Towards High-Quality 4D Scene Understanding

Dr. Fisher Yu

Visual Intelligence and Systems group
ETH Zürich

🎙 Host: 董豪 助理教授
🕐 2023年2月17日 星期五 20:00–21:00
📍 在线讲座 (Zoom ID: 857 1619 7635; Password: 966197)

Abstract

Understanding semantics and motion in dynamic 3D scenes from monocular cameras is foundational for autonomous driving and robots. The recent availability of large-scale video datasets creates new research possibilities. In this talk, I will start with our recent findings in multiple object tracking (MOT). We inspected MOT from new perspectives on generalizability and robustness. I will also discuss using the 2D tracking method for monocular 3D object tracking and video instance segmentation. Our quasi-dense 3D tracking pipeline achieves impressive improvements on the nuScenes 3D tracking benchmark and we are still making fast progress. On the video segmentation, we find that our algorithms can even beat the manual labeling accuracy. I hope the talk can encourage more investigation in 4D scene understanding.

Biography

Fisher Yu is an Assistant Professor at ETH Zürich in Switzerland. He obtained his Ph.D. degree from Princeton University and became a postdoctoral researcher at UC Berkeley. He is now leading the Visual Intelligence and Systems (VIS) group at ETH Zürich. His goal is to build perceptual systems capable of performing complex tasks in complex environments. His research is at the junction of machine learning, computer vision and robotics. He currently works on closing the loop between vision and action. His works on image representation learning and large-scale datasets, especially dilated convolutions and the BDD100K dataset, have become essential parts of computer vision research. More info is available at https://www.yf.io