Motion Model and Filtering Techniques for Scaled Vehicle Localization with Fiducial Marker Detection

Report Author: Kyle Coble

Group Partners: Sharang Kaul, Akanshu Mahajan

Abstract

This projects studies the fusion of control inputs and IMU data for developing a kinematic bicycle motion model for the KTH Smart Mobility Lab Small-Vehicles-for-Autonomy (SVEA) platform. This motion model is filtered with relative pose estimates between a camera and fiducial markers, using both an Extended Kalman Filter and a Particle Filter. The developed motion models and filters are implemented on SVEA vehicles and are tested in the Smart Mobility Lab. Pose estimates from the motion model and filters are compared against ground truth, determined by a motion capture system with sub-millimeter accuracy. The results presented provide the necessary base for development of automated vehicle control technologies on the SVEA platform with perception based on the detection of fiducial markers.

1 Introduction

Detection of fiducial markers using computer vision (CV) is an effective method for estimating the six degree-of-freedom pose of mobile robots in closed environments, including laboratories, warehouses, and test driving tracks. Reputed robotics companies and educational research labs alike utilize fiducial markers as landmarks for robot localization due to the accuracy and minimal hardware required for implementation. For these reasons, the KTH Smart Mobility Lab has begun to use fiducial markers for localization on their Small-Vehicles-for-Autonomy (SVEA) platform. This project builds on their prior work, which enabled relative pose estimates between an RGB webcam and a fiducial marker in the camera's frame of view.

Using this technology, the pose of a mobile robot could be estimated by detecting stationary fiducial markers with known positions. Conversely, the robot's pose can be estimated by placing a single fiducial marker on the mobile robot and detecting the marker with a stationary camera in a known location. Further, the robot's pose can be estimated by a camera on a second, more intelligent, mobile robot with knowledge of its own pose. The detected pose is filtered with control input signals and inertial measurement unit (IMU) data from SVEA vehicles with the Extended Kalman Filter (EKF) and the Particle filter (PF). This is done with the goal of determining the configuration that best maximizes localization accuracy and implementation flexibility, while minimizing hardware requirements on the SVEA vehicles.

The remote localization of a mobile robot from a stationary camera has applications including 'smart intersections' optimizing vehicular and pedestrian traffic. The remote localization of a mobile robot from another mobile robot has applications including a 'follow-me' control algorithm in the manner of Autonomous Vehicle (AV) fleeting. For these reasons, the localization of a remote-controlled SVEA vehicle, both through the lens of a stationary camera and of a camera mounted on a second mobile robot, was pursued in this project.

1.1 Contributions

This subsection briefly introduces the successes of this project with corresponding impacts.

- An accurate motion model has been developed for the SVEA vehicle platform, based on the kinematic bicycle model. The motion of the vehicle can be approximated from control inputs, with or without fusing IMU data. This is useful in the event of infrequent observation updates occurring from long periods of occlusion.
- Localization packages using both EKF and Particle Filter are readily implementable for various sensor configurations and use cases on SVEA vehicles. All filters were developed in a modular structure that enables substituting in various process and observation measurement sources. As an example, filtered localization of a mobile robot observing stationary fiducial markers is readily implementable.
- Relative localization of a mobile robot based on the pose of another mobile robot has been successfully implemented. This has immediate application for use in vehicular fleeting control, if coupled with path planning and following algorithms.

Note: *For all EKF applications, an open source ROS package called 'robot_localization' was used. Proper implementation required extensive troubleshooting of input sensors and tuning of parameters, such as which sensor states to include and their relevant covariances. The coordinate transformation, motion model, and particle filter nodes were developed specifically for this project, based on relevant references described in Sections 2 and 3. These nodes will be added to the Smart Mobility Lab's default package for SVEA vehicles as a basis for further development of AV technologies as discussed in Section 7.*

2 Theoretical Background

This section describes and references the prior research that this project is built upon.

