Taking Advantage of Road Users Occluding the Road: Supporting Camera-based Road Tracking in Shared Spaces using Radar Doppler Measurements

Bianca Forkel, Philipp Berthold and Mirko Maehlisch

***Abstract*— Conventional road tracking approaches rely on direct measurements of the road. Using camera, LiDAR, or radar sensors, they detect lane markings, textures, road boundaries, curbs, or obstacles. However, all require a direct line of sight to the road. In shared spaces, e. g., a busy campus road, a lot of pedestrian groups are occluding the road for all sensors. Hence, we present a novel measurement concept to support road tracking in presence of other road users. In fact, we are taking advantage of them: Where other road users are moving, there might be the road. Therefore, we probabilistically identify dynamic objects based on radar Doppler measurements. Considering the measurement uncertainties, we then build a probability map of road user paths. This is used to derive probable road areas. Integrating this measurement into an existing camera-based road tracking, we evaluate the accuracy of the resulting road estimate using an HD map.**

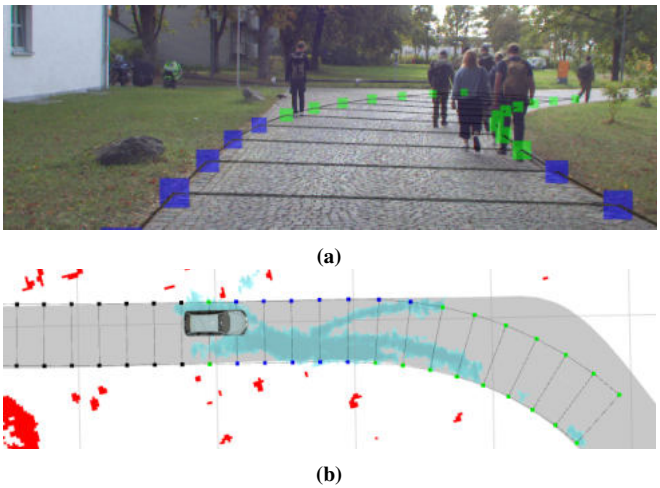
I. Introduction

Without map knowledge, a key task of autonomous driving is to perceive the road to drive on. In prior work, we proposed a road tracking framework [1] with measurement methods utilizing camera and LiDAR [2]. These methods target dirt roads, rural paved roads, and highways. Within these fields of application, we could demonstrate a high degree of robustness to different road appearances, weather and light conditions.

Nevertheless, the scenarios have one thing in common: they allow a direct, unobscured view of the road ahead. While dirt roads are usually free of other road users, the required safety distance ensures free sight on highways – at least on the next 50 m. There was no need to explicitly handle other road users. In this respect, an opposing scenario are narrow paths on a busy university campus. At peak times, a high number of pedestrian groups obscure the road especially for the camera. Figure 1a shows an example, where the camera is no longer able to observe the road course; the previous camera-based and LiDAR-supported road tracking would fail.

However, in this paper, we propose a new radar-based measurement method that exploits the behavior of other road users and turns their presence into an advantage: Since the other road users covering the road are in turn following the road, their positions provide information about the further road course. Hence, we detect dynamic objects using radar Doppler measurements and accumulate probable road user positions over time. Using their paths as additional information in our road tracking allows to successfully track the road in Fig. 1 around the bend, even though it is occluded in the camera.

All authors are with the Institute for Autonomous Systems Technology (TAS) of the University of the Bundeswehr Munich, Neubiberg, Germany. Contact author email: bianca.forkel@unibw.de



(a)

(b)

Fig. 1: Example of a busy campus road. Although pedestrians obscure the road in the camera image, our road tracking for shared spaces is able to start the bend. (**a**) Estimated lane boundary points projected into the camera image. Only the blue colored points could be measured in the camera image within the last 3s. (**b**) Top down view. Next to the estimated lane boundary points (blue/green, black behind the ego vehicle), accumulated pedestrian probabilities (cyan), static obstacles (red), and the ground truth road (gray) are shown.

Related work primarily focuses on estimating road geometries based on vehicle trajectories. [3] and [4] use a database of GPS trajectories to extract and update an offline road map. For highways, [5] and [6] also utilize sensor-based vehicle tracking to obtain lane information for vehicle following applications. However, these approaches are limited when dealing with pedestrians who do not strictly adhere to defined lanes. Other approaches using radar sensors for road tracking depend on artificial structures: Some detect special reflective [7], [8] or conventional [9] road markings, while others rely on road barriers or obstacles alongside the road [10], [11]. Yet, there are no lane markings, barriers, or curbs on the shared roads we aim to track.

Our radar measurement method does not require artificial structures or tracks of other road users. Since it integrates with camera and LiDAR measurements into a larger road tracking framework, we further support more complex roads also on segments without any other road users. Hence, Section II summarizes the key ideas of our existing road tracking framework. Sections III and IV then describe our new radar-based measurement method and its fusion into the road tracking. Finally, in Section V, the resulting road tracking for shared spaces is evaluated quantitatively using an HD map.

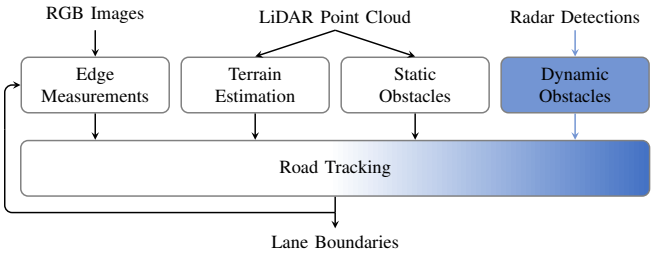


Fig. 2: System architecture of our road tracking for shared spaces. The contribution of this work is shown in blue.

II. Road Tracking Framework

As discussed, we add a new measurement method to a road tracking framework developed in previous work. The original camera-based and LiDAR-supported road tracking for paved roads and dirt roads was introduced in [1]. In [2] it was then extended for more complex road geometries and by more robust camera measurements and additional LiDAR measurements. The main concepts of the resulting road tracking framework used for this work are as follows:

- A road is described by its boundary lines.
- One boundary is chosen as the reference lane boundary. It is modeled by equidistantly spaced points connected by clothoid pieces. We choose a point distance of $l = 2\,\mathrm{m}$.
- Based on the reference lane boundary points, points on all other lane boundaries are defined by their orthogonal distance to the reference lane boundary.
- The 6D poses of the reference lane boundary points and the corresponding distances to the other lane boundaries are estimated by recursive state estimation. Therefore, we set up and solve an optimization problem.
- The residuals include internal evaluations of the consistency of the estimated points to the clothoid model, the smoothness of the road, and constraints. In addition, the deviation of the estimated lane boundary points to several different measurements is evaluated.
- The main measurements are performed in a camera image to ensure high accuracy. Using the previous state estimate as prior knowledge, local measurement windows are set up around the projected pixel positions of the estimated lane boundary points. Within these measurement windows, edge lines are extracted using directed edge detection with CRONOS [12], [13].
- Since the distance of the measured edge lines to the camera is unknown, an estimated ground elevation determines the height of the road. We are using a probabilistic terrain estimation based on ground points from LiDAR measurements [14], [15].
- Static obstacles mapped in a LiDAR-based obstacle grid [14] provide additional support. They help to locate the road between street posts, railings, or trees.

