The Ko-PER Intersection Laserscanner and Video Dataset

Elias Strigel, Daniel Meissner, Florian Seeliger, Benjamin Wilking, and Klaus Dietmayer

***Abstract*— Public intersections are due to their complexity challenging locations for drivers. Therefore the german joint project Ko-PER - which is part of the project initiative Ko-FAS has equipped a public intersection with several laserscanners and video cameras to generate a comprehensive dynamic model of the ongoing traffic. Results of the intersection perception can be communicated to equipped vehicles by wireless communication. This contribution wants to share a dataset of the Ko-PER intersection to the research community for further research in the field of multi-object detection and tracking. Therefor the dataset consists of sensordata from the laserscanners network and cameras as well as reference data and object labels. With that dataset, we aim to stimulate further research in this area.**

I. Introduction

Intersections are accident black spots. Therefore driver assistance is needed in this areas. The aim of this publication is to introduce a laserscanner and video camera dataset gathered at a public urban intersection to the ITS community to promote further research in the field of road user detection, classification, and tracking. Additionally the dataset provides reference data facilitating the evaluation and benchmarking of algorithms. Since the intersection perception system (ISP) has been designed and installed within the joint project Ko-PER [1], the development of algorithms to perceive the road users at the intersection was one major aim of the project. In [2] and [3] the video and laserscanner based object recognition and tracking algorithms including evaluation with reference data are presented. A 3D model of the public intersection can be seen in Fig. 1.

In the published dataset, each object causes multiple laserscanner measurements and is represented by multiple pixels in the camera images. Therefore, as in [3], the data set is highly suitable to develop and evaluate tracking algorithms which do not require a point target assumption. Recently, these extended object tracking algorithms, like such proposed by [4], [5], [6], [7] attracted a lot of attention in the multi-object tracking community. The sufficiency of the presented intersection perception system to estimate the number and states of extended objects has already been shown in [8]. Here a gamma-Gaussian-inverse Wishart (GGIW) PHD filter is used to track pedestrians, bikes, and vehicles on their way through the intersection.

Further, to the knowledge of the authors the Ko-PER intersection dataset is the first dataset containing temporal and spatial calibrated laserscanner and camera data of a permanently installed intersection perception system.

E. Strigel, D. Meissner, F. Seeliger, B.Wilking, K. Dietmayer are with the Institute of Measurement, Control, and Microtechnology, Ulm University, Ulm, Germany, elias.strigel@uni-ulm.de, daniel.meissner@uni-ulm.de

Fig. 1. Public intersection in Aschaffenburg (Germany) used for the test system (Picture is kindly provided by the Wuerzburg Institut for Traffic Sciences GmbH, http://www.wivw.de).

II. Related Work

Several intersection perception systems (IPS) were developed in research projects addressing intersection collision avoidance applications in the recent years. In the United States of America, the Cooperative Intersection Collision Avoidance Systems-Stop Sign Assist (CICAS-SSA) program, made use of radar sensors and laserscanners to acquire road user data at rural intersections, [9], [10]. The used sensors were mounted at street level. In Europe the SafeSpot subproject INFRASENS involved laserscanners, cameras and RFID-systems to detect road users in urban areas, [11]. The laserscanners were mounted at street level, too, which made the IPS prone to occlusions caused by nearby passing pedestrians. This drawback was solved in Intersafe2 by mounting laserscanners on higher top-view-positions, [12]. This concept was adopted within the Ko-PER project. Due to the changed mounting position the appearance of the data changed, as well. This made the development of new object extraction algorithms necessary, [13]. In comparison to Intersafe2, Ko-PER extended the IPS by low- and high resolution cameras, to gather further classification information and information about vulnerable road users, e.g a pedestrian intending to cross the street, [14], [3].

Within the recent research programes IPSs are used to provide a solid information base for intersection collision avoidance systems, see e.g. [9], [12], [15]. In addition IPSs are very useful to gather sufficient naturalistic driving data for parameter determination in the development process of intersection sited driver assistance applications. Using an IPS for Cooperative Awareness in combination with car-to-X (C2X) communication and localization techniques, solves the availability problem of sufficient communicated information at equipped intersections.

III. Test Site and Sensor Setup

This section briefly summarizes the sensor setup installed at a public intersection in Aschaffenburg, Germany to perceive the intersection scene. A detailed description of the intersection perception system is given in [16]. The intersection is a four-way crossing and illustrated in Fig. 2. Its main road features two straight ahead lanes and a separate left-turn lane for each direction. The branch roads have one lane per direction and a left-turn lane on one side. Additionally, the main road has a separate bicycle lane and the intersection is surrounded by sidewalks on all except one side.

