Jan Čech
Center for Machine Perception
Department of Cybernetics, Faculty of Electrical Engineering
Czech Technical University in Prague

Bachelor's Thesis Review

Prague, May 28, 2021

Title: Calibration of multiple cameras for autonomous driving
Author: Martin Jaroš
Supervisor: Ing. Martin Matoušek, Ph.D.
Date received: 24/05/2021

The thesis presents a method for calibrating a rig of non-overlapping field of view cameras. After the introduction, Chapter 2 introduces the image formation camera and projection models including the radial distortion. Then an intrinsic camera calibration is presented in Chapter 3. Multiple camera calibration method is presented in Chapter 4. Experimental validation is included in Chapter 3 and Chapter 4. Chapter 5 describes the experimental setup. Finally, Chapter 6 concludes the thesis.

For positive points, I would stress that the calibration of cameras with non-overlapping field of views is an important problem in autonomous driving. The thesis presents the calibration method in a compact way, the principle ideas are comprehensible. The text refers to a related literature. The method is validated in both simulated data and real data captured in the laboratory. Results suggest that the method is well designed and probably implemented correctly. The problem is not trivial and requires a mathematical understanding.

Nevertheless, the thesis suffers from several weaknesses.

1. Lack of reproducibility. I do not believe it is possible to reimplement the method based on the thesis. At the beginning, the author made an effort to be formal, using mathematical equations. However, this effort is lost almost completely later on. There are many vague and unclear points, unexplained symbols. For instance, neither the projection equation nor the optimization criterion are given explicitly. Details on the optimization are missing completely. It is only mentioned that Python scipy.optim was used. It should be detailed which optimization algorithm was used.

2. Experimental setup is not described with enough details. The synthetic data are not described at all. For the real data, cameras and lenses are not specified. Then, e.g., dimensions of the acquisition room, size of the calibration targets are not given.

3. Experiments are not very convincing. For the intrinsic camera calibration, the reprojection error is low (around 1 px), however for all models, a single instance is given. Multiple data should have been tested, or the dataset should have been split to provide an idea of error distribution (variance). For the rig calibration, results on both synthetic and real data are reported. However, the synthetic results are only commented that the method converges to the given parameters, and that when perturbed, the reprojection error is below 1 px. This is not acceptable. Especially for the synthetic data, a thorough quantitative evaluation should have been done. For the real data, a comparison is made with a "different" method, but it is not revealed what the reference method is.

4. Minor justification and discussion on the results. Results are only listed in the tables, beside showing the reprojection errors of the targets in several images, and left without comment. Reprojection error remains high for certain instances. The reasons are not discussed. In conclusion, an optimization is mentioned as a suspect. However, it is unclear why it always works on synthetic data and sometimes fails on real data. It should also be discussed/analyzed, where the source of inacuracy is, e.g., an impact of target detection inacuaracy. Degeneracy and minimal configuration should be discussed.

5. Review of the literature is missing.

6. The title promises multiple camera calibration, but a pair of cameras only was considered. It should have been at least sketched how to extend the work to calibrate more than two camera rigs.

7. The thesis would benefit from proofreading. The text is full of typos and grammatical errors. Mathematical typesetting is not satisfactory. Matrices and vectors are typeset as scalars. Notation is inconsistent, e.g., capital letters for 3D scene points are mixed with 2D image points between Chapters 3 and 4.

In summary, considering the difficulty of the problem and Covid-19 pandemic restrictions that certainly complicated work on the thesis, I suggest assessing the thesis

C – good.

Ing. Jan Čech, Ph.D.

Question for the defense:

1. The result in Fig. 3.2 indicates a systematic error in the calibration. The error vectors are pointing to one side. Can you explain the observation?

2. In the first paragraph of Sec. 4.6, it is stated that there is a (length) ambiguity of translation vector m. And that the ambiguity is resolved by measuring the real distance between the cameras. What is meant by "the real distance between the cameras" exactly? I believe it is the distance between camera projection centers. However, I have no idea how to measure this distance in the real world. There are other distances in the scene that would fix the scale, e.g., the calibration pattern size, which is easy to measure.