Figure 2 shows the complete system architecture including the new extensions to exploit the behavior of dynamic objects measured by radar detections. The following sections give some more details on the road tracking framework.

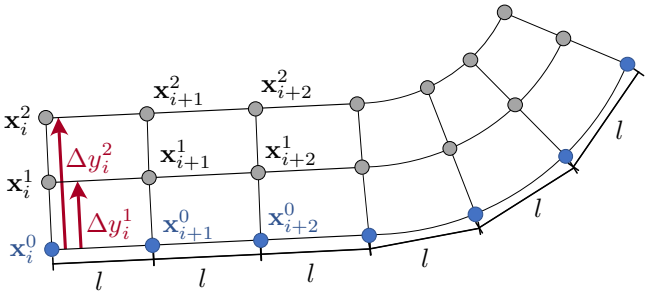


Fig. 3: Exemplary model of a road with $n = 2$ lanes. The road is described by point sequences at the lane boundaries. The lane boundary points are defined by reference lane boundary points (blue) in a constant distance and respective orthogonal offsets to the other lane boundaries (red). The figure was already published in [2].

A. Road Model and State Vector

A road consisting of n lanes and a reference lane boundary of length $m \cdot l$ (meters) is described by its reference lane boundary points $\mathbf{x}_i^0$ with $i \in \{0, \dots, m\}$. The remaining lane boundary points $\mathbf{x}_i^k$ ($k \in \{1, \dots, n\}$) on the other lane boundaries are defined by the offsets $\Delta\mathbf{y}_i = \left[\Delta y_i^1, \dots, \Delta y_i^n\right]$ to the reference lane boundaries. Fig. 3 shows an example.

Following this model, a state vector $\mathbf{x}$ consisting of a vector

$$\mathbf{x}_i = \left[p_x, p_y, p_z, q_x, q_y, q_z, q_w, c_h, c_v, \Delta\mathbf{y}\right]^\mathsf{T} \tag{1}$$

for every reference lane boundary point has to be estimated. ${}^{\text{odom}}\mathbf{p}_i = \left[p_x, p_y, p_z\right]^\mathsf{T}$ and ${}^{\text{odom}}\mathbf{q}_i = \left[q_x, q_y, q_z, q_w\right]^\mathsf{T}$ specify the position and orientation (in quaternions) of the reference lane boundary point in a fixed dead reckoning coordinate system (*odom*). $c_{h,i}$ and $c_{v,i}$ indicate the horizontal and vertical curvatures to the next reference lane boundary point. As the ego vehicle moves forward, new points are added in the front. Points behind the car are excluded from $\mathbf{x}$.

B. State Optimization

To estimate $\mathbf{x}$, the optimization problem

$$\begin{aligned}\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2} \sum_i \Big[&\mathbf{r}^2_{\text{cloth.}}(\mathbf{x}_i) + \mathbf{r}^2_{\text{smooth}}(\mathbf{x}_i) + \mathbf{r}^2_{\text{limits}}(\mathbf{x}_i) + \\ &\mathbf{r}^2_{\text{img}}(\mathbf{x}_i) + \mathbf{r}^2_{\text{terrain}}(\mathbf{x}_i) + \mathbf{r}^2_{\text{obstacles}}(\mathbf{x}_i) + \\ &\mathbf{r}^2_{\text{dyn.}}(\mathbf{x}_i) \Big]\end{aligned} \tag{2}$$

is solved using the Ceres Solver [16].

The first three residuals concern the road geometry. The residual $\mathbf{r}_{\text{cloth.}}$ causes the road estimate to consist of clothoids. It calculates the deviation of two consecutive lane boundary points to the simplified clothoid equation

$$\mathrm{f}(l) = \frac{1}{2} c_i l^2 + \frac{1}{6} \mathring{c}_i l^3 = \frac{1}{2} c_i l^2 + \frac{1}{6} (c_{i+1} - c_i) l^2 \tag{3}$$

in both horizontal and vertical direction, where $\mathring{c}$ is the curvature change. The residual $\mathbf{r}_{\text{smooth}}$ punishes larger jumps in the curvatures and the lane width from one lane boundary point to the next. This way, a natural geometry is ensured. Further, if only one boundary of the road is observable, the

road is continued on the basis of the other lane boundary keeping the same road width. If none of the lane boundaries can be observed, the last curvature is retained. The residual $\mathbf{r}_{\text{limits}}$ limits the possible curvatures (of every lane boundary point, not only the reference points) and lane widths. This excludes impossible parameters.

The second three residuals are based on different kinds of measurements. The residual $\mathbf{r}_{\text{img}}$ takes into account image measurements performed for every lane boundary point individually. Therefore, the deviation of the lane boundary point projected into the camera image to an edge line measured in this camera image is evaluated. Here we consider not only the most recent measurements, but all measurements from the last 3 s. The residual $\mathbf{r}_{\text{terrain}}$ covers the elevation of the road. It minimizes deviations in the z-height p_z and the normal vector of the lane boundary point to a continuous terrain model from a terrain estimation. As the change in the vertical curvature is slightly minimized by $\mathbf{r}_{\text{smooth}}$, the elevation profile might also support finding the road since mostly the road is the most even part of the surrounding. The residual $\mathbf{r}_{\text{obstacles}}$ punishes static obstacles within the road boundaries. Assuming that the road to be tracked is free of static obstacles, obstacles next to the road can be used to delineate the possible road position. Increasing costs with decreasing distance of an obstacle to the road center ensures that the road estimate is moved away from static obstacles. Especially posts or railings right next to the road hence often lead to a correction in the estimated road even before it can be measured in the camera image accurately enough. Lastly, $\mathbf{r}_{\text{dyn.}}$ represents the new residual for taking into account the behavior of dynamic obstacles. It will be introduced in the next chapter. More details on all other residuals can be found in [2]. Note that for shared spaces we do not use the LiDAR range measurements described there.

