Sensor Fusion-Based Object Tracking For Self-Driving Cars Using Kalman Filter

Abstract—**Autonomous vehicles to have fully functional autonomy, there is a need to process the data defining the surrounding environment. Hence, sensors like LiDAR, RADAR and camera are used. Tracking is one of such autonomous function where it is defined as, to follow the trail or movements of (someone or something), typically to find them or note their course. In achieving this sensor fusion have promised minimum error in estimation for challenging to the conditions too. Hence this paper provided an overview of tightly coupled fusion techniques for Kalman filter, Extended Kalman filter and Unscented Kalman filter for LiDAR RADAR sensor for object tracking, for a road scenario of variant Radius of curvature for comparing there performance.**

Keywords : RADAR, LiDAR, Sensor Fusion, Extended and Unscented Kalman Filter, Radius of curvature.

I. Introduction

In todays fast-moving world, automation is a necessity in every field to improve the quality of life, it allows us to have quality products and services at faster rates and minimize human error, time and cost. Autonomous Vehicles boon to human life to reduce ever-increasing road accidents and traffic congestion's, this led to the need for self-driving cars(SDC).

An SDC is a vehicle that uses multiple sensors, artificial intelligence, sensor fusion, and computer vision to travel without any human interference. To understand the surrounding environment, there is a need to predict the values like position, velocity, yaw angle, and yaw rate of preceding obstacles eg: pedestrian, bicycle, car, etc. Further actions like accelerating, decelerating, braking of our ego vehicle, are done depending on necessary values, these values are acquired using LiDAR, RADAR, a Depth Camera, and Ultrasonic sensors to achieve safe and smooth driving experience.

The depth camera measures the distance of each point by transmitting invisible near-infrared light and measuring its time of flight after it reflects off the objects. It is used across a variety of applications including automotive, SLAM, robotics, unmanned aerial vehicles and more.

The ultrasonic sensor is a device that measures the distance to an object using ultrasonic sound waves. An ultrasonic sensor uses a transducer to send and receive ultrasonic pulses that relay back information about an object's proximity. It is used across a variety of applications including automotive, robotics and more.

LiDAR, which stands for Light Detection And Ranging, is a method that measures the relative distance to object by illuminating the target with laser light and measuring the reflected light with a sensor. It is used across a variety of applications including automotive, mapping, robotics, security, smart cities and more.

RADAR, which stands for Radio Detection And Ranging, with the help of radio waves it measures radial distance, radial velocity, and yaw angle. It is used across a variety of applications including military, aircraft, ship, space, automotive, robotics, and more.

Sensor fusion is the process of combining two or more sensor data to improve performance. Visual perception, to understand the surrounding environment of self-driving cars, generally sensors used for this purpose are LiDAR, RADAR and Camera. The output of visual perception, localization and map are given to path planning algorithms for avoiding obstacles by tracking the desired trajectory, also decides the shortest path from starting point to destination point. A further output of visual perception and path planning are given to car control i.e. steering angle, acceleration, and brake.

In multiple sensor data fusion process, individual sensor error is taken care of. LiDAR and RADAR gives error in position, velocity and yaw angle i.e. turning angle in the z-direction. Estimation algorithms like Kalman filter(KF), Extended Kalman filter(EKF) and Unscented Kalman filter(UKF) are used for multiple sensor data fusion and sensor error reduction, it has numerous applications like navigation, control of vehicles, etc.

KF is an algorithm which takes measurements from sensor containing noise as input and produces estimations which are more accurate than the sensor measurements. It can generally handle linear equations, whereas EKF and UKF algorithms can handle non-linear equations. These algorithms generally consist of 3 steps initialization, prediction, and update.

This paper deals with the sensor fusion of LiDAR and RADAR sensors for visual perception of self-driving cars, we used fused data of both sensors, which yields better results than individual sensor data. A brief of the different methods or techniques employed for multi-sensor data fusion as reported by various authors is summarized below.

