Name: Yilun Xie

Project Advisor: Dr. Zhang, Fumin

Group Name: XXLS

Technical Review for Deep Visual Odometry for 3D Mapping using an Autonomous Indoor Blimp

Introductory:

The project is for technology of indoor aircraft navigation. The goal of this project is to improve the current sensor-based odometry to a visual-based odometry using the most advanced supervised learning tools like deep neural network. This review paper will mainly focus on the software aspect of the project, more specifically, the potential applications of different deep neural network models. Different type of existing technological resources and their potential uses will be explored and compared, and similar application of such technology will also be examined.

Commercial Applications:

The ideal product is expected to generate dense 3D visual maps of an indoor location, useful for navigation, inspection and marketing purpose. There one existing products that shares at least two, if not all design goal of the GT-MAB, is the world's first fully autonomous consumer drone, Skydio R1, developed by a startup called Skydio. It uses AI and a mix of cameras and sensors to pilot itself and follow a subject through complex environments. The product costs about 1999 dollars. Since this drone is operated in an open environment that is a more challenging environment than indoor, the odometry used in this drone is highly comparable to the odometry we about to incorporate, despite the sensor difference on the drone and the GT-MAB. There maybe some constructive design language that we can learn from the review of Skydio R1 by Verge[1].

Supporting Technology:

Since deep visual odometry is a fairly new area of research and just recently became a heated topic, there are numerous papers of related field been published in recently years. Five very import sub-topic in visual

[1] SKYDIO R1 AUTONOMOUS DRONE REVIEW: CRUISE CONTROL. [Online]. Available: https://www.theverge.com/2018/4/5/17195622/skydio-r1-review-autonomous-drone-ai-autopilot-dji-competitor

odometry are feature extraction, feature matching, convolution network motion estimation and local optimisation. And here's some established resources that could potentially be useful to our project. The *Guided Feature Selection for Deep Visual Odometry*[2] offers some perspective on how we extract feature point from obtained image and how we can match these features between images that are taken from a slightly different angle. The *DeepVO: Towards End-to-End Visual Odometry with Deep Recurrent Convolutional Neural Networks*[3] explores some options in how to choose the appropriate recurrent convolutional neural networks for gaining information from the perceived image. For motion estimation, we may looked into the *Exploring Representation Learning With CNNs for Frame-to-Frame Ego-Motion Estimation*[4], which explained how motion update could be performed just by two images that was taken at consecutive seconds. *Convolutional networks for real-time 6-DOF camera relocalization*[5] also provides us with necessary tool to fine tune our odometry using visual relocalization. The lecture notes[6] from CS 4476 Intro to Computer Vision would also be a reliable resource from which we can learn basic image correction and manipulation.

Building Modules:

Although an initial version of blimp vehicle is provided with basic functionality of reaching target locations, there are still possible requirement for additional camera, IMU sensor or proximity sensor. And there's also potential requirement for a more integrated overall design of the whole blimp vehicle to make it more cost efficient, consume less energy and operate safer. A monitor app on the smartphone is also worth considering for easier debugging and more human interactions. The most important aspect of the project, however, is the software: how we process the visual perceived image, how we update our odometry using motion update from displacement of image features and how we relocalize using surrounding environment.

[2] Fei Xue, Qiuyuan Wang, Xin Wang, Wei Dong, Junqiu Wang, Hongbin Zha, "Guided Feature Selection for Deep Visual Odometry", The 14th Asian Conference on Computer Vision 2018

[3] Sen Wang, Ronald Clark, Hongkai Wen, Niki Trigoni, "DeepVO: Towards End-to-End Visual Odometry with Deep Recurrent Convolutional Neural Networks", 2017 IEEE International Conference on Robotics and Automation (ICRA 2017)

[4] Gabriele CostanteMichele ManciniPaolo ValigiThomas A. Ciarfuglia, "Exploring Representation Learning With CNNs for Frame-to-Frame Ego-Motion Estimation", IEEE Robotics And Automation Letters 2016

[5]Alex KendallMatthew Koichi GrimesRoberto Cipolla, "Convolutional networks for real-time 6-DOF camera relocalization", ArXiv2015

[6] Frank Dellaert, CS 4476: Intro to Computer Vision, Georgia Institute of Technology, Fall 2019