

MỞ RỘNG KHẢ NĂNG TÁI TẠO 3D THỜI GIAN THỰC CỦA NEURALRECON TRONG MÔI TRƯỜNG ĐỘNG

Võ Đình Trung- 22521571

Trương Phúc Trường - 22521587

Tóm tắt

- Lớp: CS519.P11
- Link Github của nhóm: <https://github.com/votrung654/CS519.P11>
- Link YouTube video: <https://youtu.be/6rFDN0k93BI>



Trương Phúc Trường
22521587



Võ Đình Trung
22521571

Giới thiệu

- **Bối cảnh công nghệ và hạn chế:**
 - Công nghệ tái tạo 3D từ video đơn ảnh ngày càng phát triển, ứng dụng trong AR/VR, robot, thiết kế, giám sát.
 - NeuralRecon là phương pháp tiên tiến nhưng chỉ tập trung vào môi trường tĩnh.
- **Vấn đề và thách thức:**
 - Môi trường thực tế thường có đối tượng động như người, xe cộ, gây nhiều khó khăn cho tái tạo 3D.
 - Cần giải pháp mới để xử lý môi trường động một cách chính xác và nhất quán.
- **Mục tiêu tổng quát:** Mở rộng NeuralRecon để tái tạo 3D trong môi trường động.

Giới thiệu

Trong đó:

1. NeuralRecon:

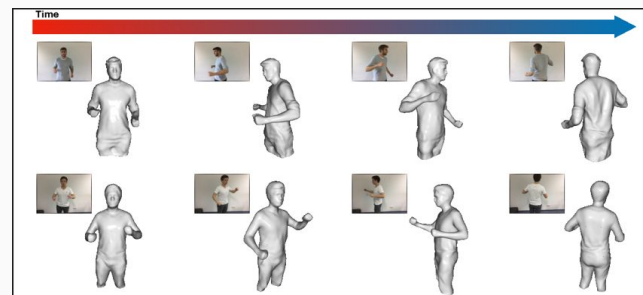
- Mạng học sâu được thiết kế để tái tạo 3D thời gian thực từ video đơn ảnh.
- Sử dụng phương pháp hợp nhất đặc trưng hình học để tạo ra mô hình 3D chính xác.

2. Tái tạo 3D:

- Quá trình xây dựng mô hình 3D từ dữ liệu 2D như ảnh hoặc video.
- Công nghệ hỗ trợ: Neural Radiance Fields (NeRF), Shape Completion.



Tái tạo 3D từ video đơn ảnh ở môi trường tĩnh

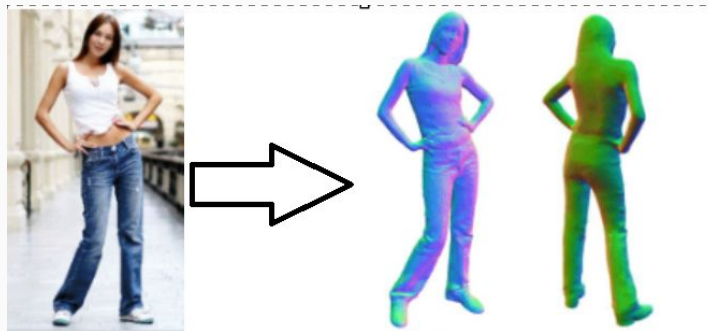


Tái tạo 3D từ video đơn ảnh ở môi trường động

Giới thiệu

Phát biểu bài toán

- **Input:** Video đơn ảnh của một bối cảnh động, chứa cả các đối tượng tĩnh và động (ví dụ: con người, phương tiện giao thông,...).
- **Output:** Mô hình 3D chính xác và nhất quán của toàn bối cảnh, bao gồm cả các phần tĩnh và động.



Ví dụ về kết quả của bài toán tái tạo 3D

Mục tiêu

1. Nghiên cứu tổng quan các phương pháp tái tạo 3D, phát hiện và theo dõi đối tượng.
2. Tích hợp module phát hiện và theo dõi đối tượng động vào NeuralRecon.
3. Phát triển phương pháp tái tạo 3D riêng cho đối tượng động và hòa trộn mượt mà giữa đối tượng động và tĩnh.

Nội dung và Phương pháp

- **Nội dung 1: Nghiên cứu tổng quan**
 - Tìm hiểu các công trình về tái tạo 3D từ video đơn ảnh trong môi trường động.
 - Nghiên cứu các phương pháp object detection và tracking: Mask R-CNN, DETR, Siamese Networks.
 - Phân tích kiến trúc NeuralRecon để xác định điểm cải tiến.
- **Nội dung 2: Thiết kế module tích hợp**
 - Lựa chọn mô hình object detection và tracking phù hợp.
 - Tinh chỉnh mô hình trên dataset chứa cảnh động.
 - Xây dựng giao diện kết nối giữa các module phát hiện, theo dõi và pipeline của NeuralRecon.

Nội dung và Phương pháp

- **Nội dung 3: Tái tạo 3D cho đối tượng động**
 - Nghiên cứu phương pháp shape completion, Neural Radiance Fields (NeRF).
 - Tối ưu phương pháp tái tạo 3D để đạt độ chính xác và hiệu suất thời gian thực.
- **Nội dung 4: Hòa trộn đối tượng động và tĩnh**
 - Nghiên cứu kỹ thuật Poisson blending và feathering để giảm artifacts.
 - Đảm bảo tính nhất quán và hiệu quả thời gian thực.
- **Nội dung 5: Đánh giá và so sánh**
 - Thử nghiệm trên dataset benchmark như KITTI.
 - So sánh với NeuralRecon gốc bằng các tiêu chí: độ chính xác, tính nhất quán, hiệu suất thời gian thực.

Kết quả dự kiến

- Hệ thống NeuralRecon mở rộng, xử lý được môi trường động.
- Module object detection và tracking tích hợp hiệu quả vào NeuralRecon.
- Phương pháp tái tạo 3D riêng biệt cho đối tượng động, hòa trộn mượt mà với môi trường tĩnh.
- Báo cáo chi tiết kết quả thực nghiệm và so sánh hiệu quả phương pháp.

Tài liệu tham khảo

- [1]. Jiaming Sun, Yiming Xie, Linghao Chen, Xiaowei Zhou, Hujun Bao: NeuralRecon: Real-Time Coherent 3D Reconstruction from Monocular Video. CVPR 2021: 15598-15607
- [2]. Luca Bertinetto, Jack Valmadre, João F. Henriques, Andrea Vedaldi, Philip H.S. Torr: Fully-Convolutional Siamese Networks for Object Tracking. ECCV 2016: 850-865
- [3]. Richard A. Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J. Davison, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Andrew Fitzgibbon: KinectFusion: Real-Time 3D Reconstruction and Interaction Using a Moving Depth Camera. ISMAR 2011: 127-136
- [4]. Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, Sergey Zagoruyko: End-to-End Object Detection with Transformers. ECCV 2020: 213-229
- [5]. Christian B. Choy, Danfei Xu, JunYoung Gwak, Kevin Chen, Silvio Savarese: 3D-R2N2: A Unified Approach for Single and Multi-view 3D Object Reconstruction. ECCV 2016: 628-644
- [6]. Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, Ren Ng: NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. ECCV 2020: 151-174
- [7]. Patrick Pérez, Michel Gangnet, Andrew Blake: Poisson Image Editing. SIGGRAPH 2003: 313-322
- [8]. Andreas Geiger, Philip Lenz, Raquel Urtasun: Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite. CVPR 2012: 3354-3361