2.1 Kinematic Bicycle Motion Model

The kinematic bicycle model (Fig. 2, Sec. 3.1) was selected over the dynamic bicycle model because it is simpler to implement, requires less computation, and yields more accurate results across a wide range of speeds [Kong et al. 2015]. The kinematic model only requires system identification of two parameters, far fewer than the dynamic model. Additionally, exclusion of linear and angular acceleration terms in the kinematic bicycle model reduces integration, and thus computation, requirements. Finally, the kinematic model was found to be appropriate for modelling low-speed and zero-speed vehicle operation, the typical realms of operation for SVEA vehicles. For these reasons, the recommendation of [Kong et al.] to use the kinematic bicycle model was followed for modelling the motion of the SVEA vehicles.

2.2 Fiducial Detection

Fiducial markers provide a relatively simple to implement and low cost method for landmarked localization of mobile robots [Mutka et al. 2008]. Through geometric analysis of image frames containing fiducial markers, the relative poses between cameras and markers can be calculated. With knowledge of either the marker's or the camera's coordinates in a global or map reference frame, the other's coordinates can be calculated. This project differs from the recommendations given by [Mutka et al.] in that square fiducial markers (Fig. 1) were chosen over circular markers due to the Smart Mobility Lab's prior implementation using square markers with the SVEA platform.

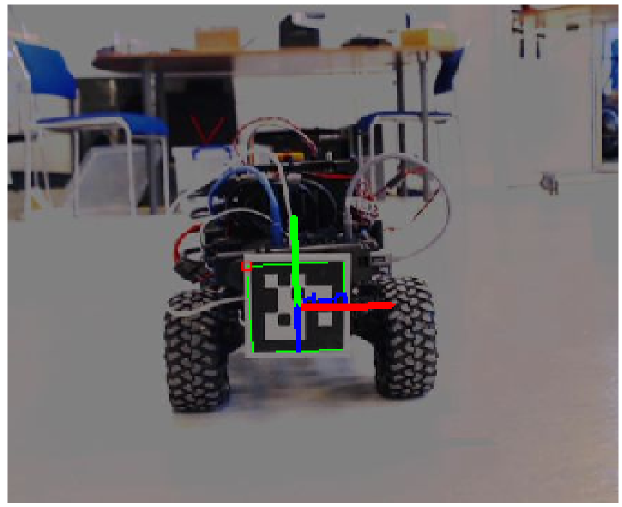


Figure 1: CV Detection of a Square Fiducial Marker mounted on a SVEA

2.3 Filtering Techniques

Popular choices for nonlinear filters used in mobile robot localization include the Extended Kalman Filter (EKF), the Unscented Kalman Filter (UKF), and the Particle Filter (PF) [Konatowski et al. 2016]. The EKF offers a relatively simple and lower computational implementation for filtering nonlinear state estimates given Gaussian noise and an initial pose estimate. The UKF provides more adaptability to high degrees of non-linearity than the EKF does, but has higher computational requirements. The PF is not affected by nonlinear or multi-modal distributions, which makes it an ideal filtering choice in global localization problems. The PF filter was found in the simulations conducted by [Konatowski et al.] to minimize root-mean-square-error as compared to the UKF and EKF. Considering this, it was determined to implement both the EKF, a simpler implementation as a proof of concept, and the PF, for the accuracy and the flexibility to expand to global localization.

3 Methodology

3.1 Kinematic Bicycle Motion Model

As mentioned in Section 2.1, the kinematic bicycle model was chosen to model the motion of the SVEA vehicles based on the findings by [Kong et al.] The equations of motion for a front wheel drive vehicle following the kinematic bicycle model are expressed in Figure 2.

$$
\begin{aligned}
\dot{x} &= v\cos(\psi + \beta) \\
\dot{y} &= v\sin(\psi + \beta) \\
\dot{\psi} &= \frac{v}{l_r}\sin(\beta) \\
\dot{v} &= a \\
\beta &= \tan^{-1}\left(\frac{l_r}{l_f + l_r}\tan(\delta_f)\right)
\end{aligned}
$$

