Learning Deep Representation for Place Recognition in SLAM

Problem Statement

Learning Deep Representation for Place Recognition in SLAM

Problem Description

Place recognition is one of the most fundamental topics in computer vision and robotics communities, where the task is to accurately and efficiently recognize the location of a given query image.
In navigation, robotic mapping and odometry for virtual reality or augmented reality, simultaneous localization and mapping (SLAM) is the computational problem of constructing or updating a map of an unknown environment while simultaneously keeping track of an agent's location within it.
In the context of this problem, place recognition refers to recognising whether a place has been visited previously or not. For robot navigation with vision, the task of SLAM is an important task. The entire SLAM process relies on recognizing the places the robot has already visited to achieve visual loop closure detection. Closing loops for pose graph optimization, by recognizing previously mapped places is an essential step for performing SLAM.
The basic idea behind the problem is to use the Simultaneous Localization and Mapping (SLAM) for making a map of the local environment. Based on the robot navigation, loop closure needs to be detected so that the robot recognizes whether it has visited a particular place previously or not.

Work Accomplished

At present, we have completed the Monocular Visual Odometry which means tracking the navigation path (trajectory) on a graph. We are doing it using the

KITTI Visual Odometry Dataset which is publicly available for download. In the dataset, the sequence of images are available. The image sequences are related to the navigation on roads. Also, the ground truth poses are available in a TXT file along with the dataset. Currently, we are able to traverse the sequence of images and display them as a video and simultaneously track the path (trajectory) followed in navigation using the ground truth poses. The trajectory followed in one of the image sequences is shown in Figure 1.

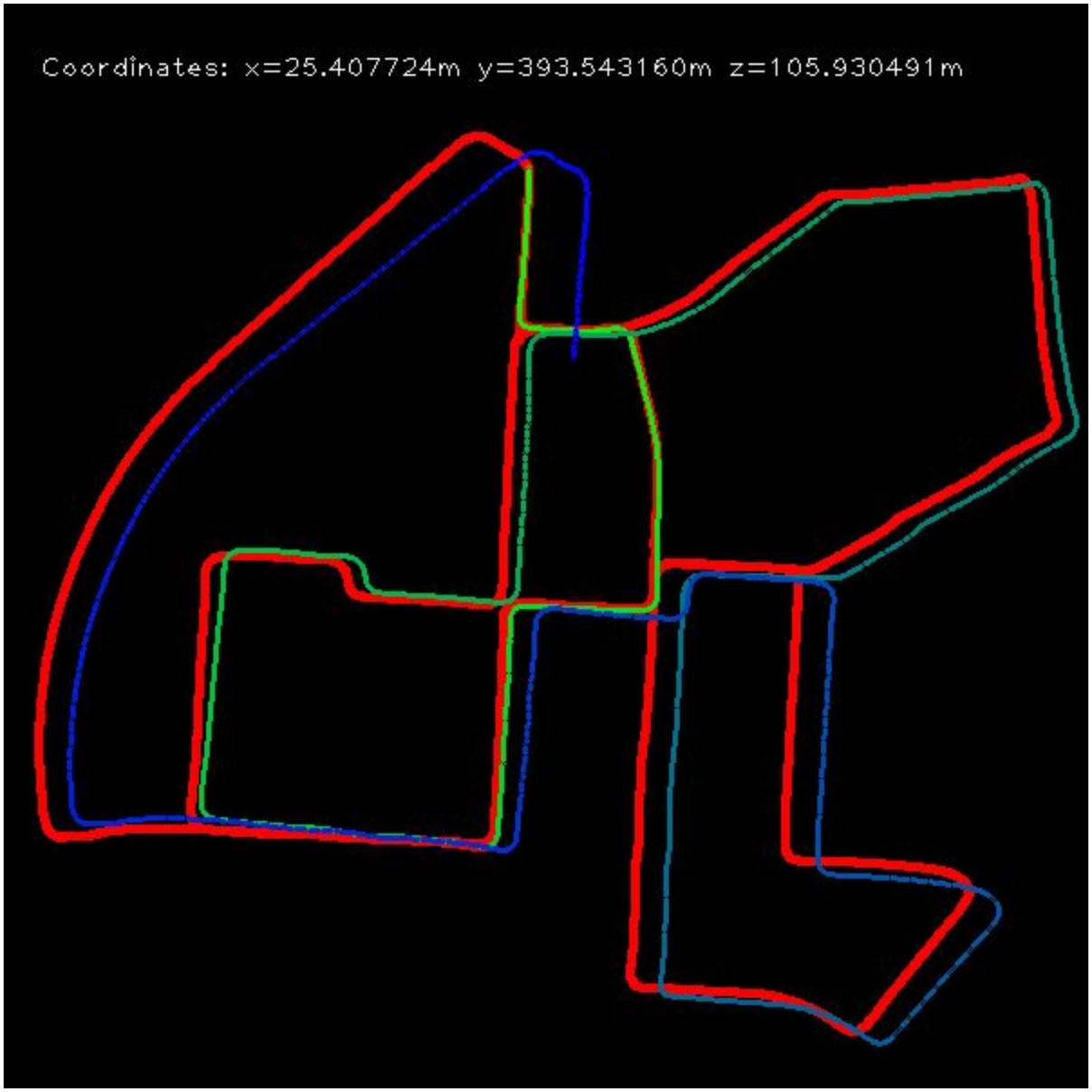

Figure 1: Trajectory tracked from an image sequence

Parallely, we are working on developing a Deconvolution Net (which we are not sure currently if it can be achieved or not). On developing the Deconvolution Net, it can be used to implement the project to detect the trajectory (navigation path) of the robot in real-time using camera.

Contribution & Future Role/Responsibility of Team Members

In the time span after Mid Sem 1, the contribution of the team members in the work accomplished till now was the following:

- Ajinkya Bedekar: Processing of image frames available in the dataset
- Devansh Anhal: Feature Tracking
- Dhruva Agarwal: Updating the graph
- Harsha Deuri: Preprocessing of images

For the future role/responsibility in completing the project, the work left has been divided amongst the team members which is mentioned below:

- Ajinkya Bedekar: Feature Extraction of a Place that the Robot is Visiting
- Devansh Anhal: Loop Closure Detection
- Dhruva Agarwal: Developing the Deconvolution Net
- Harsha Deuri: Tracking the Real Time Trajectory (Navigation Path)

Future Perspectives

For the future perspectives, we will be doing the loop closure detection to recognise whether a place has been visited previously or not. One of the proposed solutions to achieve this is to keep track of the coordinates that are being plotted on the graph. If a particular coordinate is plotted again, the robot can give the output that it has visited that place previously. Another proposed solution is to extract the features of the place that the robot is visiting and store them in the memory. If the features of a particular place matches the features that are stored in the memory, then the robot can conclude that it has visited that place earlier.

Also, it is not always possible to have the ground truth poses beforehand. So, we will also try to track the path of navigation in real time without the availability of ground truth poses (but we are not sure right now whether it can be achieved or not). On achieving this, the robot can track its trajectory (navigation path) in real-time and can plot it on a graph simultaneously.

References

1. Chen, Z., Lam, O., Jacobson, A., Milford, M.: Convolutional neural network-based place recognition. arXiv preprint arXiv:1411.1509 (2014)
2. Noh, H., Hong, S., Han, B.: Learning deconvolution network for semantic segmentation. In: ICCV, pp. 1520–1528 (2015)
3. Milford, M.J., Wyeth, G.F.: SeqSLAM: visual route-based navigation for sunny summer days and stormy winter nights. In: ICRA, pp. 1643–1649. IEEE (2012)