 데이터를 표시하기 위해 Unity 씬을 실행하는 Meta Quest 3 VR 헤드셋으로 시스템을 이식했습니다. 이는 VR을 사용하여 기존 원격 조작 시스템의 기능을 확장하여 더 큰 효과를 얻기 위해 사용된 이전 연구 [1], [17], [8]의 후속 연구였습니다.

몰입감과 사용 편의성.

VR 로봇 데이터: 비교 기준선을 마련하기 위해 TF 동기화 로봇 모델, 포즈 게시, 센서 시각화(LiDAR, RGBD 포인트 클라우드, 이미지용) 등 RViz의 몇 가지 기능을 그림 6과 같이 VR 헤드셋에 통합했습니다. 사용자는 헤드셋과 손 상호작용을 통해 포즈 목표를 전송하여 VR 인터페이스에서 직접 로봇 명령을 전송할 수 있습니다. 이 시스템은 Voxblox 재구성의 실시간 스트리밍을 위해 모바일 GPU 기반 메시 생성 기능을 통해 확장되었습니다. 이를 통해 사용자는 RViz에서 접근 가능한 것과 동일한 센서 정보를 몰입형 3D 환경에 표시할 수 있습니다. 경량화를 위해 VR 장면은 헤드셋에서 기본적으로 실행되도록 최적화되었으며, ROS와의 TCP 연결만으로 메시지 트래픽을 처리합니다. 이를 통해 헤드셋은 가자국 PC나 헤드셋 테더링 없이 로봇의 ROS 네트워크에 직접 연결하여 상호작용을 시작할 수 있습니다.



그림 7. 2.5D 핸드헬드 패널렉스 뷰어(왼쪽)와 완전 몰입형 360° (오른쪽)를 사용한 Radiance Field VR 장면.

VR Radiance Fields: Radiance Field 렌더링을 사용하기 위해 2.5D 핸드헬드 뷰어가 제공되며, 이 뷰어는 렌더링된 세계를 사용자에게 보여줍니다. 이 뷰어는 깊이 데이터를 사용하여 시차 효과를 생성하여 VR 헤드셋이 렌더링을 보는 동안 더욱 풍부한 깊이감을 제공합니다. RViz 카메라와 유사하게, 뷰어는 사용자가 머리를 움직일 때마다 새로운 렌더링을 요청할 수 있으며, 이는 장면 내 로봇 데이터의 위치와 동기화됩니다. ActionServer에 대한 요청은 TCP 연결을 통해 직접 라우팅되며, 렌더링된 결과는 시차 렌더링을 위해 GPU 셰이더로 직접 스트리밍됩니다. 더욱 몰입감 있는 경험을 위해 360° 구형 렌더링도 제공되어 사용자가 마치 환경에 실제로 존재하는 듯한 스케일감을 느낄 수 있습니다. 뷰어는 이미지 확대/축소, 해상도, 로봇의 라이브 카메라 피드와 같은 제어 기능도 제공합니다. 그림 7은 배치된 핸드헬드 Radiance Field 뷰어와 구형 렌더링 모드에서의 뷰를 보여줍니다. RViz 플러그인과 마찬가지로 Radiance Fields의 깊이 데이터는 동적 장면 폐색을 제공하는 데에도 사용되어 로봇이나 포인트 클라우드와 같은 더 가까운 3D 객체를 Radiance Field 렌더링 앞에 렌더링할 수 있습니다.

IV. 실험

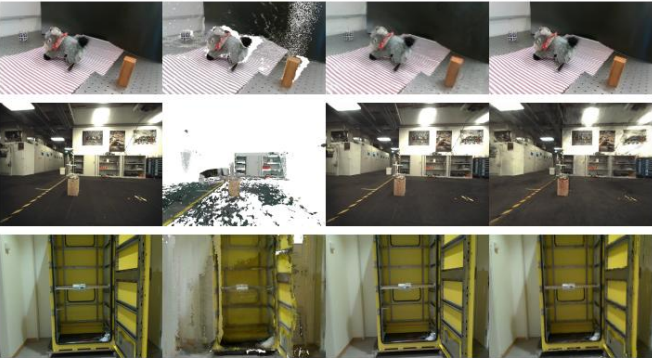
이 파이프라인에는 세 가지 주요 구성 요소가 있으므로 세 가지 실험 세트가 수행되었습니다. 먼저, 로봇에 독립적인 작동을 보장하고 재구성 결과를 비교하기 위해 고정 팔, 모바일 베이스, 모바일 팔에서 데이터를 수집했습니다. 각 데이터 세트는 Voxblox, NeRF, 3DGS를 사용하여 처리되었습니다.

