Stefano Ruggiero – project 2

Multi-Object Tracking and Sensor Fusion with Kalman and Particle Filters

Objectives

The objective of this project is to develop a multi-object tracking (MOT) system using sensor fusion on the NuScenes dataset. Specifically, the project combines Lidar and camera data to improve accuracy in tracking multiple objects across frames, handling challenges such as noise, occlusions, and varying environmental conditions. Two estimation techniques, the Kalman filter for sensor fusion and the particle filter for motion prediction, are implemented to test their effectiveness in object tracking scenarios.

Methods

1. **Data Loading and Sensor Preprocessing**: Using the NuScenes dataset, which provides synchronized data from multiple sensors, Lidar data was transformed into the camera frame to enable a consistent coordinate system for fusion. This alignment, achieved through calibration using intrinsic and extrinsic parameters, allows the data from different sensors to be accurately combined, a necessary condition for reliable sensor fusion.
2. **Bounding Box Projection and Data Association**: Camera frames were used to detect 2D bounding boxes around objects, while 3D Lidar points were projected onto these frames. The association of Lidar points with detected objects was achieved using a cost matrix based on Euclidean distances, minimized via the Hungarian algorithm. This association process enables each Lidar point to be matched with an object, creating a consistent object-to-point mapping across frames.
3. **Kalman Filter for Sensor Fusion**: The Kalman Filter's measurement matrix maps only positional information, while process and measurement noise matrices were configured to account for Lidar and camera sensor variances. This configuration enables iterative updates to position and velocity estimates based on new measurements.
4. **Particle Filter for Motion Prediction**: To address potential non-linear motion, a particle filter was implemented for prediction purposes. This filter initializes particles around the object positions from the Kalman-filtered data and updates particle positions based on random variations to simulate possible movements. Weights are assigned to particles based on their proximity to actual measured positions, and systematic resampling was employed to retain particles around the most likely positions. This approach aims to capture more complex motion patterns beyond the assumptions of linearity inherent in the Kalman Filter.
5. **Evaluation and Visualization**: The tracking system's accuracy was evaluated using the Root Mean Square Error (RMSE) between predicted positions and ground truth data. Visualization functions were developed to display bounding boxes, projected Lidar points, and associations on camera images, providing a way to verify the associations and fused data visually.

Results

The RMSE values obtained from tracking across different scenes were notably high, indicating limited accuracy in the system's tracking performance. The results are as follows:

- RMSE for scene 1 (scene-0061): 842.10
- RMSE for scene 2 (scene-0103): 1202.75
- RMSE for scene 3 (scene-0553): 1173.12
- RMSE for scene 4 (scene-0655): 1435.79

- RMSE for scene 5 (scene-0757): 592.14
- RMSE for scene 6 (scene-0796): 2284.67
- RMSE for scene 7 (scene-0916): 1357.44
- RMSE for scene 8 (scene-1077): 1413.59
- RMSE for scene 9 (scene-1094): 1334.38
- RMSE for scene 10 (scene-1100): 1070.78

The average RMSE across all scenes was 1270.68, a high error margin that suggests significant challenges in achieving accurate object tracking with this implementation.

Analysis of Results

The high RMSE values may stem from several factors inherent in the chosen tracking and fusion approach:

1. **Linear Assumptions in Kalman Filter**: The Kalman Filter relies on linear assumptions and constant velocity, which may not suit real-world, complex motion patterns, especially in highly dynamic scenes. This limitation could cause substantial inaccuracies when objects follow non-linear trajectories, contributing to the high RMSE values.
2. **Data Association Errors**: The Euclidean-based cost function used in the Hungarian algorithm for associating points to bounding boxes may not handle overlapping or densely clustered objects effectively. This could result in incorrect associations, leading to compounded errors in position estimates across frames.
3. **Particle Filter Variability**: While particle filters can manage non-linear movements, their accuracy is sensitive to the number of particles and noise parameters. In this implementation, the particle filter may not have been tuned optimally for the variety of object motions present in the dataset, reducing its ability to accurately predict future positions and resulting in higher RMSE.

Conclusion

This project established a functional sensor fusion and tracking pipeline, but the results indicate significant limitations in accuracy and robustness. The Kalman Filter's linear assumptions and the particle filter’s limited impact highlight the need for alternative methods. This project illustrates the challenges of multi-object tracking in dynamic environments and underscores the importance of matching filtering techniques to real-world complexities for reliable tracking outcomes.