Combining Structure from Motion with visual SLAM for the Katwijk Beach dataset

Marlon B. de Jong[ID]
Thesis supervisor: Arnoud Visser[ID]

Intelligent Robotics Lab, University of Amsterdam, The Netherlands

1 Introduction

In 2015 the Katwijk beach was used to imitate a Martian landscape to allow for a planetary rover to collect an abundance of data from multiple sensors [4]. This dataset is available for public use, with the intention of applying and testing Simultaneous Localization And Mapping (SLAM) techniques. For this thesis the focus was on the data collected from the stereo camera mounted on the rover. The research goal was to create an accurate 3D map with a correct location description. To reach this goal three SLAM techniques were applied to the stereo camera data. These techniques were: visual SLAM using Point clouds, Structure from Motion, and a combination of both.

2 Visual SLAM using Point Clouds

Visual SLAM was performed by first rectifying the left and right images of the stereo camera, followed by combining both images into a disparity map. The result is a point cloud, with for each point a distance estimate and a color. The point clouds are down-sampled into a 3D grid, with an average color for each cube. Following, the cubes are matched with the Iterative Closest Point algorithm [2]. By combining all point clouds one can create a 3D map and estimate the trajectory of the rover.

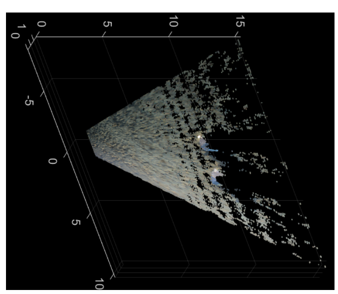

Fig. 1: Example of merged point clouds, depicting two stones in 3D.

The Visual SLAM maps were accurate when applied locally, but when applied on longer trajectories it is difficult to maintain a correct location estimation, because a majority of the points in the cloud display patterns which can be found on multiple locations.

3 Structure from Motion

Structure from Motion works differently in that it uses a limited number of key points in the landscape to map the environment, rather than the uniform distributed point clouds down-sampled in a grid. In this thesis Speeded-Up Robust Features (SURF) was chosen to identify the key points, because it is invariant to scale and photometric variations [1].

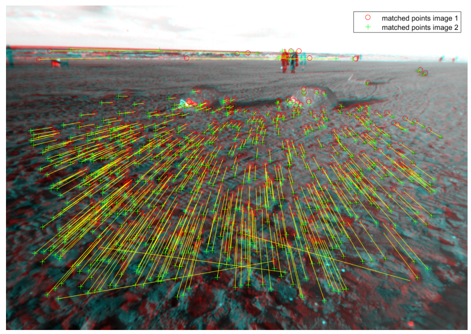

Fig. 2: Visualisation of the matching key points across two images.

The result is a trajectory estimate more accurate than the visual SLAM estimate, however the 3D map was visually incomprehensible from a human standpoint. This is because the map consisted of colourless points. Furthermore are the features close by more prominently represented, the stones were in most instances not close enough to register, as can be seen in Figure 2.

4 Combination

Both techniques have aspects which could be improved. The camera position estimate of visual SLAM is vulnerable to accumulation errors. Structure from Motion suffers from difficult map comprehension. A possible solution to both these problems would be to use the map representation from visual SLAM and the camera position estimate from Structure from Motion, to get the best of both approaches. The results of this combination can be found in the thesis [3].

5 Conclusion

Using the point clouds from visual SLAM and the location estimations from Structure from Motion, this thesis was able to create a visually more comprehensible map with a more accurate location estimation of the rover. The results indicate that by combining the techniques a better performance on both mapping and localisation can be achieved, therefore one would encourage further testing on the combination of these techniques.

References

1. Bay, H., Tuytelaars, T., Van Gool, L.: Surf: Speeded up robust features. In: European conference on computer vision. p. 407. Springer (2006)
2. Chen, Y., Medioni, G.: Object modelling by registration of multiple range images. Image and Vision Computing **10**(3), 145–155 (1992)
3. de Jong, M.B.: SLAM at the Katwijk beach. Bachelor thesis, Universiteit van Amsterdam (July 2020)
4. Hewitt, R.A., Boukas, E., Azkarate, M., Pagnamenta, M., Marshall, J.A., Gasteratos, A., Visentin, G.: The Katwijk beach planetary rover dataset. The International Journal of Robotics Research **37**(1), 3–12 (2018)