The intersection is observed by 14 SICK LD-MRS 8-layer research laserscanners and eight monochrome CCD cameras (Baumer TXG-04) with different viewpoints. The sensors are installed at infrastructure components like lamp posts and traffic lights and are mounted at least 5m above the ground. A sketch of the mounting positions as well as the simulated field of view (FOV) of the sensors [17] is given by Fig. 2 and Fig. 6. All sensors are triggered in hardware thus, each measurement is associated with a timestamp which corresponds to the acquisition time of the measurement (UTC).

A. Laserscanners

Four laserscanners cover the central intersection widespreaded (see Fig. 2(a)), two scanners observe the sidewalks along the main road (see Fig. 2(c)), and eight sensors observe three egresses of the intersection (see Fig. 2(b), 2(d), and 2(e)). The laserscanners synchronously operate with a frequency of 12.5 Hz. Since they scan their environment within 80 ms, not all measurements are acquired to the same time. This has to be considered by using the provided timestamps. An example of a measurement gathered by the laserscanner network is given by Fig. 3. Being able to refer to the measurement geometrically, a highly accurate map is illustrated in the background of Fig. 3. The map is provided in the form of a *Matlab* `Figure` in the east, north, up (ENU) coordinate system and comprises the $[x, y]$ position of lane markings and street boundaries.

B. Cameras

The cameras are monochrome cameras with a resolution of 656×494 pixels and a Pentax H416 lens with focal length of 4.2 mm. To fulfill data protection restrictions, only two of the eight cameras are included in the dataset. For these two cameras it is guaranteed that no personal data is gathered. Example images provided by the two cameras are shown in Fig. 4 and 5.

Fig. 6 shows the cameras' mounting position and a simulation of their FOV. The optical axis form an angle of approx.

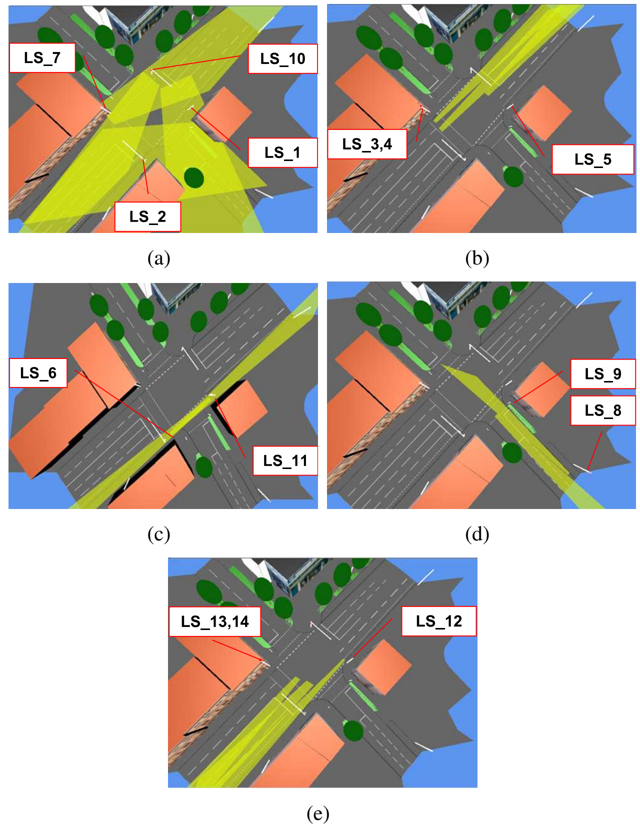


Fig. 2. Mounting positions of laserscanners and their simulated FOV.

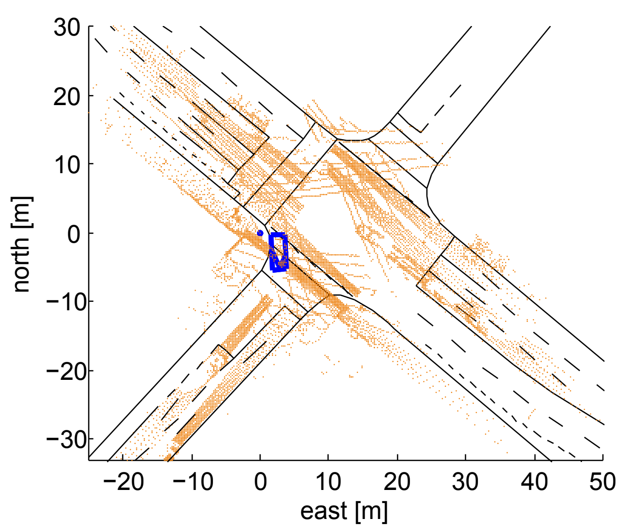


Fig. 3. Aligned lasercanner data, reference data of a vehicle, and highly accurate map of intersection.