광도 품질을 평가했습니다. 둘째, 온라인 작업을 테스트하기 위해 재구성 및 렌더링 시간을 측정했습니다. 마지막으로 시청 경험을 검증하기 위해 사용자 연구가 수행되었습니다. RViz 메시와 NeRF 렌더를 비교하기 위해 수행된 VR 대응 제품.

A. 데이터 세트

- 1) 정적 팔: 정적 Franka Panda 팔이 사용되었습니다. Intel Realsense 435i RGBD 카메라. 암은 고정되어 있습니다. 금속 테이블과 두 면이 벽으로 둘러싸여 있어 대상 스캐닝 영역을 20cm x 20cm로 설정합니다. 세 번째 면 이 지역은 팔이 있는 곳이기 때문에 탐험할 수 없습니다. 장착, 스캔 범위를 전면으로 제한 방위각은 90 ° 에 불과했습니다. 세 개의 표적이 포착되었습니다. 고운 털과 기준 큐브, 나무로 만든 당나귀 장난감 블록. 포피는 고무파 디테일을 테스트하는 데 사용되었습니다. 기준점에서 텍스처 정확도를 확인했고 블록이 3D 기하학적 재구성을 확인하는 데 사용됩니다.
- 2) 모바일 베이스: Anybotics Anymal은 걷기에 사용되었습니다. 대규모 실험실 환경을 중심으로. 장면의 규모는 약 15m x 10m 크기의 이 물체는 전면 및 후면 1440p RGB 카메라로 촬영되었습니다. 위치 측정은 CompSLAM[9]을 사용하여 수행되었으며, 나무 받침대를 중심으로 진행되었습니다. 로봇이 궤도를 돌았습니다. 받침대 위에는 유리 그릇이 있었습니다. 각 시스템의 투명하고 반사적인 재구성 테스트 능력.
- 3) 모바일 암: DynaArm은 Intel Realsense L515를 추가로 탑재한 Anybotics Anymal에 부착되었습니다. 손목에 장착되었습니다. 포피 대상은 노란색이었습니다. 중앙에 금속 막대가 있는 스위치보드 캐비닛 화면과 스위치가 달린 상자를 들고 있었습니다. 대략 1m x 1m 크기로 Anymal 스캐닝을 통해 전면에서 캡처 팔을 2m 떨어진 곳에서 움직여서 캡처하기 전에 상자 안에. 팔이 상자 안으로 들어가면서 조명 조건을 약간 변경하여 얼마나 잘 테스트하는지 시스템은 동적 색상에 적응합니다.

나. 품질



지상 진실 Voxblox 네르프 3DGS

그림 8. 고정 암(왼쪽), 이동식 베이스(가운데) 및 모바일 암(오른쪽). 맨 위 행은 지상 실측 이미지를 보여주고 두 번째 행은 행은 Voxblox 재구성을 보여주고 세 번째 행은 NeRF를 보여줍니다. 재구성을 보여주고, 아래 줄은 3DGS 재구성을 보여줍니다. NeRF 모바일 데이터 세트 모두에 대해 최고 품질의 결과를 생성했습니다. 3DGS는 정적 팔 데이터 세트에 대해 거의 완벽한 이미지를 만들었습니다.

