EXTENDING NEURALRECON FOR REAL-TIME 3D RECONSTRUCTION IN DYNAMIC ENVIRONMENTS

Võ Đình Trung[1] *and* **Trương Phúc Trường**[1]

[1] University of Information Technology, Vietnam National University Ho Chi Minh City

What ?

We introduce a framework to extend the capabilities of NeuralRecon for real-time 3D reconstruction in dynamic environments.

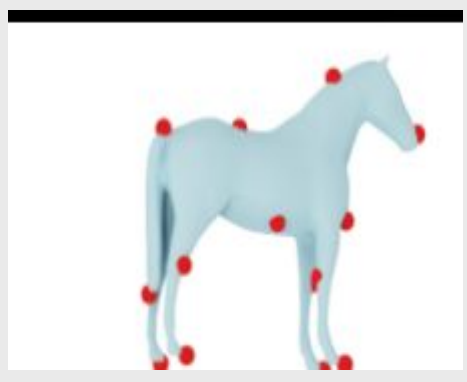
- Proposes a method to integrate object detection and tracking modules into NeuralRecon's pipeline.
- Aims to develop a solution for reconstructing 3D scenes with both static and dynamic objects.
- Evaluate the proposed framework on benchmark datasets with moving objects.

Why ?

- Real-time 3D reconstruction is crucial for AR/VR, robotics, and autonomous systems.
- Existing 3D reconstruction methods struggle with dynamic objects, creating artifacts and inaccuracies.
- Our work addresses a key limitation of NeuralRecon by expanding its applicability to dynamic scenarios.
- The developed system could be used to reconstruct scenes more realistically and reliably.

Overview

Object Detection/Tracking → Dynamic 3D Reconstruction → Static 3D Reconstruction



Description

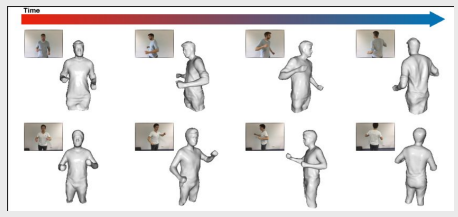
1. Object Detection and Tracking

- Integrate pre-trained object detection model (e.g., Mask R-CNN) to detect moving objects in each frame.
- Apply a tracking model (e.g., Siamese Networks) to establish the trajectory of each detected object.
- Fine-tune these pre-trained models to adapt to the particular datasets and ensure accurate results.

2. Dynamic and Static 3D Reconstruction real time

- Reconstruct separate 3D models for both dynamic (moving objects) and static scene elements using NeuralRecon.
- Explore object-level 3D reconstruction method such as Shape Completion or NeRF for the dynamic objects.
- Develop a seamless fusion mechanism to merge the dynamic object and static scene in a consistent and coherent 3D model.
- Apply blending or feathering to make the transition smooth and natural.

Static 3D Reconstruction real time

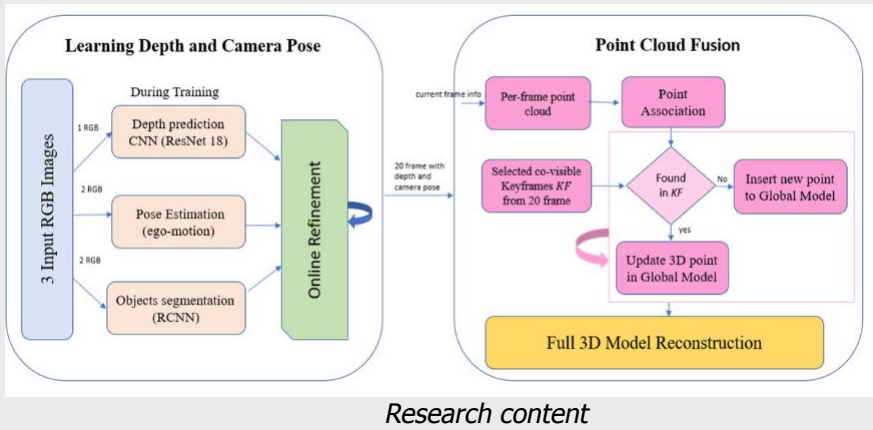


Dynamic 3D Reconstruction real time

3. Dataset:

- In this research, we utilized the following datasets to evaluate our system's performance:

1. **ScanNet**:
   - **Source**: ScanNet is a large-scale dataset containing over 1,500 indoor scenes captured using RGB-D scanning technology.
   - **Annotations**: Includes RGB images, depth maps, camera poses, and reconstructed 3D models.
   - **Size**: Over 1,500 scenes with millions of corresponding images and depth maps.
2. **7-Scenes**:
   - **Source**: 7-Scenes is a smaller dataset focusing on indoor office environments, captured using an RGB-D camera.
   - **Annotations**: Provides RGB-D image sequences, camera poses, and 3D models of the scenes.
   - **Size**: Contains 7 different scenes, each with multiple video sequences.
3. **Replica**:
   - **Source**: Replica is a high-quality dataset containing 3D models of real-world indoor environments.
   - **Annotations**: Offers detailed 3D models with material and lighting information, along with RGB images and depth maps from various camera viewpoints.
   - **Size**: Comprises 18 highly detailed indoor scenes.



Research content

NII

Võ Đình Trung – Trường Đại học Công nghệ Thông tin - ĐHQG TP.HCM
TEL : 0914854406 Email : 22521571@gm.uit.edu.vn