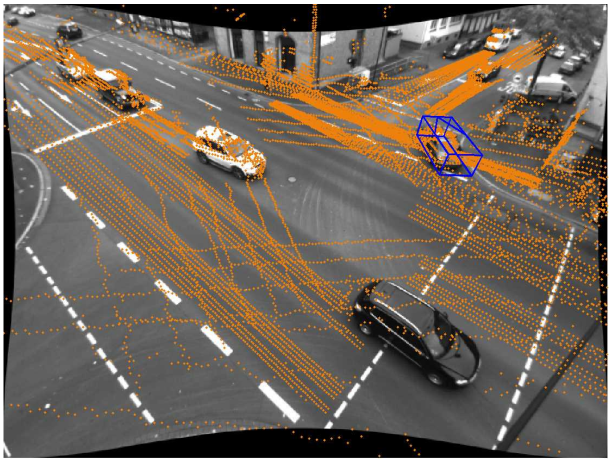


Fig. 4. Image of SK_1 and projected laserscanner measurements.

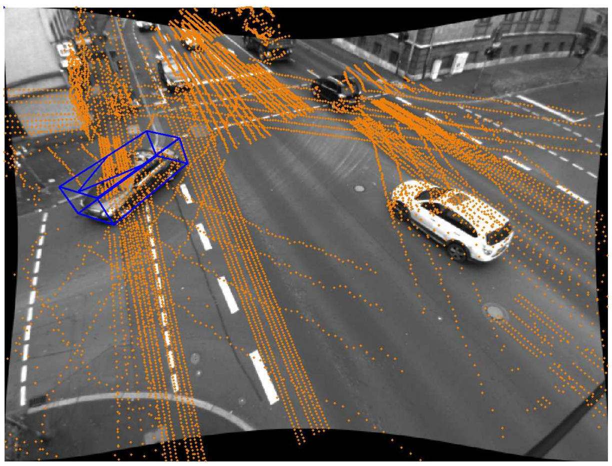


Fig. 5. Image of SK_4 and projected laserscanner measurements.

180 degrees to reduce the risk of occlusions. Throughout the paper and the dataset the cameras are named as in Fig. 6. The operation frequency of the cameras is 25 Hz in phase with the laserscanners.

IV. Dataset

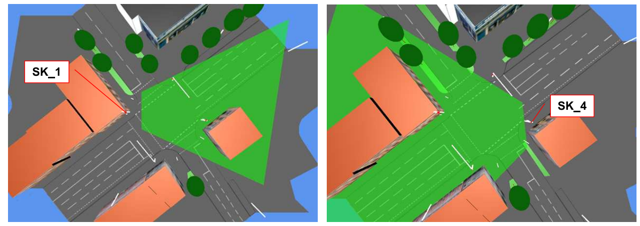
The dataset can be downloaded from `www.uni-ulm.de/in/mrm/forschung/datensaetze.html` and features the content below. All the raw data as well as the calibration parameters are stored in `.mat` files for processing with *Matlab*.

A. Content

The dataset comprises:

- raw laserscanner data
- undistorted camera images
- reference data of selected vehicles
- object labels

While the object labels are provided for `Sequence1` with a duration of 6:28 minutes, reference data for two cars



(a) (b)

Fig. 6. Mounting positions of cameras and their simulated FOV.

performing a right turn and a straight ahead maneuver is included in `Sequence2` and `Sequence3`. Laserscanner measurements and camera images are provided for all sequences.

The `Sequence1` has been recorded in the afternoon and contains several hundred road users. The majority of objects are cars but also trucks, buses, two-wheelers, and pedestrians pass the intersection. Using the provided object labels, this sequence is sufficient to test, evaluate, and benchmark multi-object detection, classification, and tracking methods. Due to the huge number of objects which differ in their type, extend, shape, and texture the sequence poses a challenge for state of the art perception methods. To reduce the size of the files the sequence has been splitted into four parts (`1a, ..., 1d`) with equal duration.

For `Sequence2` and `Sequence3` highly accurate reference data of one car in each case is available. It includes ground truth data for the car's state and can be used to evaluate the absolute estimation error.

The data to each sequence is stored in a folder which contains a `.mat` file for each sensor type. The data in the `.mat` files is structured as shown in Tab. I. Each

TABLE I
Sensor Data streams

```
result =
    stream1: [1x1 struct]
    stream2: [1x1 struct]
    ...
```

available stream contains a `time` (timestamps), `data`, and `converterMethodName` field which specifies the data structure (see Tab. II). To each timestamp a source dependent

TABLE II
Content of Sensor Data Stream

```
result.stream1 =
    time: [1x9674 double]
    data: {1x9674 cell}
    convertMethodName: '<CONVERTER>'
```

`data` element is provided. The content of `data` is described in the subsequent sections.