표 I
다양한 로봇 구성의 광도 측정 비교
및 재건 방법

데이터 세트 방법 PSNR ↑ SSIM ↑ LPIPS ↓				
정적 암	복스블록스	15.42	0.4969	0.5507
	네르프	18.07	0.5309	0.4917
	3DGS	34.12	0.9288	0.2298
모바일 베이스	복스블록스	6.85	0.4091	0.8051
	네르프	25.55	0.8718	0.1988
	3DGS	20.18	0.7853	0.3953
모바일 암	복스블록스	16.94	0.514	0.5295
	네르프	22.16	0.6473	0.1478
	3DGS	24.59	0.7414	0.2126

플랫폼이 로봇에 독립적이라도 하려면 다음 세 가지가 필요합니다. 다양한 배포가 실행되었습니다. 세 가지 설정 모두 훈련되었습니다. Voxblox, Nerfacto를 사용하여 동일한 매개변수를 사용하여 Splatfacto, Nerfacto 및 Splatfacto가 기본입니다. NerfStudio의 NeRF 및 3DGS 방식입니다. 각각 피크 신호 대 잡음비(PSNR)를 기준으로 평가된 구조적 유사성 지수 측정(SSIM) 및 학습 지각 이미지 패치 유사성(LPIPS)[18]. PSNR은 SSIM은 장면의 아티팩트를 측정하는 데 사용됩니다. 조명 및 대비와 같은 기능의 유사성 LPIPS는 패치의 네트워크 활성화를 측정하여 사람들이 유사하다고 말할 수 있는 수준을 추정합니다. 실제 결과는 다음과 같습니다. 그림, 기준 Voxblox 재구성, NeRF 및 3DGS 그림 8에 표시되어 있으며 광도 비교는 다음과 같습니다. 표 I에 표시됨. 각 데이터 세트에 대한 최상의 결과 방법 굵은 글씨로 표시되어 있습니다.

- 1) 정적 팔: 팬더 팔 데이터는 거의 완벽했습니다. 포즈를 취하면 매우 높은 질감의 결과를 생성할 수 있었습니다. 충실도. 그림 8에서 기준 큐브의 마커는 스캔하기에 충분히 선명합니다. 이동성이 부족합니다. 장면 주변에 제한된 시야를 제공했습니다. 결과적으로 NeRF 재구성의 품질이 낮아집니다. 그러나 작은 장면 크기는 놀라울 정도로 고품질의 3DGS를 의미합니다. 재구성. 이는 가우시안이 전체 공간을 밀접하게 덮고 있어 장면이 잘 초기화되었기 때문입니다. 약간의 과잉. 모든 지표(PSNR, SSIM, LPIPS)에 걸쳐 3DGS 재구성은 두 배 더 잘 수행되었습니다. NeRF. 캡처된 영역은 적절하게 표현되었지만, Voxblox는 불완전한 메시지를 생성하여 가장 낮은 점수를 받았습니다. 모든 지표. 이러한 결과는 표 I에 나와 있습니다.
- 2) 모바일 베이스: Radiance Fields, 특히 NeRFs, 장면 중심에서 멀리 떨어진 광선 안정성에 문제가 발생합니다. 소음 속에서. Voxblox조차도 먼 거리의 데이터를 포착하는 데 어려움을 겪었습니다. 대부분 불완전한 재구성을 생성합니다. 훈련 중에 NeRF 방법은 다음을 통해 노이즈를 부드럽게 할 수 있었습니다. 포즈 최적화를 통해 최상의 결과를 얻을 수 있습니다. 또한, Voxblox에서 대상 유리 그릇이 완전히 없어졌습니다. 재구성은 NeRF에서 명확하게 포착되었지만 3DGS로 대략적으로 캡처했습니다.
- 3) 모바일 암: 스위치보드 캐비닛이 생성되었습니다. NeRF 및 3DGS에서 LPIPS에 대한 최고 품질의 결과 재구성은 인간이 인식하는 가장 높은 수준을 나타냅니다.