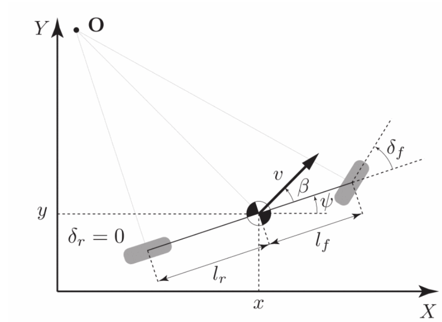


Figure 2: Kinematic Bicycle Model Equations and Diagram [Kong et al. 2015]

Where x and y are the coordinates of the vehicle center of mass, v and a are the speed and acceleration of the vehicle, ψ is the inertial heading angle, and β is the angle of the center of mass velocity relative to the longitudinal axis of the car. l_f and l_r represent the front and rear axle distance from the center of mass. The front wheel steering angle is represented by δ_f, while the rear wheel steering angle, δ_r, is set to zero for the front-steered SVEA vehicles.

3.2 Filtering Techniques

EKF and PF algorithms are well established and, in this project, are based on those in [Thrun et al. 2006]. Illustrations of the implemented filter configurations can be seen in Figures 3 and 4. All inputs entering the left side of a filter block are configured as process/prediction steps and all inputs entering the bottom are configured as observation updates.

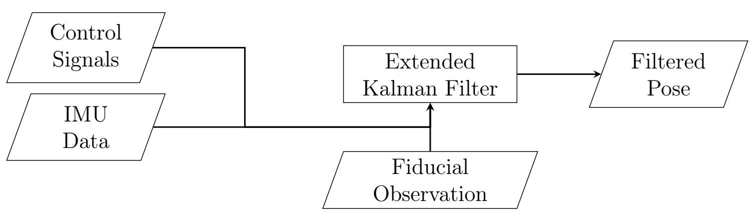


Figure 3: Extended Kalman Filter Flowchart

Note that all inputs into the EKF node were used as observation measurements due to limitations of the 'robot_localization' EKF package, despite intuition suggesting that IMU and control inputs should be used in a prediction step. This is discussed further in Section 4.4. The implementation of the particle filter uses the fused control and IMU data from the EKF, dubbed 'Filtered Odometry', to calculate prediction steps.

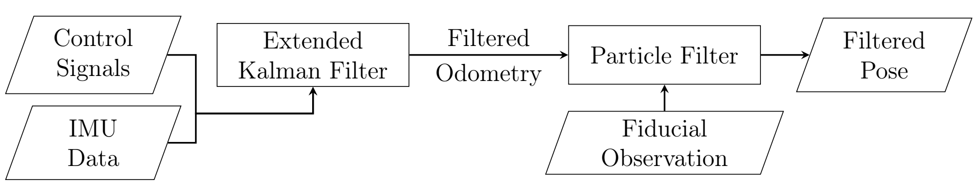


Figure 4: Particle Filter Flowchart

4 Experimental Implementation

4.1 Experimental Structure

Figure 5 demonstrates the logic used to obtain the pose estimates, the ‘True Pose’, and the estimation error as a SVEA vehicle is driven around the Smart Mobility Lab. ‘IMU/Control Estimate’ represents the ‘Filtered Odometry’ obtained from fusing IMU data with control inputs and can be thought of as the pose estimate from the motion model. The Qualisys motion capture system gives a sub-millimeter accurate ‘SVEA 1 True Pose’, referred to as ‘True Pose’. Qualisys also gives the true pose of the stationary camera used in most experiments and the second SVEA vehicle used in the Two Vehicle (Mobile Camera) experiments. The ‘Detected SVEA 1 Pose’ is the estimate of the first SVEA vehicle’s pose based on fiducial marker detection and the camera’s or second SVEA’s true pose. This pose is filtered with the ‘IMU/Control Estimate’ to give the ‘Filtered SVEA 1 Pose’, and is compared against the ‘True Pose’ for the ‘Error Calculation’. This structure, relying on the ROS ‘tf2’ package, enables all poses to be visualized and compared in the ‘Global/Map Coordinate Frame’.