C. Image Measurements

The image measurements, which form the main measurement, are obtained by locally extracting edges from a single-channel image (refer to [2] on how to increase the contrast compared to a gray image). The edge detection is based on the last state estimate of a lane boundary point. For each lane boundary point, a horizontal and a vertical measurement window are set up around its estimated position, as it can be seen in Fig. 4. In each measurement window, we are looking for edges in the expected direction – derived from the projected position of point estimates – using a rotated filter mask. In case multiple edges exceed a magnitude threshold, we choose the innermost edge. More details on the image measurements can again be found in [2]. This directed, local edge detection is very efficient and allows to locate the road precisely. However, it requires a roughly correct road estimate first. If the estimate is too far off, the road can no longer be measured or it is even measured incorrectly. Thus, additional measurements are required to ensure a rough localization of the road at all times – especially after long occlusion or occlusion in curves as caused by groups of pedestrians, which frequently occur on campus roads.

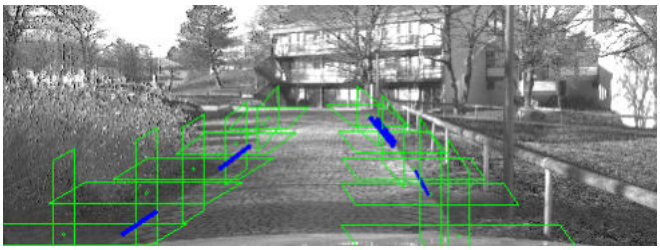


Fig. 4: Example of the lane boundary detection using CRONOS image measurements [12], [13]. The estimated lane boundary points are projected into the image (green points) and used to set up local measurement windows for edges in a certain direction (green boxes). The resulting measured edge lines are drawn in blue.

III. Probabilistic Radar-based Accumulation of Road User Paths

In addition to cameras and LiDAR sensors, our autonomous vehicle is equipped with radar sensors: five short-range sensors (*smartmicro UMRR-96*) around the vehicle and one far-range sensor (*smartmicro UMRR-11*) at the front center. Each radar sensor S provides a list $\mathcal{Y}_{k,\text{S}} = \left[\mathbf{y}_1, \ldots, \mathbf{y}_{N_\text{S}}\right]$ of point measurements $\mathbf{y}$, so-called *detections*, at every time step k. A detection denotes a point of particularly high local reflectivity to radio waves and includes measurements of

- range r_D (with measurement noise σ_r),
- azimuth angle α_D (with measurement noise σ_α), and
- radial speed $\dot{r}_\text{D}$ (with measurement noise $\sigma_{\dot{r}}$).

Range and azimuth angle give an Euclidean point of the detection (in sensor coordinates):

$$\overset{\text{sensor}}{\begin{bmatrix} x_\text{D} \\ y_\text{D} \end{bmatrix}} = \begin{bmatrix} r_\text{D} \cdot \cos(\alpha_\text{D}) \\ r_\text{D} \cdot \sin(\alpha_\text{D}) \end{bmatrix}. \tag{4}$$

Figure 5 shows an accumulation of such detection points over time. It turns out that – unlike with LiDAR points [2] – the positions of the radar detections cannot be used directly to determine the boundaries of an open road. Further, the detections are subject to noise and clutter measurements.

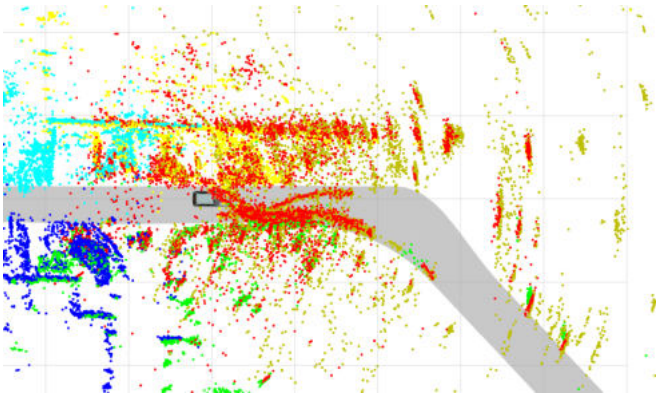


Fig. 5: Radar detections in the scene depicted in Fig. 1 accumulated over 10 s. Detections measured by the far-range sensor at the front center are colored in brown. Detections measured by the short-range sensors are colored in red (front center), yellow (front left), green (front right), light blue (rear left), and dark blue (rear right). The ground truth road is shown in gray.

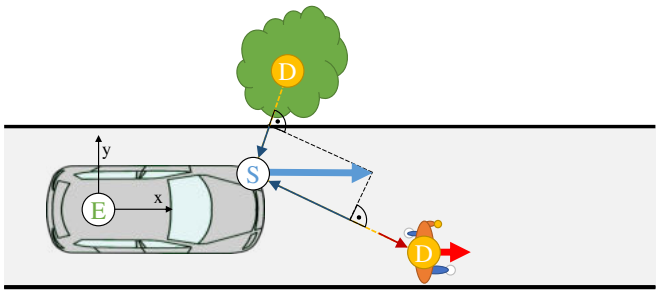


Fig. 6: The velocity vector $\mathbf{v}_S$ of the sensor (blue arrow) is projected into the radial line-of-sight (dark blue arrows) to the detections (yellow). The pedestrian moves along the road (red arrow); its velocity vector invokes an additional radial speed component in the projection to the line-of-sight (dark red arrow) and increases the radial speed measurement accordingly.

However, in addition to the position of a detection, the radar sensors are able to measure the radial speed of a detection towards the sensor based on the Doppler-deduced frequency shift. This way, it is possible to differentiate whether the detection originates from a static or a dynamic object. This allows us to detect pedestrians or other road users on our road without the need for a multi-object tracking.

Since we want to exploit the behavior of other road users to show us where to drive on, we are interested in their paths. Hence, we first identify dynamic objects using radial speed measurements. Therefore, the egomotion of our own vehicle has to be compensated. Then, we obtain a spatial probability map for every detection taking into account the range and angle uncertainties of the detection. Finally, we accumulate the probability maps of all detections from all sensors over time. The result is a grid map containing the probability that a grid cell has ever been used by another road user.

A. Compensating the Egomotion

In order to transform a detection into world coordinates, we utilize an egomotion estimation to obtain the transformation ${}^{\text{odom}}\mathbf{H}_{\text{ego}}(t)$ from the vehicle rear axis to our dead reckoning coordinate system *odom* at measurement time t. Further utilizing the static transformation matrix ${}^{\text{ego}}\mathbf{H}_{\text{sensor}}$ from the radar sensor to the ego vehicle (rear axis), the detection position in world coordinates can be calculated:

$$ {}^{\text{odom}}\begin{bmatrix} x_D \\ y_D \end{bmatrix} = {}^{\text{odom}}\,\mathbf{H}_{\text{ego}}(t) \cdot {}^{\text{ego}}\,\mathbf{H}_{\text{sensor}} \cdot {}^{\text{sensor}}\begin{bmatrix} x_D \\ y_D \end{bmatrix} . \quad (5) $$

For better readability, we will omit the specification of the coordinate system for coordinates and vectors in the *odom* coordinate system from now on.

