

Q1: Why we need to do data association in tracking module?

A1: We use an object detector to provide the Kalman Filter with N measurements. As the detector results could be noisy, we need to design a data association mechanism to decide which detection to pair with a predicted object state and which detections to treat as outliers.

Q2: Which method does this module use for the data association? How it works?

A2:

This module use Mahalanobis distance to measure the affinity between predictions and detections. This distance m measures the difference between predicted detections and actual detections weighted by the uncertainty about the prediction as expressed through the innovation covariances. In the experiments, it shows that the Mahalanobis distance provides better tracking performance than the 3D-IOU.

Given the distances between all predictions and detections, we solve a bipartite matching problem to find the optimal pairing by employ a greedy algorithm with an upper bound threshold value to solve this problem. Compared to the Hungarian algorithm as used in the AB3DMOT baseline, the greedy approach performs better.

It also adopted the orientation correction approach from the AB3DMOT baseline. Specifically, when the angle difference between the detection and prediction is between 90 and 270 degrees, it rotate the prediction's angle by 180 degrees before calculating the Mahalanobis distance.

Q3: How does this module compute the covariance matrix for Kalman filter? Any other method to compute the covariance? What's the difference?

A3: This module use the statistics of the training set data to estimate the initial state covariance, the process and observation noise covariance. There are also other methods to compute the covariance, such as using the identity matrices and heuristically chosen scalars to build the covariance matrices of the Kalman Filter as in AB3DMOT. The difference between using the statistics of the training set data to estimate the initial state covariance and identity matrices is that the former method may have different outputs at different dimensions.