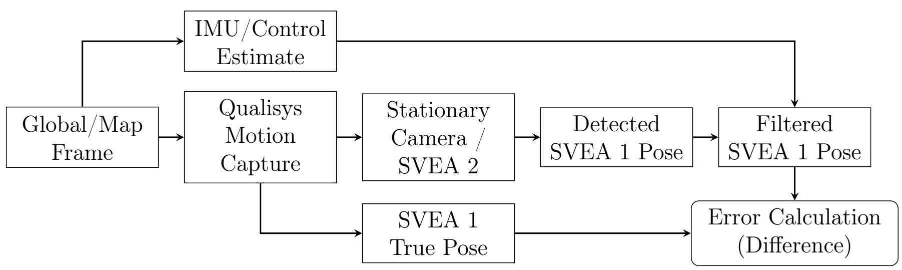


Figure 5: Experimental Setup - Coordinate Frame Diagram

4.2 SVEA Description

The Small-Vehicles-for-Autonomy (SVEA) platform of the KTH Smart Mobility Lab was used to implement and test the motion models and localization filters developed in this project. One representative configuration of the SVEA vehicles can be seen in Figure 6. The fiducial marker placed on the side of the vehicle enables stationary cameras, or those mounted on other SVEA vehicles, to estimate the pose of this SVEA vehicle. This marker can also be affixed to the rear of the vehicle, as a license plate would be, for the two vehicle tests. The reflective markers are used by the infrared motion capture system, Qualisys, to determine the pose of the vehicle in the global/ map reference frame. This pose is dubbed the ‘True Pose’ and is the benchmark by which the accuracy of all pose estimates is measured.

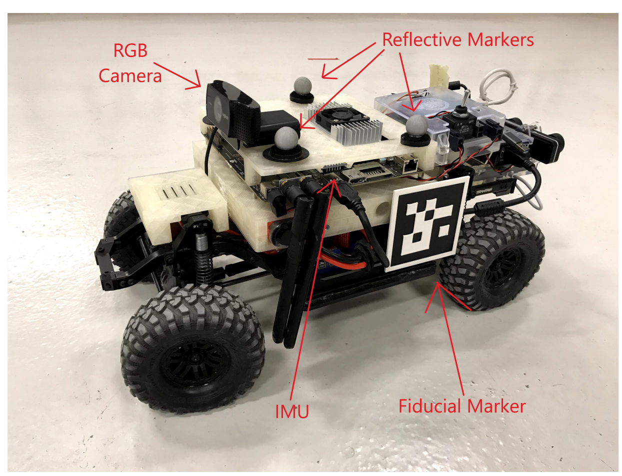


Figure 6: SVEA Vehicle with Relevant Components Labelled

4.3 ROS Package Dependencies

This project was heavily dependent on a few key ROS packages. First and foremost is the KTH Smart Mobility Lab 'svea_starter' package that enables remote or programmed control of the SVEA vehicle and publishes sensor measurements and executed control commands. The 'aruco_detect' package of the 'fiducials' open-source group of packages was also critical to this project. This package detects the fiducial marker in the camera's frame and calculates the coordinate frame transform between the camera and the marker. This transformation is converted through a series of 'tf2' coordinate frame transformations and is formatted as a 'PoseWithCovarianceStamped' ROS message type in the global/map frame. Additionally, 'qualisys' was used to determine 'True Pose' as explained in Section 4.2 and 'video_stream_opencv' was used for all video streaming from RGB cameras.

4.4 Robot_Localization Package

The 'robot_localization' package contains an EKF node that was used in this project for all EKF applications by configuring a .yaml text file. Necessary configurations for this file include sensor types, sensor message formats, covariance values, and more advanced settings including how differentiation and integration of various sensor readings should occur. This node appears to allow for a control input to be used in a prediction step, however in practice this is very difficult to implement. To allow the node to function properly, all inputs into the EKF node are used as observation measurements. By setting a high publish rate of the IMU readings and control inputs, 100 hz, as compared to the EKF pose estimate publish rate, 30 hz, the lack of an accurate prediction step is negligible. When considering the long periods of occlusion when the camera does not detect the fiducial marker, the updates from the IMU and control signals can be thought of to act as prediction steps that are updated by the infrequent (on the order of 10 seconds) fiducial detection measurements.