표 II

다양한 방법 간의 시간 비교

방법	반복당 시간[ms]	16.94dB[s] 의 PSNR	세우다 시간[ms]	세우다 시간[FPS]
복소블록스	1205.903	165.470	-	-
네르프	35.644	7.027	1020.131	0.980
3DGS	34.651	6.996	6.6257	151

품질. Voxblox는 대부분의 장면을 재현했지만, 지지대와 문의 일부를 포착하는 데 실패했습니다. 또한, 캐비닛 내부를 이동하면서 조명이 변함에 따라 메시의 색상이 균일하지 않았습니다. 이는 Radiance Fields의 시점 의존적 색상을 통해 포착되었습니다. 표 I에서 볼 수 있듯이, 3DGS는 PSNR과 SSIM에서 NeRF보다 우수한 성능을 보였는데, 이는 구면 고조파와의 색상 매칭이 더 우수했기 때문일 가능성이 높습니다.

C. 성능 온라인 로봇 원격

조작에는 빠른 재구성 및 시각화가 필요합니다. 이를 위해 RTX 4090에서 각 방법의 재구성 및 렌더링 시간을 비교했습니다. 실행 및 비교 전반에 걸친 반복성을 보장하기 위해 데이터를 먼저 rosbags에 기록하고 이미지의 시퀀스 ID를 파일에 저장했습니다. 이후 이 ID를 사용하여 실행 중에 데이터셋을 생성하고, 각 실행에서 동일한 타임스탬프의 동일한 이미지가 사용되도록 했습니다. 먼저, 반복당 시간을 측정했습니다. Voxblox의 경우, 이는 새로운 배치를 메시와 통합하는 데 걸리는 시간이며, Radiance Fields의 경우, 이는 한 번의 반복에 대한 학습 시간입니다. Voxblox는 1cm 복셀로 실행되었으며, 각 배치를 통합하는 데 1.2초가 걸렸습니다.

NeRF와 3DGS의 반복 시간은 각각 35.644ms와 34.651ms였습니다. 또한, 공정한 비교를 위해, Radiance 방법은 최종 Voxblox PSNR 16.94dB에 도달할 때까지 학습되었습니다. NeRF의 경우 7.027초, 3DGS의 경우 6.996초가 소요되었습니다. Radiance Field 방법은 더 높은 품질의 결과를 생성했을 뿐만 아니라, 거의 20배 더 빠르게 달성했습니다. 이러한 결과는 표 II에 요약되어 있습니다.

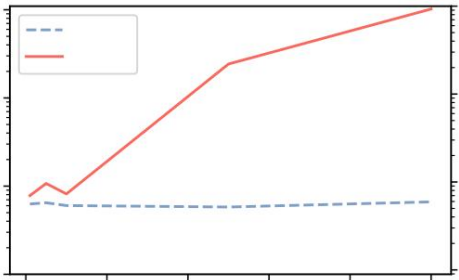


그림 9. 1024x1024 이미지에 대한 NeRF와 3DGS의 다양한 해상도 비율 렌더링 시간 비교. NeRF의 렌더링 시간은 기하급수적으로 증가하는 반면 3DGS의 렌더링 시간은 거의 일정함을 보여줍니다. NeRF 렌더링 시간의 급격한 증가를 고려하여 데이터는 로그 스케일로 표시했습니다.

둘째, 렌더링 시간이 고려되었습니다. Voxblox

지속적으로 업데이트되는 메시지를 표시하며, 시청자는 초당 30프레임으로 렌더링 속도를 제한합니다. NeRF와 3DGS의 경우, 전체 1024x1024 이미지를 렌더링하는 데 NeRF는 1020.13ms(0.98FPS), 3DGS는 6.63ms(151FPS)가 소요되었습니다. 다양한 해상도의 렌더링 시간을 비교해 보면, 3DGS는 약 6.23ms로 거의 일정하게 유지되는 반면, NeRF는 고해상도일수록 렌더링 시간이 기하급수적으로 길어집니다. 이러한 관계는 그림 9에 나와 있습니다. 빠른 렌더링이 필요한 경우, 메시나 3DGS와 같은 명시적인 표현 방식이 더 큰 해상도와 화면에 맞게 확장성이 뛰어나므로 선호됩니다.

