Proposal: Robust Direct Monocular SLAM in Dynamic Environments

Chenfeng Tu
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213
chenfent@andrew.cmu.edu

Kenan Deng
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213
kenand@andrew.cmu.edu

Zhengjia Huang
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213
zhengjih@andrew.cmu.edu

Jiayuan Li
Biomedical Engineering Department
Carnegie Mellon University
Pittsburgh, PA 15213
jiayuanl@andrew.cmu.edu

1 Introduction

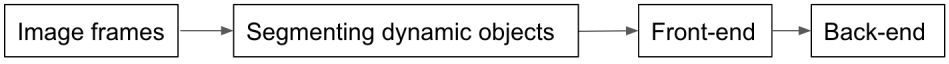
Real-time monocular Simultaneous Localization and Mapping (SLAM) and 3D reconstruction is a popular research topic which has applications in many fields. Classical approach to solve this problem consists of front-end which solves the problem of visual odometry and back-end which optimize the result from front-end. For front-end, classical methods can be divided into feature-based methods and direct methods.

Although modern SLAM systems can be real time and can achieve high accuracy, it is based on the assumption that the scene where the system operates is static. However, it is not true for many environments that SLAM systems work in. Although modern algorithms have some mechanics such as RANSAC that can handle dynamic environments to some extent, when the environment is highly dynamic, these SLAM systems will still fail.

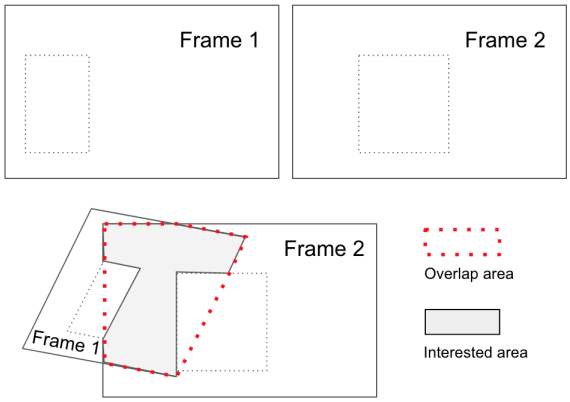
There has been some work in solving this issue in feature-based systems[1][2], which this dynamic problem to some extent. However, there has not been any solution in direct method as far as we know. In our project, we want to solve this issue in the framework of direct method.

2 Method

To address the problem of dynamic environment SLAM in the framework of direct method, the pipeline we propose is shown as below:



The framework of the system should be similar to that of LSD-SLAM. Additionally, we plug in one segmentation layer to filter out the dynamic points in different frames, this will lead to some changes in the minimization of the photometric error. The difference in the error calculating part is shown in the figure below. The dashed rectangle represents dynamic objects we detect. Some parts of the overlapped area will not be counted into the error. Hopefully, the elimination of dynamic points will make the whole system more robust to the dynamic scene.



3 FAQs

- **What are some impacts of the proposed research?**
  The proposed research should improve the stability of SLAM systems based on direct method in dynamic environments.
- **What is novel about the approach you are taking?**
  We use learning based method to eliminate dynamic points in image frames while using direct method for the front-end, which will greatly improve the robustness of the system and it has not been done before to our best knowledge.
- **What methods from class does it use?**
  We will establish a whole SLAM system based on current direct SLAM method such as LSD-SLAM. We will basically go through the whole process of a modern SLAM system from front-end to back-end. Specifically, we will deal with the nonlinear least squares problem, data association problem and we may also learn a lot about rotations and manifolds.
- **What is your metric for success?**
  We will compare camera trajectory in our method to the ground truth trajectory in the TUM dataset. We will also compare our method to our baseline method such as LSD-SLAM.
- **What are the key technical issues you will have to confront?**
  Firstly we have to get familiar with a modern SLAM system based on direct method. Secondly, we have to think out efficient ways to carve out dynamic segments from image frames. Thirdly, we have to deal with the data association problem between different image frames, especially when some parts of the frames are carved out.
- **What software or datasets will you use?**
  We may use code base from LSD-SLAM or DSO as the framework of our system. We will use TUM dataset, which includes some dynamic scenes.
- **What is your timeline?**
  03/18 - 03/31 : Figure out ways to eliminate dynamic segments and get familiar with the whole system.
  04/01 - 04/25 : Establish the system and optimize the results.
  04/25 - 04/30 : Prepare for the presentation.

References

[1] Chao Yu, Zuxin Liu, Xin-Jun Liu, Fugui Xie, Yi Yang, Qi Wei, and Fei Qiao. DS-SLAM: A semantic visual SLAM towards dynamic environments. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2018, Madrid, Spain, October 1-5, 2018*, pages 1168–1174, 2018.

[2] Berta Bescós, José M. Fácil, Javier Civera, and José Neira. Dynaslam: Tracking, mapping, and inpainting in dynamic scenes. *IEEE Robotics and Automation Letters*, 3(4):4076–4083, 2018.