

University of Cape Town

EEE4022S - Final Year Project

Spacial and Temporal Calibration of Multi-Sensor Systems

Michael Katsoulis

KTSMIC005

24 June 2020

Plagerism Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. I have used the IEEE convention for citation and referencing. Each contribution to, and quotation in, this final year project report from the work(s) of other people, has been attributed and has been cited and referenced.
3. This project report is my own work.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as their own work or part thereof.

Michael Katsoulis
22 June 2020

Contents

1	Introduction	3
1.1	Research goal	3
1.2	Literature Review	3

1 Introduction

Why we want to calibrate a system In the modern world, the need for autonomous vehicles is growing daily. They would improve the quality of human life, provide advances in science and open new opportunities. One of the problems that is being faced at the moment is that vehicles are of different shapes and sizes, this leads to the sensors on each vehicle being in different locations.

The goal of this project is to investigate the current methods of determining the spacial as well as temporal offsets between sensors on a vehicle automatically. This would allow a decrease in calibration time, ease of manufacture as well as allow vehicles to continue to operate in the case of an unforeseen incident moving a sensor. This would allow for cheaper and more robust vehicles.

Same algorithm on multiple models of vehicle.

Drones which can reconfigure in flight or continue if damaged.

Configurable vehicles which adapt to their modification.

1.1 Research goal

1.2 Literature Review

TODO why it is important to have fully calibrated systems

How do we get each form of error When a system has not been spatially calibrated properly, the readings for each sensor cannot be transformed and fitted to one another accurately. This could be because the sensors have been moved from the location where they first were or it could be that the positions of the sensors are just not known. This results in misalignment of the sensor readings and therefore, the vehicle's world view is distorted and inaccurate.

Temporal errors are caused by a wide variety of factors that influence the time at which the readings are taken from all the sensors. It is common for cameras and lidar sensors to have their own onboard clocks which may not be perfectly synchronized. There may be filters on the readings, different lengths of communication channels and other factors too. The reason it is important to have a temporally calibrated system is so that there is no changes in the data between readings. This way, the vehicle knows that the data that is being used to make decisions reflects what was happening at the same instant in time for all the sensors. There have been attempts to fix the temporal calibration problem through clock synchronization and time stamping sensor readings but this is not a long term solution in case the clocks go out of sync while the vehicle is in the field. It is also an expensive way of solving a problem which could be solved with an algorithm and some code.

From the literature that has been reviewed, there are two main methods for obtaining targets in each sensor to use as references during the calibration and evaluation stages. The first method uses markers that are of known shape, texture and that are placed into what is often a known, simple environment. This is a more hands-on, involved approach but it is often simpler to find methods of calibration. The markers are normally planar with patterns, holes and other features on them to help make them easy to distinguish from the environment. The other method is markerless calibration. In these cases, the goal is to use the information in the natural environment around the vehicle in order to calibrate the sensors. This often involves removing targets which are moving relative to the global frame and trying to isolate those which are static.

Markerless calibration is the newest form and it is what many researchers are trying to find an elegant solution to at the moment.

The most common sensors used in autonomous vehicles are lidar and stereo cameras. This calibration faces two main challenges as stated by TODO ref authors in TODO ref: calibration under arbitrary configurations

- The first is that there is a common field of view constraint. Both the lidar and the camera need to be able to see the same objects.
- The second is that two optimizations need to happen simultaneously in both spacial and temporal calibration. This makes the optimization more challenging.

In order to simplify this challenge, there were two main approaches used in the literature remove these constraints. The first is to remain stationary while capturing a 3D image and comparing that to the lidar point map. This helps because dynamic objects in the field of view can be filtered out of the lidar point map as well as the 360° image if there is more than one image taken. There is not a temporal optimization as the scene is stationary. This meant that the researchers could simply solve the spatial calibration on it's own. By having a 360° lidar and camera image, they also removed the common field of view constraint as both images were showing the same frames. A limitation to this research was that they used a target based calibration with checkerboards which had to be set up in a room surrounding the vehicle. They also did not do any filtering of dynamic objects. This leaves two areas for improvement in future research to bring these techniques out of the academic space into something that has real-world applications.

In an attempt to make lidar - camera calibration easier, TODO ref accurate calibration of lidar camera systems using ordinary boxes developed a way to use ordinary cardboard boxes as the calibration targets. This meant that calibration of the system spacially became a simpler, less intensive task. Boxes of known size were placed in a room in arbitrary configurations. Then the lidar point maps were used to find planes in the environment and from there, isolate the potential cardboard boxes. They also developed a way of predicting where the corners of the boxes were if they were between two points. This became particularly useful when the targets were further from the vehicle and the lidar point map was starting to become spread out with distance. All that was left was to find the boxes in the camera images and map the images from the lidar onto the camera images. They used the dimensions of the boxes to then finish calibrating the system spatially. TODO check if they used temporal calibration. similarly to the paragraph above, the problem with this technique is that it requires supervision and cannot be done on the fly in an unknown environment. But it provides some good ideas for improving on the need for complex planar targets while achieving the same accuracy.

In the fly Camera & LiDAR Calibration TODO cite

The main task is to find an efficient way to map point clouds from lidar sensors to the pixels of camera sensors. In TODO ref the solution to this problem was to have two steps of calibration, one coarse, object-level alignment and the other small, fine detail alignment. To do the coarse alignment, they expanded the point dots into planes and tried to find objects in the environment by matching sharp changes of distance ($\geq 30\text{cm}$ jumps) to lines in the camera image. To implement this, TODO ref obtained a 3D model from the camera images using a Structure from Movement. They then found potential objects and mapped these to the lidar point cloud. A finer alignment was then done using a Non-Uniform Rational B-Spline. Using movement to calibrate the camera-lidar system removes the dependency on high fidelity GPS and inertial measurement units.

TODO then used filters to pick up edges in both the point maps and images. Then they dilated the edges found so that there was a wider margin to find overlaps and then they found the

Another method that has been used is a spin on the hand-eye calibration used when calibrating robot arms. In hand-eye calibration, the spatial transformation from the robot "hand" to the cameral is not known. The robot then undergoes a series of known transformations which we will call A and B . Then at each pose, the position of the hand-eye can be given by the combination of the known transformation and the spatial transformation between the hand and eye (X).

So using the relationship $AX = XB$, we can determine X . In their paper, TODO ref Motion-Based Calibration of Multimodal Sensor Extrinsic and Timing Offset Estimation , TODO authors showed that a similar algorithm could be used to determine the spatial offset between a camera, lidar and navigation sensor. This meant that the authors could get a coarse calibration by estimating the motion of each of the sensors from what it has been recording before then using a similar approach to the hand-eye problem

to determine the coarse calibration parameters. Then they used objects detected in the environment to implement a fine calibration and ensure that the calibrated data lined up.

This method resulted in no need for calibration targets while also simplifying the coarse calibration step. A problem with it however is that some form of navigational sensor is needed. It did however provide a very easy way to use the same technique for different vehicles which the authors demonstrated. They also showed how the length of the data used during calibration played a role in the final accuracy. It is an algorithm that seems to be of very good practical use and from the report also seems to be the easiest to implement.

TODO How these papers influenced my ideas TODO Direction for the project

– CNNs for finding these objects?

References