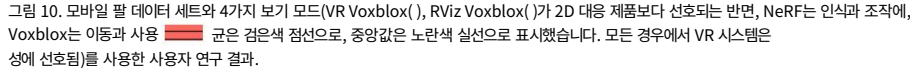
D. 사용자 연구

20명의 참가자를 대상으로 사용자 연구를 수행하여 다양한 시각화 시스템을 비교했습니다. 참가자는 로봇 시스템에 익숙하고 RViz를 정기적으로 사용하는 그룹에서 선택되었습니다. 연령은 22~32세로 평균 연령은 26세였고, 여성 참가자는 5명, 남성 참가자는 15명이었으며, 이전에 VR 시스템을 사용한 사람은 5명뿐이었습니다. 그들은 그림 8의 모바일 팔 데이터 세트를 평가하고 2D RViz Voxblox와 NeRF 장면을 VR 대응물과 비교했으며, 그 결과는 그림 10에 나와 있습니다. 이 데이터 세트는 가장 높은 광도 점수(표 I)와 따라서 가장 높은 품질을 가지고 있기 때문에 선택되었습니다. 혼란을 피하기 위해 Radiance Fields, NeRF 및 3DGS가 모두 동일한 사용자 인터페이스를 통해 제공되므로 사용자에게는 한 가지 방법만 제공되었습니다. NeRF는 이 비교를 위해 선택되었는데, 이는 인간이 인지하는 품질이 가장 근접하는 LPIPS에서 3DGS보다 점수가 높았기 때문입니다[18]. 테스트한 해상도에서는 NeRF와 3DGS의 렌더링 시간 차이가 눈에 띄지 않았습니다.

인지 측면에서는 NeRF가 Voxblox 재구성보다 선호되었으며, VR NeRF가 RViz 변형보다 약간 더 높은 점수를 받았습니다. 이는 표 I의 광도 비교 결과, Radiance Fields가 메시보다 더 높은 품질의 결과를 생성한다는 것을 뒷받침합니다. 흥미롭게도, VR 시스템은 RViz와 동일한 데이터를 보여줌에도 불구하고 인지 품질이 지속적으로 향상되었습니다. 이는 VR 시스템의 몰입감, 즉 깊이감이 기하학적 인지를 돕고 헤드 마운트 광학 장치가 텍스처 품질을 향상시키는 데 도움이 되기 때문입니다. 또한, VR 헤드셋의 광학 장치는 저해상도 이미지를 더 크고 높은 해상도로 인식할 수 있도록 하여 시스템의 렌더링 지연 시간을 줄입니다.

원격 조작 작업 측면에서는 Voxblox VR과 RViz의 3D 메시가 이동 작업에 더 적합했습니다. 사용자가 로봇의 위치를 주변 환경과 비교하여 쉽게 파악할 수 있었기 때문입니다. 그러나 조작 작업에서는 RViz NeRF가 Voxblox보다 더 적합했지만, VR 시스템은 거의 동일했습니다. 이러한 작업에서는 인지 능력과 세부적인 정보를 읽는 능력이 매우 중요하기 때문에 NeRF 시스템이 선호되며, 많은 VR 사용자들이 이상적인 설정으로 3D NeRF를 선호한다고 답했습니다. 향후 3DGS의 명시적 표현을 활용하여 2.5D 뷰어 없이 가우시안 분포를 3D로 직접 표시함으로써 이 문제를 해결할 수 있습니다.

마지막으로, VR 시스템은 RViz 시스템보다 사용성 면에서 더 선호되었습니다. 많은 사용자가 VR을 처음 시도했음에도 불구하고 말입니다. Voxblox 시스템은 사용하기 더 쉬웠습니다.



[3] M. Tancik, PP Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, JT Barron 및 R. Ng, "푸리에 기능을 통해 네트워드는 차자원 도메인에서 고 주파 함수를 학습할 수 있습니다." NeurIPS, 2020.

[5] T. Muller, A. Evans, C. Schied 및 A. Keller, "다중 해상도 해시 인코딩을 사용한 인스턴트 신경 그래픽 기본형", *ACM Trans. Graph.*, vol. 41, no. 4, pp. 102:1–102:15, 2022년 7월.