Hojoon Lee *et al.* [1] has proposed a methodology for LiDAR and RADAR data fusion. The accuracy of position is better in the LiDAR sensor, for yaw angle accuracy of RADAR is better, by the method of sensor fusion can yield better accuracy for position and yaw angle.

Hatem Harji and Rahal [2] have proposed a methodology for obstacle detection using both LiDAR and RADAR sensors, the advantage of using multiple sensors over single sensor is illustrated by making use of several tracking scenarios, generated ground truth using Mean Squared Error(MSE) method, later estimated MSE of position and velocity of LiDAR, RADAR and fused data.

Shengho Eben Li *et al.* [3] has presented a cost-effective approach to track objects using linear ultrasonic sensors, arranged 8 ultrasonic sensors linearly to expand range using EKF and UKF. The advanced tracking method(ATM) was used to compare the tracking performance of EKF and UKF.

Yassine Malej *et al.* [4] has presented an approach to track vehicles using 3D LiDAR sensor, tracking based on the anchor boxes of vehicles, recognized through a convolutional neural network(CNN) for the perception of autonomous vehicles, tested on KITTY data set, later applied KF and EKF to minimize error in CNN prediction.

Marco Chiani *et al.* [5] has discussed localization and tracking of objects utilizing a RADAR sensor, made use of signal processing to improve the accuracy of tracking of obstacles in the environment.

Ankit Manjunath *et al.* [6] has proposed a methodology for object detection and tracking for self-driving cars using a RADAR sensor. RADAR measurements of every object combined to a single object and tracked using EKF and UKF, later tested in multiple tracking scenarios.

Mobus and Kolbe [7] presented algorithms for multi-sensor data fusion of infrared and RADAR sensor data using Interacting Multiple Model(IMM) filter and KF, IMM filter applied for infrared sensor while KF for RADAR, shown that fused data of RADAR and Infrared considerably increases range and accuracy compared to a single sensor for tracking.

II. Methodology

Tightly coupled multiple sensor data fusion is a method in which multiple sensors are dependent upon each other in a intertwined method, hence the output of one section is feed to other, influencing once outcome on the other sensor.

1. Kalman filter

KF is an algorithm which takes measurements from sensor to reduce unwanted variance in the output data. It can generally handles linear equations. Filters involves 3 steps i.e. Initialization, Prediction, and Update is presented in figure 1.

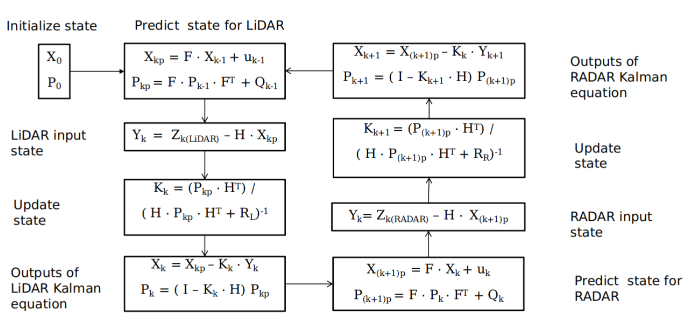

Fig. 1. Kalman Filter

1.a Initialization

In Initialization state, Used Constant Speed (CS) motion model, hence considered position velocity (px, py, vx, vy) in the state vector. Where X_0 represents the initialized object state vector, we initialized the objects position i.e. px, py using either LiDAR or RADAR sensor data be taken for initialization depending on the sequence, for reference RADAR data is taken as initialization data. And object's velocity i.e. vx, vy to 0. P_0 represents the initialized state co-variance matrix, we initialized diagonal elements of P_0 to unity, because variables of the state vector are independent of each other.

1.b Prediction

$$X_{kp} = F \;\; X_{k-1} + u_{k-1} \tag{1}$$

$$P_{kp} = F \;\; P_{k-1} \;\; F^T + Q_{k-1} \tag{2}$$

X_{kp} represents predicted object state, F represents state transition matrix, it is designed based on state transition equations of linear motion model as in equation (1) . P_{kp} represents the predicted state co-variance matrix. u_{k-1} represents process noise, it gives us information about the car we are tracking is accelerating, decelerating or braking. Q_{k-1} represents the process noise co-variance matrix, it gives information about the error in the acceleration shown in equation (2).