The radial speed measurement indicates how fast the sensor and a detection are moving towards each other. Since our sensor is moving itself, we need to subtract the influence of the sensor's motion to obtain the actual radial speed of an object in the world. Figure 6 shows the principle following [17]. The velocity vector $\mathbf{v}_S$ of the sensor over ground is computed considering the current speed v_E and yaw rate $\dot{\psi}_E$ of the vehicle as well as the lever arm from the kinematic reference point of the vehicle to the sensor [18]:

$$ \mathbf{v}_S = \begin{bmatrix} \cos\psi_E \cdot v_E - \dot{\psi}_E \cdot (y_S - y_E) \\ \sin\psi_E \cdot v_E + \dot{\psi}_E \cdot (x_S - x_E) \end{bmatrix} . \quad (6) $$

x_E, y_E, and ψ_E denote the current position and yaw angle of the ego vehicle in *odom* coordinates as provided by the egomotion estimation ${}^{\text{odom}}\mathbf{H}_{\text{ego}}(t)$. x_S and y_S denote the current position of the sensor in *odom* coordinates. To obtain the radial component of the velocity vector $\mathbf{v}_S$, it is now projected into the line of sight of the sensor towards the detection. Using the angle

$$ \Theta = \operatorname{atan2}(x_D - x_S, y_D - y_S) , \quad (7) $$

that points from the sensor to the detection in *odom* coordinates, the projection is given by:

$$ \dot{r}_S = \left[\cos\Theta, \sin\Theta\right] \cdot (-\mathbf{v}_S) . \quad (8) $$

The projection of the sensor motion $\dot{r}_S$ now returns the influence of the vehicle motion on the radial speed measurement to a specific detection position. Its comparison with the actual measured radial speed gives the egomotion-compensated radial speed

$$ \Delta\dot{r} = \dot{r}_D - \dot{r}_S . \quad (9) $$

It indicates dynamic objects: If the egomotion-compensated radial speed of a detection is greater or less than zero, the measured object invoked an additional term on the radial speed beyond the egomotion. Hence, the object is likely to be moving. However, a radial speed matching the egomotion does not necessarily imply a static object: the object might also move tangentially to the radar sensor.

B. Identifying Dynamic Objects

Apart from being caused by a dynamic object, a mismatching radial speed measurement can also result from erroneous measurements of static objects. Thus, the calculation of the probability that a specific detection originates from a moving object depends on two further factors. Firstly, even a static object might invoke minor movements, e. g. a bush blowing in the wind. Therefore, we set a threshold of $\dot{r}_{\text{lim}} = \pm 0.4\,\text{m/s}$ to exclude low radial speed mismatches. Secondly, the radial speed measurement is subject to measurement noise, such that this threshold needs to be implemented probabilistically. Figure 7 illustrates the computation of the dynamic object probability. The red line represents an exemplary (egomotion-compensated) radial speed measurement of $\Delta\dot{r} = -0.35\,\text{m/s}$. The blue graph shows the normal distribution $\mathcal{N}(x,\, \mu,\, \sigma^2)$ of the exemplary (exaggerated) measurement uncertainty. It provides the likelihood that a specific radial speed was the true radial speed behind the noisy radial speed measurement. This inversion of the measurement uncertainty is possible for normal distributions. The yellow area denotes the region, in which the true radial speed is considered to belong to a dynamic object. The probability $p_{\text{dyn.}}$ that the true radial speed is inside this region corresponds to its area below the

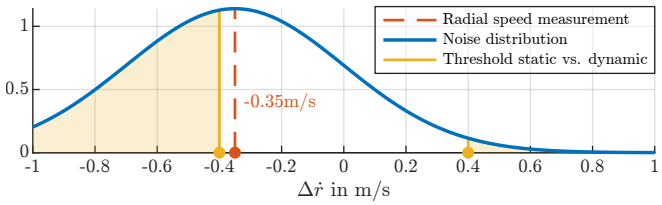


Fig. 7: The computation of the probability that a measured radial speed (dashed red line) originates from a dynamic object. The measurement uncertainty is modeled with a Gaussian distribution (blue line). For illustration purposes, it is exaggerated to $\sigma_{\dot{r}} = 0.35\,\text{m/s}$. The yellow area under the Gaussian distribution gives the probability that the true radial speed exceeds the radial speed thresholds (yellow lines) that separate dynamic from static objects.

normal distribution:

$$\begin{aligned} p_{\text{dyn.}} &= \int_{-\infty}^{\dot{r}_{\text{lim}}^{-}} \mathcal{N}(u,\, \Delta\dot{r},\, \sigma_{\dot{r}}^2)\, \mathrm{d}u + \int_{\dot{r}_{\text{lim}}^{+}}^{+\infty} \mathcal{N}(u,\, \Delta\dot{r},\, \sigma_{\dot{r}}^2)\, \mathrm{d}u \\ &= \int_{-\infty}^{\dot{r}_{\text{lim}}^{-}} \mathcal{N}(u,\, \Delta\dot{r},\, \sigma_{\dot{r}}^2)\, \mathrm{d}u + \left(1 - \int_{-\infty}^{\dot{r}_{\text{lim}}^{+}} \mathcal{N}(u,\, \Delta\dot{r},\, \sigma_{\dot{r}}^2)\, \mathrm{d}u\right) \\ &= \Phi\left(\frac{\dot{r}_{\text{lim}}^{-} - \Delta\dot{r}}{\sigma_{\dot{r}}}\right) + \left(1 - \Phi\left(\frac{\dot{r}_{\text{lim}}^{+} - \Delta\dot{r}}{\sigma_{\dot{r}}}\right)\right), \end{aligned} \tag{10}$$

utilizing the cumulative distribution function $\Phi(\cdot)$

$$\Phi(x) = \frac{1}{2}\left(1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right), \tag{11}$$

which in turn utilizes the error function $\text{erf}(\cdot)$

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-u^2} \mathrm{d}u\,. \tag{12}$$

The dynamic object probability $p_{\text{dyn.}}$ yields $\approx 46\,\%$ in the illustrated example. In reality, we set the radial speed measurement uncertainty to $\sigma_{\dot{r}} = 0.1\,\text{m/s}$. The dynamic object probability yields $\approx 31\,\%$ in this case. With a hard threshold, this object would have been gated out as a static object, although there is still some likelihood for it to be dynamic.

One last aspect to consider is that a specific percentage of all detections are random clutter measurements. Even though its radial speed measurement might indicate a dynamic object, a clutter measurement is not caused by a real object. Hence, in addition to the dynamic object probability, we introduce a clutter probability $p_{\text{clutt.}}$. We determined the clutter probability to be $p_{\text{clutt.}} = 5\,\%$ for our sensors.

