

# Self-supervised Object Motion and Depth Estimation from Video

Libing Zeng

In this project, I will implement this new paper, Self-supervised Object Motion and Depth Estimation from Video. Currently, the authors have not publicly released the code.

The aim of this paper is to solve the motion of individual object from video in the context of autonomous driving (i.e. the video is taken by a camera installed on a moving car). Due to the entanglement of object movement and camera ego-motion, it is challenging to estimate the individual object motion from video.

In this paper, they propose a self-supervised learning framework for estimating the individual object motion and the monocular depth from video by modeling the object motion as a 6 degree-of-freedom rigid-body transformation and leveraging the instance segmentation mask to introduce the information of object. In order to eliminate the scale ambiguity of motion prediction, they impose a novel geometric constraint loss term.

In this final project, I am planning to implement this paper using PyTorch and generate results similar to those present in the paper. The list of goals are as follows.

1. November 15  
the first draft of presentation. Getting familiar with this research topic and understanding the paper.
2. November 25  
the first version of implementation. Generating some results from the implementation
3. December 1  
update the draft of presentation with problems encountered and results generated.
4. December 4  
final presentation, final implementation, and final results.