State transition equations are represented in the equation (3 - 6).

$$p'_{\mathrm{x}} = p_{\mathrm{x}} + v_{\mathrm{x}}\Delta t + a_{\mathrm{x}}\Delta t^2 \tag{3}$$

$$p'_{\mathrm{y}} = p_{\mathrm{y}} + v_{\mathrm{y}}\Delta t + a_{\mathrm{y}}\Delta t^2 \tag{4}$$

$$v'_{\mathrm{x}} = v_{\mathrm{x}} + a_{\mathrm{x}}\Delta t \tag{5}$$

$$v'_{\mathrm{y}} = v_{\mathrm{y}} + a_{\mathrm{y}}\Delta t \tag{6}$$

Δt represents the rate of change of time,(for this experiment Δt is 1s) a_x represents an acceleration in the x-direction, which can be acquired by taking the difference in position in the x-direction of the previous and present state, a_y represents an acceleration in the y-direction, which can be acquired by taking the difference in position in the y-direction of the previous and present state.

1.c Update

$$Y_{\mathrm{k}} = Z_{\mathrm{k}} - HX_{\mathrm{kp}} \tag{7}$$

$$K_{\mathrm{k}} = (P_{\mathrm{kp}}H^{\mathrm{T}})/(HP_{\mathrm{kp}}H^{\mathrm{T}} + R) \tag{8}$$

$$X_{\mathrm{k}} = X_{\mathrm{kp}} + K_{\mathrm{k}}Y_{\mathrm{k}} \tag{9}$$

$$P_{\mathrm{k}} = (I - K_{\mathrm{k}}H)P_{\mathrm{kp}} \tag{10}$$

Y_k represents the innovation matrix. Z_k represents LiDAR and RADAR sensor measurements. H represents a measurement matrix, it is a changing unit of the state vector, to match sensor measured values as in equation (7). K_k represents the Kalman gain as in equation (8). P_{kp} represents the state co-variance matrix. R_{LiDAR} represents a LiDAR sensor co-variance matrix, whose values are given by sensor manufacturers in the data-sheet. similarly R_{RADAR} represents a RADAR sensor co-variance matrix, whose values are given by sensor manufacturers in the data-sheet. I represent the identity matrix as in equation (9 and 10).

2. Kalman filter and Extended Kalman filter

EKF is an algorithm is a extented form of simple Kalman filter which is more accurate than the sensor measurements as It generally handle non-linear equations. It too involves 3 steps i.e. Initialization, Prediction, and Update. For LiDAR sensor Prediction, update and RADAR sensor prediction we used the KF algorithm. Only for the RADAR sensor update, we used the EKF algorithm. A Tightly coupled form of KF and EKF is presented in figure 2

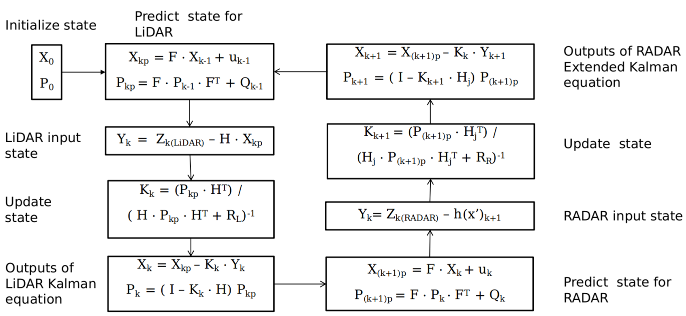

Fig. 2. Kalman And Extended Kalman Filter

As Initialization and Prediction steps follow as same as KF is presented in 1.a and 1.b in Kalman filter section. Hence update part is as follows.