4.5 Kinematic Bicycle Model

The SVEA platform reports the speed and steering control inputs as 8-bit integers on the range [-127,127]. However, there is a dead zone in the speed control on the range [-20,20] where the servomotors do not have enough power to propel the vehicle forwards or backwards. By accounting for maximum speeds and steering angles, the reported 8-bit speed and steering values are converted into relevant linear and angular velocity units through the kinematic bicycle model and formatted as a 'TwistWithCovarianceStamped' ROS message type.

4.6 Particle Filter

The particle filter is implemented based on the algorithm in [Thrun et al.] with sensor inputs as shown in Figure 4. The 'robot_ localization' EKF node is used to fuse control inputs, converted using the kinematic bicycle model, and raw IMU data into a 'Filtered Odometry' message. This message is used for linear motion prediction/process updates, with Gaussian process noise, in the particle filter. This was done to take advantage of the fusion of control and IMU data completed while implementing the EKF and to make the particle filter more configurable for other applications. The use of the 'Odometry' ROS message type here enables other process update input sources to be implemented with this particle filter node. Both multinomial and systematic resampling methods were implemented, but the multinomial method was used during all experimental data collection for consistency.

4.7 Covariance Values

Selection of proper covariance values for all sensors, processes, and updates were critical for proper calibration and operation of both filters. Covariance values for raw IMU data came from the manufacturers data sheet, while values of the control inputs were manually tuned to give approximately equal weight between IMU and control inputs. Observation covariance values were typically set lower than the true accuracy of the fiducial observation pose estimates. This is so the observation updates could correct the poor pose estimates after long periods of occlusion of the fiducial marker from the cameras frame of view. Process covariance values for the particle filter were calculated and tuned based on measured error of the motion model, and adjusted for the 30 hz frequency of process updates.

5 Experimental Results

This section details the experimental results obtained with various configurations of physical setups, localization filters, and parameter settings. In all experiments, SVEA vehicles were driven around using remote control and no knowledge of the 'True Pose' determined by the Qualisys motion capture system was used in the pose estimates. The paths taken by all relevant pose estimates and the 'True Pose' are tracked over time and displayed in X-Y plots, as seen in Figures 7 - 10. All error measurements in this section are reported as the Euclidean distance (omitting heading) from the relevant pose estimate to the 'True Pose'.

5.1 Extended Kalman Filter

Figure 7 demonstrates an implementation of the EKF filter with high process covariance values, on the order of 10x the values used for the well-tuned motion models seen in Figure 8. The motion model is very inaccurate, reaching error values up to 2.9 m. The observation update from fiducial marker detection corrects the 'Filtered Pose' estimate by 1.5m to an error of only 0.2 m. These results show a successful integration of the fiducial detection observation updates with the EKF filter. The 0.2 m 'Filtered Pose' error will be seen to occur regularly and comes from miscalibration of static coordinate transforms, such as the transform representing the mounting location of the fiducial marker on the SVEA.