C. Obtaining Spatial Probability Maps

Alongside the radial speed measurement, the position measurement is also affected by measurement noise. Thus, in order to calculate probabilities for road areas based on radar detections, we need to take into account their spatial measurement noise. The uncertainty ${}^{\text{sensor}}\mathbf{R}$ of the Euclidean coordinates (in sensor coordinates) is invoked by the polar measurement uncertainties and given by the transformation:

$${}^{\text{sensor}}\mathbf{R} = \mathrm{R}(\alpha_{\mathrm{D}}) \cdot \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & (2r_{\mathrm{D}} \cdot \tan(\sigma_\alpha/2))^2 \end{bmatrix} \cdot \mathrm{R}(\alpha_{\mathrm{D}})^{\mathsf{T}}, \tag{13}$$

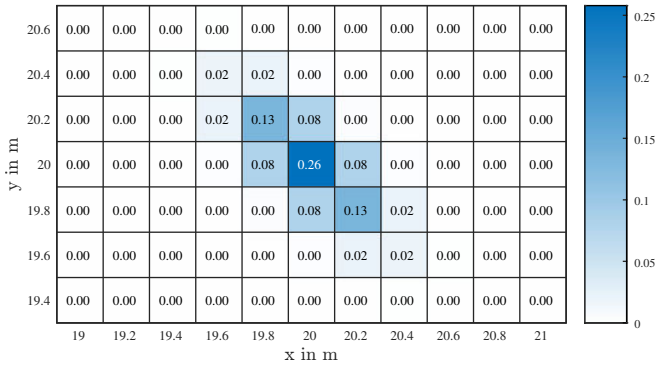


Fig. 8: The spatial probability distribution of a detection in 28 m distance to the sensor. The detection is located at $x_{\mathrm{D}} = 20\,\text{m}, y_{\mathrm{D}} = 20\,\text{m}$, the sensor is located at $x_{\mathrm{S}} = 0\,\text{m}, y_{\mathrm{S}} = 0\,\text{m}$ (bottom left). Each grid cell shows the probability, that the detection originates from a point within this grid cell.

which uses the rotation matrix $\mathrm{R}(\cdot)$:

$$\mathrm{R}(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}. \tag{14}$$

We set $\sigma_r = 0.1\,\text{m}$ and $\sigma_\alpha = 0.5°$. The corresponding uncertainty matrix in *odom* coordinates is computed by

$$\mathbf{R} = \mathrm{R}\left(\psi_{\mathrm{E}}(\mathrm{t}) + {}^{\text{ego}}\,\psi_{\mathrm{S}}\right) \cdot {}^{\text{sensor}}\,\mathbf{R} \cdot \mathrm{R}\left(\psi_{\mathrm{E}}(\mathrm{t}) + {}^{\text{ego}}\,\psi_{\mathrm{S}}\right)^{\mathsf{T}}, \tag{15}$$

using the orientation of the radar sensor ${}^{\text{ego}}\psi_{\mathrm{S}}$ in ego vehicle coordinates (contained in ${}^{\text{ego}}\mathbf{H}_{\text{sensor}}$) and the heading of the vehicle $\psi_{\mathrm{E}}(t)$ in *odom* coordinates (contained in ${}^{\text{odom}}\mathbf{H}_{\text{ego}}(t)$).

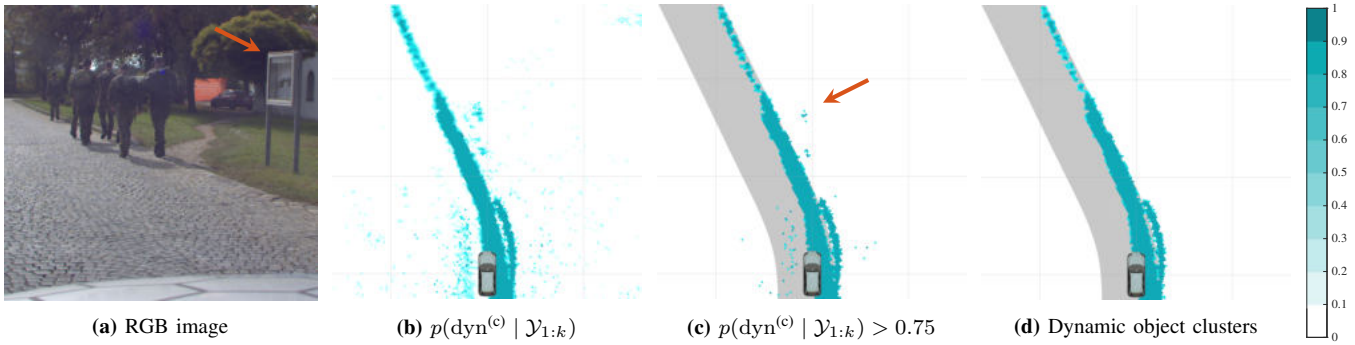
Our goal is to obtain a probabilistic grid map of dynamic objects. In accordance with our static obstacle grid [14], this map is built in stationary *odom* coordinates and divided into grid cells of size $0.2\,\text{m} \times 0.2\,\text{m}$. Taking into account measurement uncertainties, a detection is inserted into the map by calculating the probability $p(\text{occ}^{(c)} \mid \mathbf{y}_i)$ for each grid cell c that it is occupied by the detection $\mathbf{y}_i$. Using the probability density function of the multivariate normal distribution, the likelihood of a grid cell c located at position $\left[x_{\mathrm{C}}, y_{\mathrm{C}}\right]^{\mathsf{T}}$ to be occupied by a detection located at $\left[x_{\mathrm{D}}, y_{\mathrm{D}}\right]^{\mathsf{T}}$ with measurement uncertainty $\mathbf{R}$ is given by:

$$\begin{aligned} p'(\text{occ}^{(c)} \mid \mathbf{y}_i) &= \mathcal{N}\left(x = \begin{bmatrix} x_{\mathrm{C}} \\ y_{\mathrm{C}} \end{bmatrix}, \mu = \begin{bmatrix} x_{\mathrm{D}} \\ y_{\mathrm{D}} \end{bmatrix}, \sigma^2 = \mathbf{R}\right) \\ &= \frac{\exp\left(-\frac{1}{2}\left(\begin{bmatrix} x_{\mathrm{C}} \\ y_{\mathrm{C}} \end{bmatrix} - \begin{bmatrix} x_{\mathrm{D}} \\ y_{\mathrm{D}} \end{bmatrix}\right)^{\mathsf{T}} \mathbf{R}^{-1} \left(\begin{bmatrix} x_{\mathrm{C}} \\ y_{\mathrm{C}} \end{bmatrix} - \begin{bmatrix} x_{\mathrm{D}} \\ y_{\mathrm{D}} \end{bmatrix}\right)\right)}{\sqrt{(2\pi)^2 \det(\mathbf{R})}}. \end{aligned} \tag{16}$$

An approximation of the probability $p(\text{occ}^{(c)} \mid \mathbf{y}_i)$ then results from normalizing the likelihoods of all grid cells $\mathcal{C}$:

$$p(\text{occ}^{(c)} \mid \mathbf{y}_i) = \frac{p'(\text{occ}^{(c)} \mid \mathbf{y}_i)}{\sum_{c' \in \mathcal{C}} p'(\text{occ}^{(c')} \mid \mathbf{y}_i)}. \tag{17}$$

Figure 8 visualizes the probability map of an exemplary detection in 28 m distance to the sensor. Especially the high angular uncertainty of a radar sensor leads to a large spread of the spatial probabilities.