2.a Update

$$Y_{\text{k}} = Z_{\text{k(LiDAR)}} - HX_{\text{kp}} \tag{11}$$

$$K_{\text{k}} = (P_{\text{kp}}H^{\text{T}})/(HP_{\text{kp}}H^{\text{T}} + R_{\text{LiDAR}}) \tag{12}$$

$$X_{\text{k}} = X_{\text{kp}} - K_{\text{k}}Y_{\text{k}} \tag{13}$$

$$P_{\text{k}} = (I - K_{\text{k}}H)P_{\text{kp}} \tag{14}$$

Y_k represents the innovation matrix. $Z_{k(LiDAR)}$ represents LiDAR sensor measurements. H represents a measurement matrix, it is a changing unit of the state vector, to match sensor measured values. K_k represents the Kalman gain. P_{kp} represents the state co-variance matrix. R_{LiDAR} represents a LiDAR sensor co-variance matrix, whose values are given by sensor manufacturers in the data-sheet. I represent an identity matrix as in equation (11 to 14).

$$Y_{\text{k+1}} = Z_{\text{k+1(RADAR)}} - h(x')_{\text{k+1}} \tag{15}$$

$$K_{\text{k+1}} = (P_{\text{(k+1)p}}H^{\text{T}}{}_{\text{j}})/(H_{\text{j}}P_{\text{(k+1)p}}H_{\text{j}}{}^{\text{T}} + R_{\text{RADAR}}) \tag{16}$$

$$X_{\text{k+1}} = X_{\text{(k+1)p}} - K_{\text{k+1}}Y_{\text{k+1}} \tag{17}$$

$$P_{\text{k+1}} = (I - K_{\text{k+1}}H_{\text{j}})P_{\text{(k+1)p}} \tag{18}$$

As in equation (15 to 18) are of update step of EKF, for RADAR measurement estimations. $Z_{k+1(RADAR)}$ represents RADAR sensor measurements. $h(x')_{k+1}$ represents a non-linear function, it contains changing the unit of the state vector, to match sensor measured values. H_j represents the Jacobian matrix, to approximately linearize the non-linear function $h(x'$, in this matrix we partially differentiate each variable of state vector with the sensor measurement variables. R_{RADAR} represents a RADAR sensor co-variance matrix, whose values are given by sensor manufacturers in the datasheet.

3.Unscented Kalman filter

UKF is an algorithm of probabilistic methods and more sutable to handle non-linear equations. It involves 3 steps i.e. Initialization, Prediction, and Update. Both LiDAR and RADAR are used for tracking; employing Unscented Kalman filter

3.a Prediction

It has 3 sub-steps i.e. Generate sigma points, Predict sigma points and Predict mean and co-variance as shown in figure 3.

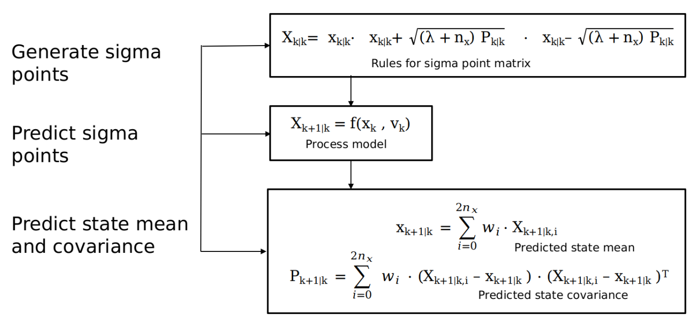

Fig. 3. Unscented Kalman Filter's Prediction state.

3.a.1 Generate sigma points

$$X_{\text{k}} = x_{\text{k}} \cdot x_{\text{k}} + \sqrt{(\lambda + n_{\text{x}})}P_{\text{k}} \cdot x_{\text{k}} - \sqrt{(\lambda + n_{\text{x}})}P_{\text{k}} \tag{19}$$

Where, $\lambda = 3 - n_x$

Used Constant Turn Rate Speed(CTRS) motion model, hence considered position, velocity, yaw angle, and yaw rate. $X_{k|k}$ represents a matrix with sigma points at time k. $x_{k|k}$ represents object state mean. we initialize the objects position i.e. p_x, p_y using LiDAR and RADAR sensor also object's velocity, yaw angle and yaw rate i.e. v, ψ, $\dot{\psi}$ to 0. λ represents scaling factor, it decided how far we need to spread sigma points. n_x represents the state vector dimension. $P_{k|k}$ represents state co-variance matrix, we initialize diagonal elements of it to unity, because variables of the state vector are independent of each other as in equation (19).