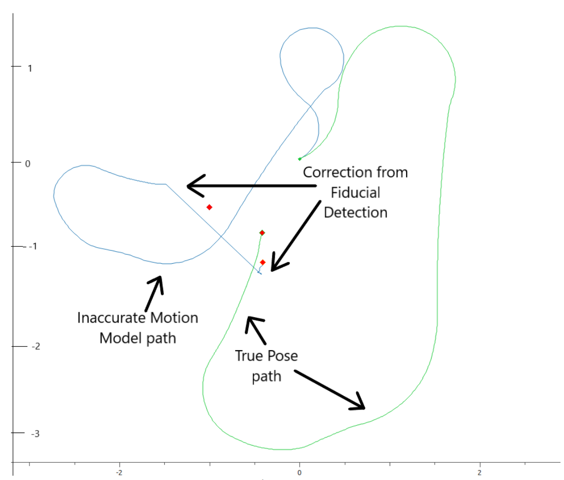
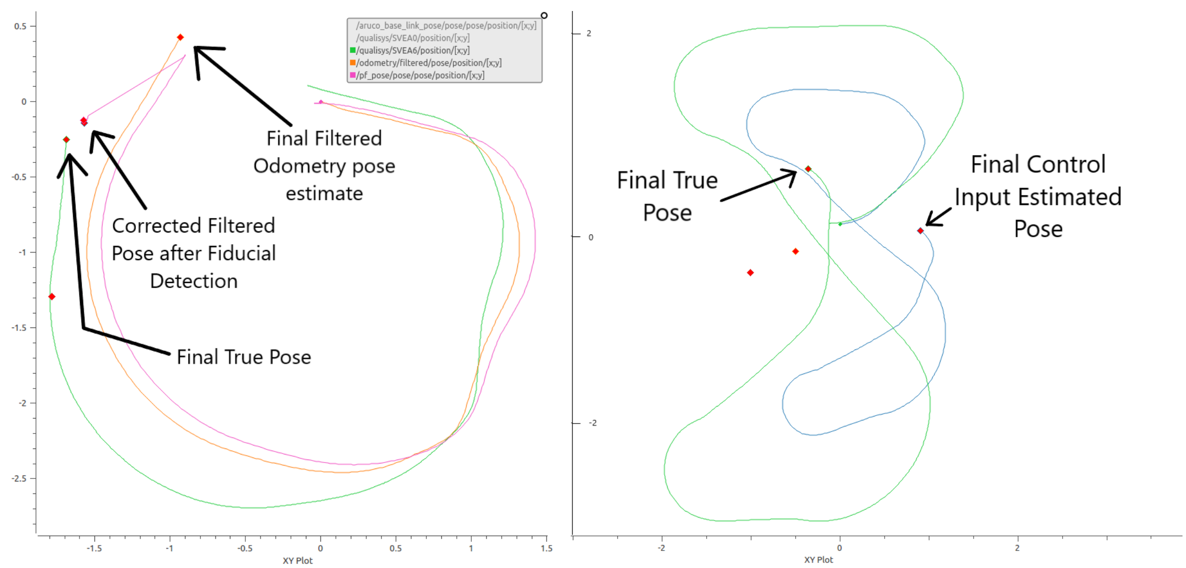


Figure 7: Extended Kalman Filter Localization with High Process Covariance

5.2 Motion Model and Particle Filter

Figure 8a demonstrates the high level of accuracy of the well-tuned motion model during long periods of occlusion. As detailed in Section 4.5, this model was obtained by fusing IMU data with control inputs, converted through a kinematic bicycle model. After driving the majority of a ∼2.5 m diameter loop over ∼10 seconds, the 'Filtered Odometry' pose estimate can be seen to have an error of 1.0 m from the 'True Pose'. The particle filter corrects this error when the fiducial marker is detected, resulting in a 'Filtered Pose' estimate error of 0.2 m. Figure 8b shows a similar level of error, 1.1 m, obtained by excluding IMU data from the motion model and using only control signals for the prediction steps. Contrarily, the motion model accumulates massive error when excluding the control inputs and using only IMU data for the motion model, as will be seen in Figure 9a. This shows that a reasonable motion model can be obtained by only using control inputs with a kinematic bicycle model.