(a) RGB image **(b)** $p(\mathrm{dyn}^{(c)} \mid \mathcal{Y}_{1:k})$ **(c)** $p(\mathrm{dyn}^{(c)} \mid \mathcal{Y}_{1:k}) > 0.75$ **(d)** Dynamic object clusters

Fig. 9: Road user probability map and dynamic object clusters used for road tracking in an example scene. (**a**) RGB image of the scene. (**b**) Accumulated road user probabilities, color coded from white for low probabilities to cyan to petrol for high probabilities. Pedestrians have left a thick trace, but there are also some clutter measurements. (**c**) Accumulated road user probabilities above a threshold of 0.75. Some of the clutter measurements have insufficient probability. However, one large wrong blob remains: A reflection of the pedestrians from a mirroring sign (orange arrows). (**d**) This blob is removed along with the remaining clutter grid cells by a clustering algorithm.

D. Accumulating Road User Probabilities

Finally, from the probability $p_{\text{dyn.}}$ that a detection was caused by a dynamic object, its spatial probability distribution $p(\mathrm{occ}^{(c)} \mid \mathbf{y}_i)$, and the clutter probability $p_{\text{clutt.}}$, we can now determine the probability $p(\mathrm{dyn}^{(c)} \mid \mathbf{y}_i)$ that inside a grid cell c a road user is measured given the detection $\mathbf{y}_i$:

$$p(\mathrm{dyn}^{(c)} \mid \mathbf{y}_i) = p_{\text{dyn.}} \cdot p(\mathrm{occ}^{(c)} \mid \mathbf{y}_i) \cdot (1 - p_{\text{clutt.}}) \,. \tag{18}$$

Given all detections $\mathcal{Y}_k$ of all sensors S at time step k, the probability that at least once a dynamic object was measured within a grid cell c is calculated using the product of all statistically independent counter-probabilities:

$$p(\mathrm{dyn}^{(c)} \mid \mathcal{Y}_k) = 1 - \left(\prod_{\mathrm{S}} \prod_{i \in [1; N_{\mathrm{S}}]} \left(1 - p(\mathrm{dyn}^{(c)} \mid \mathbf{y}_i)\right) \right) . \tag{19}$$

By recursively accumulating all time steps in the same manner, we finally obtain the probability that a dynamic object was ever measured in a grid cell c:

$$p(\mathrm{dyn}^{(c)} \mid \mathcal{Y}_{1:k}) = 1 - \left(\left(1 - p(\mathrm{dyn}^{(c)} \mid \mathcal{Y}_{1:k-1})\right) \cdot \left(1 - p(\mathrm{dyn}^{(c)} \mid \mathcal{Y}_k)\right) \right) . \tag{20}$$

An example for the resulting accumulated probability map is shown in Fig. 9b.

IV. Integration of a Road User Probability Map into Road Tracking

To integrate the road user paths into our road tracking framework, we need to extract discrete objects from the probability map. If we take all grid cells with a road user probability $p(\mathrm{dyn}^{(c)} \mid \mathcal{Y}_{1:k})$ above 75 %, apart from the road user traces there are still some false positives. As it can be seen in the example in Fig. 9, these include clutter measurements and reflections. Especially the reflections may cause a relatively high probability, as they originate from real dynamic objects. To filter these out, we would need to choose a very high threshold of up to 90 %, but this would result in a loss of lookahead: At larger distances, the measurement uncertainty is extended over several grid cells, which leads to lower probability values for the individual grid cells, as it was shown in Fig. 8. However, the false positives are sparse and can be easily separated from the dense pedestrian traces. Thus, for road tracking, we perform a clustering on the probability map binarized with a threshold of 75 %. Using the Spaghetti clustering algorithm [19] as implemented by OpenCV [20], clusters of grid cells connected by an 8-way connectivity are extracted. To remove false positives, we then omit all clusters consisting of less than 10 grid cells. While true positives will form larger clusters over time due to their movement, false positives will remain at the same point and can be removed this way. Figure 9d shows the resulting grid map for the example scene. Another possibility to identify false positives would be to adapt the approach introduced in [21] that considers the position of detections relative to each other to derive free space evidences.

Since we assume road users are moving on the road, we want our road estimate to include the grid cells used by dynamic objects. Therefore, we introduce a cost for each grid cell that was used by dynamic objects according to the clustering step, which is outside of the current road estimate.

Figure 10 shows the principle, which is carried out for each road section individually: For every reference lane boundary point $\mathbf{x}_i^0$, all dynamic object cell positions O in between the straight line between $\mathbf{x}_i^0$ and the other lane boundary

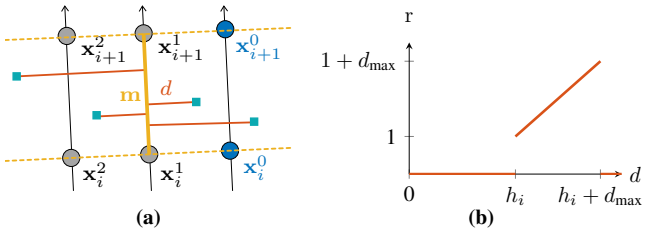


Fig. 10: Schematic example of our dynamic object residuum. (**a**) For all dynamic object cells (cyan) within a road segment (dashed yellow), the distance d (red) to the road center line $\mathbf{m}$ (solid yellow) is evaluated. (**b**) The cost caused by dynamic object cells is a function $\mathrm{r}(d)$ of their distance from the road center line.

