Hybrid Camera Pose Estimation

Wenjie Niu

July 2rd,2018

Abstract

This paper aims to solve the pose estimation problem of calibrated pinhole and generalized cameras w.r.t. a Structure-from-Motion (SfM) model by leveraging both 2D-3D correspondences as well as 2D-2D correspondences. While traditional approaches either focus on the use of 2D-3D matches, known as structure-based pose estimation or solely on 2D-2D matches (structure-less pose estimation).

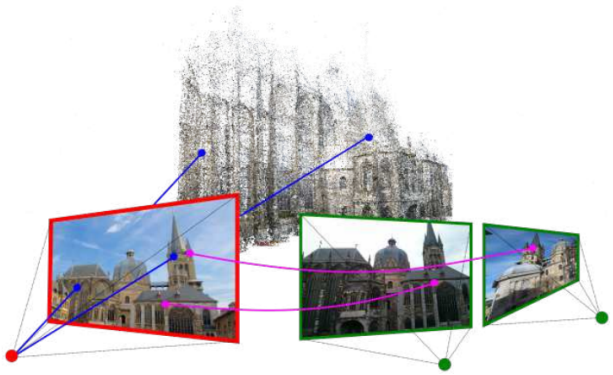

Figure 1. Visualization of 2D-2D matches (pink) and 2D-3D matches (blue) used by one of our hybrid pose solvers. The query camera is represented in red and SfM cameras in green.[1]

1. Introduction

Camera pose estimation, *i.e.*, estimating the position and orientation of a given image, is a central step in 3D computer vision approaches such as SfM [6], Simultaneous Localization and Mapping (SLAM) [3], and visual localization [2]. In addition, camera pose estimation plays an important role in applications such as selfdriving cars [4] and augmented reality [5].

The availability of both structure-based and structureless camera pose estimation techniques leads to a set of interesting questions: Are they mutually exclusive, *i.e.*, is one always preferable over the other, or is there value in using both 2D-3D and 2D-2D matches for pose estimation? Is it best to use pure solvers, *i.e.*, solvers that use either 2D-3D or 2D-2D correspondences, or do hybrid solvers (*c.f.* Fig. 1) using both type of matches improve pose estimation performance? Should one decide prior to RANSAC which solver to use, or is it best to select solvers in a data-driven way during RANSAC-based pose estimation?

2. Hybrid RANSAC for Pose Estimation

RANSAC variant stops when at least one solver s has been chosen K_s times, as this means that a good solution for the current inlier ratios has been found with probability P.

3. Conclusions

This paper have posed the question whether camera pose estimation can be improved by using both 2D-2D and 2D-3D matches. To answer this, we have developed a novel framework for camera pose estimation that jointly uses different minimal solvers within a new Hybrid RANSAC scheme.

Acknowledgements. This research was funded by Google.

References

[1] F. Camposeco, A. Cohen, M. Pollefeys, and T. Sattler. Hybrid camera pose estimation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2018. 1

[2] F. Camposeco, T. Sattler, A. Cohen, A. Geiger, and M. Pollefeys. Toroidal constraints for two-point localization under high outlier ratios. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017. 1

[3] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse. Monoslam: Real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):1052–1067, 2007. 1

[4] C. Hne, L. Heng, G. H. Lee, F. Fraundorfer, P. Furgale, T. Sattler, and M. Pollefeys. 3D visual perception for self-driving cars using a multi-camera system: Calibration, mapping, localization, and obstacle detection. *Image and Vision Computing*, 68:14 – 27, 2017. 1

[5] S. Middelberg, T. Sattler, O. Untzelmann, and L. Kobbelt. Scalable 6-dof localization on mobile devices. In *European Conference on Computer Vision*, 2014. 1

[6] A. Sameer, S. Noah, S. Ian, M. S. Steven, and R. Szeliski. Building rome in a day. In *IEEE International Conference on Computer Vision*, 2009. 1