3.a.2 Predict sigma points

$$X_{\text{k+1|k}} = f(x_{\text{k}}, v_{\text{k}}) \tag{20}$$

i.e

$$X_{\mathrm{k+1|k}} = X_{\mathrm{k}} + \int_{t_{\mathrm{k}}}^{t^{\mathrm{k+1}}} g(x)dt + v_{\mathrm{k}} \tag{21}$$

x_k represents object state mean vector. v_k represents process noise, after passing each generated sigma points through process model, hence we get a matrix with sigma points at time k+1 i.e. $X_{k+1|k}$ (i.e. predicted set of sigma points in the x-y plane) expressed in equation (20) and (21).

3.a.3 Predict mean and co-variance

$$x_{\mathrm{k+1|k}} = \sum_{i=0}^{2n_{\mathrm{x}}} w_{\mathrm{i}} \cdot X_{\mathrm{k+1|k,i}} \tag{22}$$

$$P_{\mathrm{k+1|k}} = \sum_{i=0}^{2n_{\mathrm{x}}} w_{\mathrm{i}} \cdot (X_{\mathrm{k+1|k,i}} - x_{\mathrm{k+1|k}}) \cdot (X_{\mathrm{k+1|k,i}} - x_{\mathrm{k+1|k}})^{\mathrm{T}} \tag{23}$$

Where, $w_i = \lambda/(\lambda + n_x)$, i = 0,1,2,,,,,$n_x$

$X_{k+1|k,i}$ represents matrix with predicted sigma points at time k+1. w_i represents weights. n_x represents state vector dimension being expressed in expressed in equation (22) and (23)..

3.b Update

It has 2 sub-steps i.e. Predict measurement and Update state.

3.b.1 Predict measurement

$$Z_{\mathrm{k+1|k}} = h(x_{\mathrm{k+1}}) \tag{24}$$

$$z_{\mathrm{k+1|k}} = \sum_{i=0}^{2n_{\mathrm{x}}} w_{\mathrm{i}} \cdot Z_{\mathrm{k+1|k}} \tag{25}$$

$$S_{\mathrm{k+1|k}} = \sum_{i=0}^{2n_{\mathrm{x}}} w_{\mathrm{i}} \cdot (Z_{\mathrm{k+1|k,i}} - z_{\mathrm{k+1|k}}) \cdot (Z_{\mathrm{k+1|k,i}} - z_{\mathrm{k+1|k}})^{\mathrm{T}} + R \tag{26}$$

Where, $w_i = \Lambda/(\Lambda + n_x)$, i = 0,1,2,,,,,$n_x$

The equation (24) shows *h(x)* representing a non-linear function, it contains changing the unit of the state vector, to match sensor measured values. Now, we pass earlier generated each sigma points through measurement model, hence we get a matrix with sigma points at time k+1 i.e. $Z_{k+1|k}$ (i.e. measured set of sigma points in the x-y plane). $Z_{k+1|k,i}$ represents matrix with measured sigma points at time k+1in equation (25). w_i represents weights. n_x represents the state vector dimension. $Z_{k+1|k,i}$, $z_{k+1|k}$ represents the predicted measurement mean. *R* represents a sensor co-variance matrix, whose values are given by sensor manufacturers in the datasheet represented in equation (26).