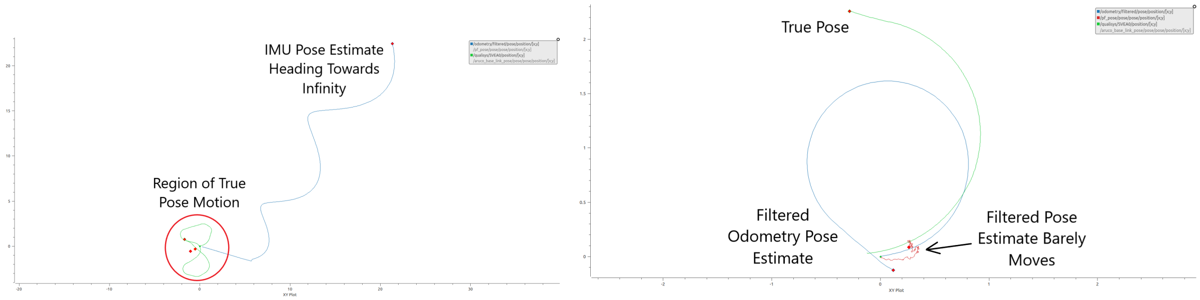
(a) Particle Filter Localization

(b) Motion Model without IMU data

Figure 8: Localization with well-tuned Motion Models

5.3 Interesting Filter Failures

Figure 9 contains the results of two test cases in which the EKF and Particle Filter failed due to improper configuration. Figure 9a depicts a test case in which the control inputs were excluded from the motion model. Even when stationary, the IMU consistently reports accelerations in the X and Y directions due to sensor bias and probable misalignment of the sensor. Without frequent zero velocity updates from the control signals cancelling this error, the motion model assumes the vehicle is moving based on these accelerations and the pose estimate moves accordingly. Figure 9b depicts a test case in which a large amount of process noise is introduced to the Particle Filter. Process covariance values of 0.15 m, when multiplied by a 30 hz update rate, lead to a 4.5 m/s prediction estimate covariance spread. This exceeds the maximum speed of the SVEA vehicle, ∼1.5 m/s, which leads to the pose estimate being determined primarily by noise. Since the Gaussian noise is centered on zero, the filtered pose estimate does not stray very far from the origin.



(a) Motion Model without Control Inputs

(b) Particle Filter Localization with Extremely Large Process Covariance

Figure 9: Results of Interesting Test Cases

5.4 Two Vehicle (Mobile Camera)

For the final experiments, the stationary camera detecting the fiducial marker is replaced by a camera mounted on a second SVEA vehicle following behind the first SVEA. The fiducial marker is mounted on the rear of the first SVEA in the spirit of a license plate. The second SVEA estimates the pose of the first SVEA by using knowledge of its own pose, control and IMU data from the first SVEA, and relative pose calculations when the fiducial marker is detected. Figure 10 demonstrates the results obtained from one such experiment in which a Particle Filter with a high particle count of 15,000 particles is used for localization. The 'Filtered Pose' estimate is corrected at two different instances when the fiducial marker is detected. The effectiveness of this implementation can be understood by comparing the final 'Filtered Odometry' pose estimate error of 3.4 m to the final 'Filtered Pose' estimate error of 0.2 m.