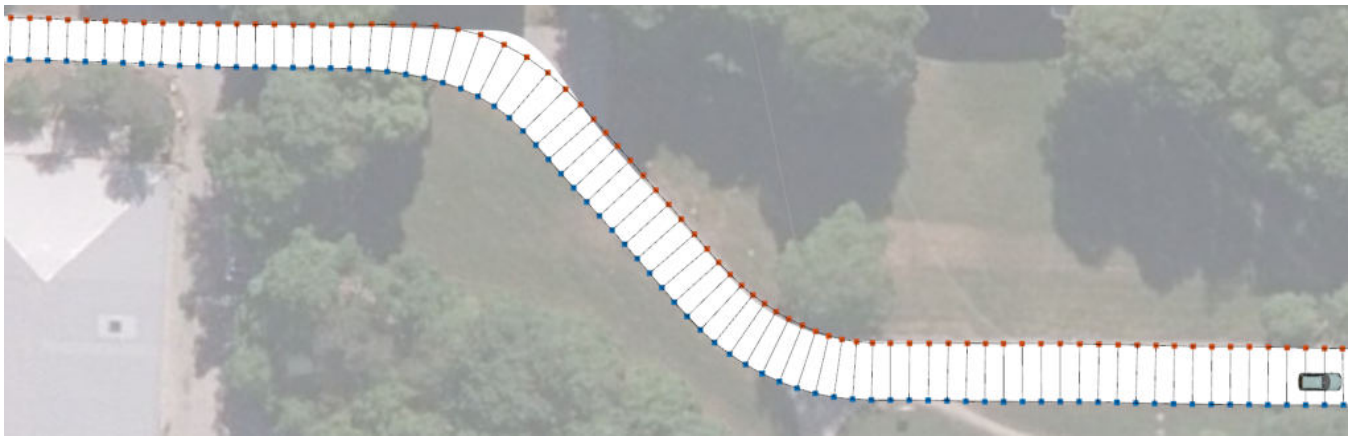


Fig. 11: Right (blue) and left (red) lane boundaries of our road estimate in comparison with the ground truth from an HD map (white).

point $\mathbf{x}_i^n$ and the straight line between the next lane boundary points $\mathbf{x}_{i+1}^0$ and $\mathbf{x}_{i+1}^n$ are collected. For each dynamic object cell, its distance $d = \mathrm{d}(O, \mathbf{m})$ to the centerline $\mathbf{m}$ of the road segment is calculated. Based on this distance, the cost function

$$\mathrm{r}(d) = \begin{cases} 0 & d \leq h_i \\ 1 + d - h_i & d > h_i \wedge d \leq d_{\text{max}} \\ 0 & d > d_{\text{max}} \end{cases} \tag{21}$$

is evaluated, where

$$h_i = \frac{1}{2} \sum_k \Delta y_i^k \tag{22}$$

is half the width of the road. Dynamic object cells with a distance less than (or equal to) the half road width are within the road estimate and therefore do not cause any cost. Dynamic object cells outside the road estimate are penalized with a cost that increases with the distance of the cell from the centerline to provide a gradient for the optimization algorithm. Cells with a distance greater than a maximum distance $d_{\text{max}} = 2\,\text{m}$ are in turn ignored. This is necessary to ensure robustness at junctions or for example when pedestrians take shortcuts through the fields. It also reduces the runtime.

The sum of the costs of all dynamic object grid cells O finally yields the dynamic object residual

$$\mathbf{r}_{\text{dyn.}}^2(\mathbf{x}_i) = \rho\left(\sum_O \mathrm{r}\left(\mathrm{d}(O, \mathbf{m}) \right) \cdot w \right), \tag{23}$$

where $w = 10$ is a weighting factor and

$$\rho(s) = \begin{cases} s^2 & s \leq 1 \\ 2s - 1 & s > 1 \end{cases} \tag{24}$$

is the robust Huber loss function [22]. The Huber loss function grows less rapidly than the quadratic loss function commonly used by the Ceres solver [16] and therefore provides additional robustness to outliers.

In order to obtain the road estimate $\hat{\mathbf{x}}$, the overall cost of all different road tracking residuals $\mathbf{r}(\mathbf{x}_i)$ for all reference lane boundary points $\mathbf{x}_i$ is minimized following Eq. (2).

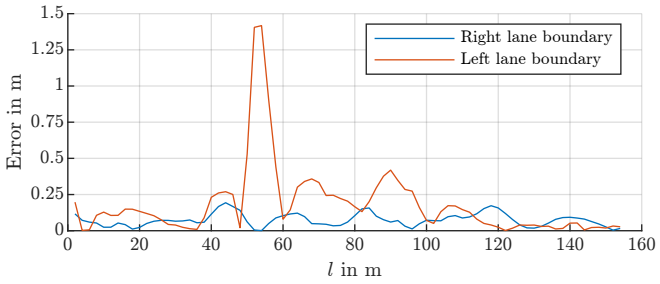


Fig. 12: Deviation of our estimated lane boundary points from the ground truth lane boundaries along the driven distance l.

V. Evaluation

To evaluate the performance of our road tracking pipeline with the new dynamic object residuum, we utilize a commercial HD map from 3D-Mapping Solutions GmbH [23] as ground truth. The map covers mainly larger roads and open space on the campus of the University of the Bundeswehr Munich, but also includes one short section of a traffic-calmed campus road. This road section is frequently used by pedestrians, such that our original road tracking measurements [1], [2] are not sufficient to track the occluded road – without the new features introduced in this paper the tracking fails at the point shown in Fig. 1. However, exploiting the pedestrian behavior, this road section is now successfully tracked from start to end.

Figure 11 shows the resulting road estimate. The right lane boundary, which is essential for vehicle control, is tracked precisely along the entire length. The left lane boundary deviates slightly further from the ground truth. On the one hand, this is because the left-hand side of the road is more difficult to observe due to the worse viewing angle. On the other hand, there is one large outlier in the first bend. Since our road model does not allow sharp changes in the road curvature, the corner is rounded off in the estimate. This can be seen in a peak in the absolute error curve shown in Fig. 12 after a third of the way. The mean absolute error (MAE) and the root-mean-square error (RMSE) are given in Table I. The estimate of the right lane boundary has an average accuracy (RMSE) of 9 cm. This is at the accuracy

TABLE I: Mean absolute error (MAE) and root-mean-square error (RMSE) of our road tracking for the campus road shown in Fig. 11.

	right boundary	left boundary	left without first bend
MAE	0.07 m	0.18 m	0.13 m
RMSE	0.09 m	0.31 m	0.17 m

limit of our ground truth evaluation: Both the evaluation of the road tracking and the generation of the ground truth are subject to localization and sensor calibration errors. Omitting the first bend that does not match the road model, the left lane boundary estimate achieves an RMSE of 17 cm. This accuracy results from the fusion of our measurement approach with precise image measurements. For comparison: On their respective datasets, [5] achieves a total RMSE of 0.5 m using vehicle tracking only, [10] gets a total RMSE of more than 1 m by detecting road barriers using radar.