3.b.2 Update state

$$T_{\mathrm{k+1|k}} = \sum_{i=0}^{2n_{\mathrm{x}}} w_{\mathrm{i}} \cdot (X_{\mathrm{k+1|k,i}} - x_{\mathrm{k+1|k}}) \cdot (Z_{\mathrm{k+1|k,i}} - z_{\mathrm{k+1|k}})^{\mathrm{T}} \tag{27}$$

$$K_{\mathrm{k+1|k}} = T_{\mathrm{k+1|k}} \cdot S_{\text{-1k+1|k}} \tag{28}$$

$$X_{\mathrm{k+1|k+1}} = x_{\mathrm{k+1|k}} K_{\mathrm{k+1|k}} \cdot (Z_{\mathrm{k+1(Sensor}} - z_{\mathrm{k+1|k}}) \tag{29}$$

$$P_{\mathrm{k+1|k+1}} = P_{\mathrm{k+1|k}} - K_{\mathrm{k+1|k}} \cdot S_{\mathrm{k+1|k}} \cdot K^{\mathrm{T}}{}_{\mathrm{k+1|k}} \tag{30}$$

The equations (27) to (30) $X_{k+1|k,i}$ represents matrix with predicted sigma points at time k+1. $x_{k+1|k}$ represents predicted state mean. $Z_{k+1|k,i}$ represent matrix with measured sigma points at time k+1. $z_{k+1|k}$ represents predicted measurement mean. $T_{k+1|k}$ represents cross-correlation. $x_{k+1|k}$ represents predicted state mean. $S_{k+1|k}$ represents the predicted measurement co-variance. $Z_{k+1(Sensor}$ represents LiDAR and RADAR sensor measurements. $z_{k+1|k}$ represents the predicted measurement mean. $P_{k+1|k}$ represents predicted state co-variance. $K_{k+1|k}$ represents Kalman gain.

III. Results and Conclusion

A simple road scenario of numeric '8' pattern is considered and is presented in figure 4 and Purple line, Red points, Green line, Pink points and Orange points represent Filter output, Sensor output of both LiDAR and RADAR, True Ground truth , LiDAR output and RADAR output respectively as shown in figure 5 of legends of results.

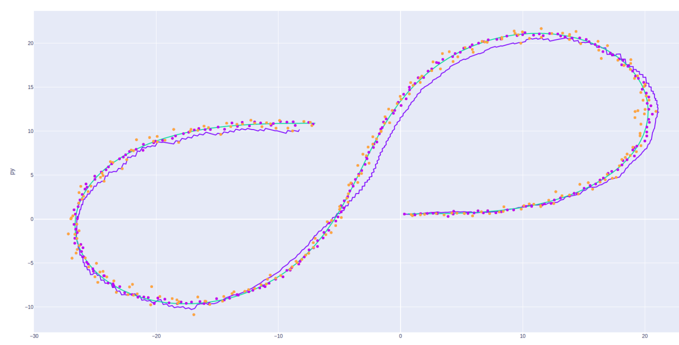

Fig. 4. Example of road scenario with various curvature and, RADAR and LiDAR data.

Fig. 5. Represents the legends for all result images.

The figures 6 to 8 represent the output of KF , EKF and UKF respectively for the curved portion of the same road scenario discussed in figure 4. One can clearly infer that as the non linear computation capability of the KF increases the Filter output completely follow the Ground truth. Hence it is evident that the non linearity of the real world scenarios can be solved using Kalman filters and its complex forms like Extended and Unscented . As here Unscented KF presents Promising output.