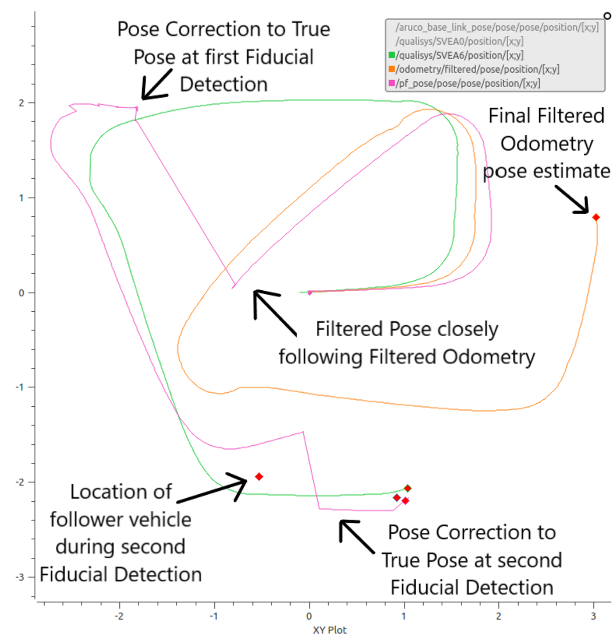


Figure 10: Mobile Camera, High Particle Count Filter

6 Experimental Conclusions

In these experiments, the Extended Kalman Filter and the Particle Filter were successfully implemented on the KTH Smart Mobility Lab's Small-Vehicles-for-Autonomy (SVEA) platform. In the base case, the prediction/process update was calculated by fusing raw IMU data and control input signals through a kinematic bicycle model. To test the merit of each of these sensory inputs in the motion model, a filtering process was run with each input individually omitted. Considering the approximations made in the conversion of the 8-bit control signals, it was anticipated that the IMU-only model would outperform the control-only model. Surprisingly, the control-only motion model performed as well as the model with both control and IMU inputs fused (Fig. 8a and 8b), while the IMU-only model accumulated massive error (Fig. 9a). This is significant, as removal of the IMU further reduces hardware requirements and cost for the ongoing project studying fleets of autonomous mobile robots with minimal hardware. In terms of filtering effectiveness, the EKF and PF proved to obtain equally accurate pose estimates when properly calibrated with smooth SVEA driving.

7 Further Expansion

The two vehicle pose estimation method developed for this experiment will be extended by the KTH Smart Mobility Lab to enable automated vehicle fleeting control. The use of fiducial markers in full sized vehicle platooning applications is an active research topic with a recent implementation on freight trucks by [Winkens and Paulus 2017], building off of their previous work [Winkens et al. 2015]. Winkens and Paulus attempt to resolve the issue that image resolution limitations make fiducial marker recognition difficult at moderate distances of 25 meters or greater, even with very large markers. They do so by tracking the shape of the marker and, at even greater distances, the body of the lead vehicle itself by using the last pose of the marker as a starting point for object detection based on optical flow estimation and global matching. This has direct correlation to the 1/10th scale SVEA vehicles, on which marker recognition falls off around 3 meters.

This 25 meter detection range for large fiducial markers coincidentally corresponds to a time gap of 1.0 second between lead and following vehicle when driven at 90 km/h, the speed limit for heavy trucks on Swedish motorways. While this confirms there is application for Winkens' and Paulus' detection method with highway platooning, their exclusion of odometry information from the lead vehicle is unnecessary. They designed their solution to eliminate communication requirements between vehicles, such that the lead vehicle could be any human-operated truck with fiducial markers affixed to the back. However, when considering the context of the fully autonomous following truck, retrofitting a speedometer data transmitter onto the lead truck is trivial. For this reason, the filters including lead vehicle control-inputs developed in this project will be used for the SVEA platooning research.

The stationary camera implementation will be expanded to study 'Smart Intersections' and methods for optimizing traffic flow. In this application, the stationary camera(s) would estimate trajectories of oncoming vehicles and would give necessary instructions for vehicles to speed up or slow down so that both vehicles could cross the intersection without colliding or coming to a halt. The ultimate use for the fleeting and intersection solutions is in an active project studying the use of centralized control for fleets of autonomous mobile robots with minimal hardware. By developing filtered localization based solely on control inputs and perception with RGB cameras detecting fiducial markers, hardware requirements on the robots becomes very low and the cost drops accordingly.

8 References

- J. Kong, M. Pfeiffer, G. Schildbach, F. Borrelli. 2015. Kinematic and Dynamic Vehicle Models for Autonomous Driving Control Design. 2015 IEEE Intelligent Vehicles Symposium; Seoul, Korea
- A. Mutka, D. Miklic, I. Draganjac, S. Bogdan. 2008. A Low Cost Vision Based Localization System Using Fiducial Markers. Proceedings of the 17th World Congress, The International Federation of Automatic Control; Seoul, Korea
- S. Konatowski, P. Kaniewski, J. Matuszewski. 2016. Comparison of Estimation Accuracy of EKF, UKF, and PF Filters. Annual of Navigation, Volume 23, Issue 1, 2016. DOI: 10.1515/aon-2016-0005
- S. Thrun, W. Burgard, D. Fox. 2006. Probabilistic Robotics. The MIT Press. Cambridge, Massachusetts.
- C. Winkens, D. Paulus. 2017. Long Range Optical Truck Tracking. Proceedings of the 9th International Conference on Agents and Artificial Intelligence. DOI: 10.5220/0006296003300339
- C. Winkens, C. Fuchs, F. Neuhaus, and D. Paulus. 2015. Optical Truck Tracking for Autonomous Platooning. International Conference on Computer Analysis of Images and Patterns. DOI: 10.1007/978-3-319-23117-4_4