The average runtime of our road tracking is 10 ms. Building the dynamic object probability map takes additional 2 ms. The clustering of this probability map takes another 4 ms on average. Hence, our approach is real-time capable. All timing measurements were conducted using a single core of an Intel© Core™ i7-8700K CPU running at 3.70 GHz.

VI. Conclusion and Outlook

In summary, we present a concept to use the behavior of road users for road tracking. This is particularly beneficial if the road is obscured by other road users and direct measurement methods fail. Our approach can take advantage of the other road users in this case. We detect dynamic objects using radar detections and favor road estimates that cover the accumulated paths of other traffic participants. This increases the robustness of the road tracking system against occlusion.

Going further, there could be drawn even more information from the other road users: The radial speed measurement also indicates their walking/driving direction. As it can be seen in Fig. 13, most pedestrians keep right. This could be exploited for a more targeted localization of the road. However, this is also more prone to error if the pedestrians do not keep right.

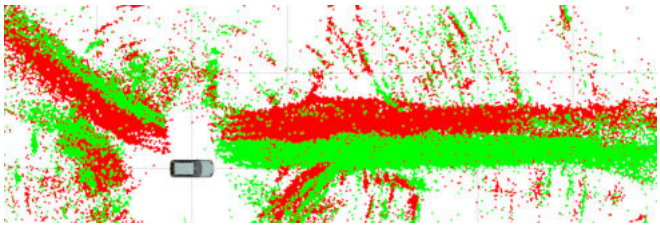


Fig. 13: Accumulated radar detections on a walkway, colored by the sign of their radial speed measurement. Green detections move away from our stationary ego vehicle, red detections are coming closer. Most pedestrians walk on their right half of the road.

Acknowledgement

This work is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr [project MORE] and the Federal Office of Bundeswehr Equipment, Information Technology and In-Service Support (BAAINBw). dtec.bw is funded by the European Union – NextGenerationEU.

References

[1] B. Forkel, J. Kallwies, and H.-J. Wuensche, "Combined Road Tracking for Paved Roads and Dirt Roads: Framework and Image Measurements," in *Proceedings of IEEE Intelligent Vehicles Symposium (IV)*, 2021, pp. 1326–1331.

[2] B. Forkel and H.-J. Wuensche, "Combined Road Tracking for Paved Roads and Dirt Roads: LiDAR Measurements and Image Color Modes," in *Proceedings of International Conference on Information Fusion (FUSION)*, 2022.

[3] W. Yang, T. Ai, and W. Lu, "A Method for Extracting Road Boundary Information from Crowdsourcing Vehicle GPS Trajectories," *Sensors*, vol. 18, no. 4, 2018.

[4] B. Chen, C. Ding, W. Ren, and G. Xu, "Automatically Tracking Road Centerlines from Low-Frequency GPS Trajectory Data," *ISPRS International Journal of Geo-Information*, vol. 10, no. 3, 2021.

[5] Y.-C. Zhang, "Road geometry estimation using vehicle trails: a linear mixed model approach," *Journal of Intelligent Transportation Systems*, vol. 27, no. 1, pp. 127–144, 2023.

[6] J. Wang, C. Li, T. Fu, L. Zhang, A. Sobhani, and J. Xue, "Automated Lane-Level Road Geometry Estimation Using Microscopic Trajectory Data," *Journal of Advanced Transportation*, 2023.

[7] Z. Feng, M. Li, M. Stolz, M. Kunert, and W. Wiesbeck, "Lane Detection With a High-Resolution Automotive Radar by Introducing a New Type of Road Marking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 7, pp. 2430–2447, 2019.

[8] T. J. Douglas and K. Sarabandi, "A Low-Profile Passive Radar Reflector for Detection and Identification of Road Markings," *IEEE Transactions on Antennas and Propagation*, 2023.

[9] T. Grebner, L. L. T. Torres, V. Janoudi, P. Schoeder, and C. Waldschmidt, "Detection of Optical Road Markings With Automotive Radar Sensors," *IEEE Sensors Letters*, vol. 7, no. 8, 2023.

[10] T. Kim and B. Song, "Detection and Tracking of Road Barrier Based on Radar and Vision Sensor Fusion," *Journal of Sensors*, 2016.

[11] F.-X. Li and Y. Zhu, "An Road Boundary Detection Algorithm Based on Radar that Can Improve Multiple-Target Tracking Performance for Autonomous Vehicles on Highway Condition," in *SAE Intelligent and Connected Vehicles Symposium*, 2023.

[12] B. Mysliwetz, "Parallelrechner-basierte Bildfolgen-Interpretation zur autonomen Fahrzeugführung," Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechik, Neubiberg, 1990.

[13] E. D. Dickmanns, *Dynamic Vision for Perception and Control of Motion*. Springer Verlag, 2007.

[14] B. Forkel, J. Kallwies, and H.-J. Wuensche, "Probabilistic Terrain Estimation for Autonomous Off-Road Driving," in *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 13 864–13 870.

[15] B. Forkel and H.-J. Wuensche, "Dynamic Resolution Terrain Estimation for Autonomous (Dirt) Road Driving Fusing LiDAR and Vision," in *Proceedings of IEEE Intelligent Vehicles Symposium (IV)*, 2022, pp. 1181–1187.

[16] S. Agarwal, K. Mierle, and Others, "Ceres Solver," http://ceres-solver.org, 2017.

[17] P. Berthold, M. Michaelis, T. Luettel, D. Meissner, and H.-J. Wuensche, "Probabilistic Vehicle Tracking with Sparse Radar Detection Measurements," *ISIF Journal of Advances in Information Fusion*, pp. 116–139, 2022.

[18] D. Kellner, M. Barjenbruch, K. Dietmayer, J. Klappstein, and J. Dickmann, "Instantaneous Lateral Velocity Estimation of a Vehicle using Doppler Radar," in *Proceedings of International Conference on Information Fusion (FUSION)*, 2013, pp. 877–884.

[19] F. Bolelli, S. Allegretti, L. Baraldi, and C. Grana, "Spaghetti Labeling: Directed Acyclic Graphs for Block-Based Connected Components Labeling," *IEEE Transactions on Image Processing*, vol. 29, pp. 1999–2012, 2020.

[20] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[21] P. Berthold, M. Michaelis, T. Luettel, D. Meissner, and H.-J. Wuensche, "Deriving Spatial Occupancy Evidence from Radar Detection Data," in *Proceedings of IEEE Intelligent Vehicles Symposium (IV)*, 2020, pp. 831–836.

[22] P. J. Huber, "Robust Estimation of a Location Parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73—101, 1964.

[23] 3D Mapping Solutions GmbH, "Ultra HD Maps." [Online]. Available: https://www.3d-mapping.de/en/our-service/ultra-hd-maps/