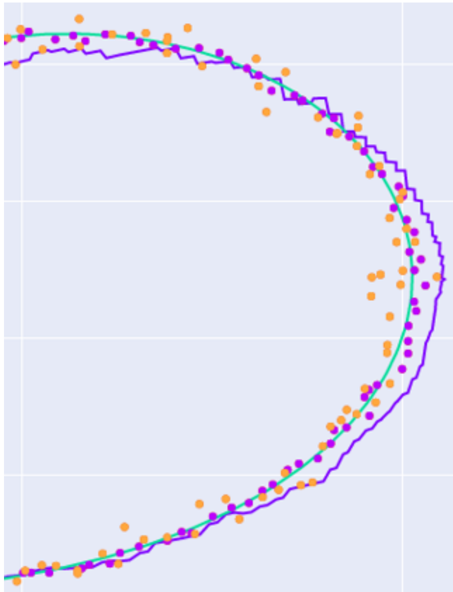

Fig. 6. Part of the road with KF

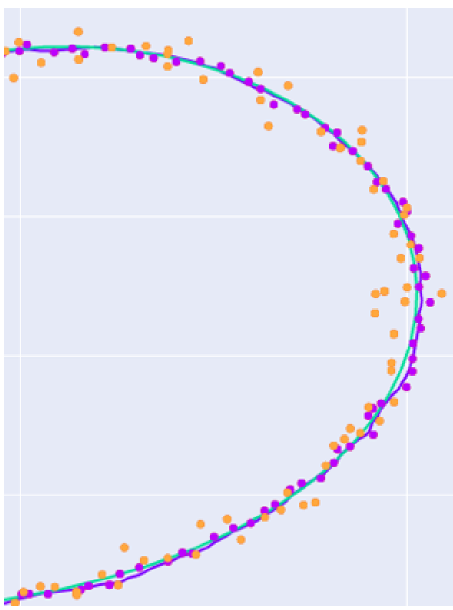

Fig. 7. Part of the road with EKF

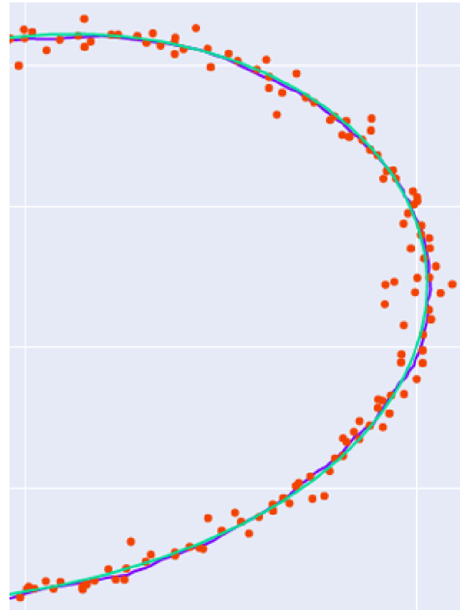

Fig. 8. Part of the road with UKF

REFERENCES

[1] Hojoon Lee, Heungseok Chae, Kyongsu Yi, "A Geometric Model based 2D LiDAR/Radar Sensor Fusion for Tracking Surrounding Vehicles, *IFAC-PapersOnLine*, Volume 52, Issue 8, 2019, Pages 130-135, ISSN 2405-8963

[2] Hatem Hajri and Mohamed-Cherif Rahal, Real Time Lidar and Radar High-Level Fusion for Obstacle Detection and Tracking with evaluation on a ground truth, July 2018

[3] Shengbo Eben Li, Guofa Li, Jiaying Yu, Chang Liu, Bo Cheng, Jianqiang Wang, Keqiang Li, Kalman filter-based tracking of moving objects using linear ultrasonic sensor array for road vehicles, *Mechanical Systems and Signal Processing*, Volume 98, 2018, Pages 173-189, ISSN 0888-3270

[4] Y. Maalej, S. Sorour, A. Abdel-Rahim and M. Guizani, "Tracking 3D LIDAR Point Clouds Using Extended Kalman Filters in KITTI Driving Sequences," *2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates,* 2018, pp. 1-6. doi: 10.1109/GLOCOM.2018.8647803

[5] M. Chiani, A. Giorgetti and E. Paolini, "Sensor Radar for Object Tracking," in *Proceedings of the IEEE,* vol. 106, no. 6, pp. 1022-1041, June 2018. doi: 10.1109/JPROC.2018.2819697

[6] A.Manjunath, Y. Liu, B. Henriques and A. Engstle, "Radar Based Object Detection and Tracking for Autonomous Driving," *2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM), Munich,* 2018, pp. 1-4. doi: 10.1109/ICMIM.2018.8443497

[7] R.Mobus and U. Kolbe, "Multi-target multi-object tracking, sensor fusion of radar and infrared," *IEEE Intelligent Vehicles Symposium, 2004, Parma, Italy,* 2004, pp. 732-737. doi: 10.1109/IVS.2004.1336475