[7] M. Tancik, E. Weber, E. Ng, R. Li, B. Yi, J. Kerr, T. Wang, A. Kristoffersen, J. Austin, K. Salahi, A. Ahuja, D. McAllister, A. Kanazawa, “Nerfstudio: 신경 복사장 개발을 위한 모듈 식 프레임워크”, ACM SIGGRAPH 2023 컴퍼런스 회의록, ser. SIGGRAPH '23, 2023.

[9] S. Khattak, H. Nguyen, F. Mascarih, T. Dang 및 K. Alexis, "자하 환경에서의 회복성 있는 로봇 자체 추정을 위한 보완적 다중 모드 센서 융합", 2020 무인 항공기 시스템 국제 컨퍼런스 (ICUAS), 2020, pp. 1024–1029.

[11] H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart 및 J. Nieto, "Voxblox: 온보드 MAV 계획을 위한 증분 3차원 유클리드 부호 거리장", 2017 IROS, 2017.

[13] Y. Pan, Y. Kompis, L. Bartolomei, R. Mascaro, C. Stachniss 및 M. Chli, "Voxfield: 온라인 계획 및 3차원 재구성을 위한 비투영 부호 거러장", 2022 IROS, 2022.

[15] T. Chen, O. Shorinwa, W. Zeng, J. Bruno, P. Dames 및 M. Schwa-ger, "Splat-nav: 가우스 splatting 맵에서의 안전한 실시간 로봇 탐색", 2024.

[17] L. Meng, J. Liu, W. Chai, J. Wang 및 MQ-H. Meng, "인간-장면 상호 작용을 통한 가상 현실 기반 로봇 원격 조작", *Proce-dia Computer Science*, 제226권, 141-148쪽, 2023, 생체모방 지능 및 로봇공학 국제 컨퍼런스 회의록.

[18] R. Zhang, P. Isola, AA Efros, E. Shechtman 및 O. Wang.

첫째, 고정 팔부터 여러 대의 카메라를 갖춘 이동식 기지까지 다양한 로봇 배치에 대한 시스템의 적응성을 입증한 결과, 각 설정에서 Radiance Fields는 메시보다 더 높은 품질의 재구성을 생성했습니다. 둘째, 3DGS의 대폭 향상된 렌더링 속도는 온라인 사용 시 메시 렌더링과 유사하면서도 동일한 품질을 달성하는 데 더 짧은 시간만 소요되었습니다. 마지막으로, 시각화 시스템을 비교한 사용자 연구는 사용자가 기존 메시보다 Radiance Fields의 품질을 선호한다는 것을 보여줍니다. 조작 작업의 경우, 이러한 렌더링이 기존 시스템보다 더 선호되었습니다. 전반적으로 VR 시스템은 더 나은 경험을 제공했으며, VR에서 직접 3DGS와 같이 명시적으로 3D Radiance Fields를 적용한 VR 버전이 최적의 원격 조작 설정임을 시사합니다.

전반적으로, 본 연구는 다양한 환경에 적용 가능한 견고한 파이프라인을 제공하고, 고충실도 재구성 기능을 제공하며, 몰입형 VR 환경을 활용하여 향상된 사용자 경험을 제공함으로써 최첨단 로봇 원격 조작 기술의 발전에 기여합니다. 본 연구 결과는 로봇공학과 Radiance Fields 기술의 융합 분야에서 미래 연구에 유망한 방향을 제시합니다.

참고문헌

- [1] P. Stotko, S. Krumpfen, M. Schwarz, C. Lenz, S. Behnke, R. Klein, M. Weinmann, "모바일 로봇을 이용한 물리학적 원격 조작 및 실시간 탐색을 위한 VR 시스템", 2019 IROS. IEEE, 2019년 11월.
- [2] B. Mildenhall, PP Srinivasan, M. Tancik, JT Barron, R. Ramamoorthi 및 R. Ng, "Nerf: 뷰 합성을 위한 신경 광도장으로 장면 표현", ECCV, 2020.