B. Laserscanner Data

In case of the point measurements of the laserscanners each element of the `data` structure (Tab. II) contains the high level measurement information for one timestep. Tab. III introduces the relevant data fields.

TABLE III

HIGH LEVEL INFORMATION OF LASERSCANNER DATA

```
result.LS_1.data{1} =
    numPoints: 1060
      number of measurement points in current scan
    availableFeatures: {10x1 cell}
      available features for each measurement point
    features: [1060x10 double]
      matrix of feature values with
      dimension (numPoints x numFeatures)
```

The `availableFeatures` for laserscanner measurements are summarized in Tab. IV. The described data is given in sensor coordinates. To transform the laserscanner coordinate system $C_{ls,i}$ to a common coordinate system C_w for all sensors at the intersection, a homogeneous transformation matrix $T^w_{ls,i}$ is provided for each laserscanner i. The system C_w is a ENU-system, where the x-axis points to east, the y-axis to north, and the z-axis upwards.

TABLE IV

CONTENT OF LASERSCANNER MEASUREMENT POINTS

```
result.LS_1.data{1}.availableFeatures =
    FEAT_X_POS x position [m]
    FEAT_Y_POS y position [m]
    FEAT_Z_POS z position [m]
    FEAT_RADIAL_DIST radial distance [m]
    FEAT_AZIMUTH_ANGLE azimuth angle [rad]
    FEAT_ELEVATION_ANGLE elevation angle [rad]
    FEAT_LAYER layer ID = 0,...,7
    FEAT_ECHO_NUM echo ID = 0,...,2 per laser pulse
```

Using $T^w_{ls,i}$ the $pos_{ls,i} = [x, y, z]^T$ measurements of each scanner $i = 1, \ldots, 14$ can be transformed to C_w which facilitates the determination and visualisation of the measurement point cloud of the laserscanner system plotted in Fig. 3.

$$pos_{w,i} = T^w_{ls,i} \cdot pos_{ls,i} \tag{1}$$

Since additionally the extrinsic and intrinsic calibration of the cameras to C_c is known, the laserscanner measurements can be projected into the camera images (see eg. Fig. 4). Example code to load, transform and visualize the scanner data can be found in the code of the provided `DataSetViewer`.

C. Camera Data

For each camera the undistorted images are stored in a folder named after the corresponding camera data stream (Tab. I). By opening an element of the `data` cell (see Tab. V), the name of the image can be accessed.

TABLE V

HIGH LEVEL INFORMATION OF CAMERA DATA

```
result.KAB_SK_1.data{1} =
    image: 'KAB_SK_1_1384779301359985.bmp'
    labels: []
    source: []
```

Similar to the laserscanners, the transformation of the world coordinate system C_w to the i^{th} camera coordinate system $C_{ca,i}$ is given by the equation:

$$pos_{ca,i} = T^{ca}_{w,i} \cdot pos_{w,i} \tag{2}$$

with the homogeneous transformation matrix $T^w_{ca,i}$ which is stored in the `Cam_Calib.mat` in the `calib` folder of the `DataSetViewer`. This file also contains the intrinsic calibration parameters for the cameras. The principal axis of the camera coordinate system is pointing down the z-axis. After transforming the laserscanner measurements into the camera coordinate system they can be normalized to their z-compontens. With this intrinsic calibration parameters, the projection of the normalized measurements is given by:

$$\begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} \tag{3}$$

The focal lenght and the principal point is denoted by f_x, f_y, c_x and c_y. Example code to load, transform, and visualize the images is also provided within the `DataSetViewer`.

D. Object Labels

For `Sequence1` a set of object labels is provided. This is generated by manually inspecting the sensor data of each sensor for each time step, including map information. During the labeling process, a box was placed around each object in different frames. The position and dimensions were adjusted by considering all available information: each camera view, the laser information and the map. Between two labeled frames, the poses of the objects were interpolated. Each labeled object includes a unique track id and an object class. Thus, the labeled data is sufficient to evaluate environment perception methods including multi-object tracking and classification algorithms. In Fig. 7, 8 and 9 one frame of labeled data is shown. The different label classes are shown in different colors: Blue boxes represent cars, green boxes represent pedestrians, yellow boxes represent trucks, and black boxes represent bikes.

Similar to the laserscanner measurements, one position measurement is provided to each timestamp in `time` (Tab. III). The `availableFeatures` are shown in Tab. VII.

The content of the labeled `Sequence1` is summarized in Tab. VI.

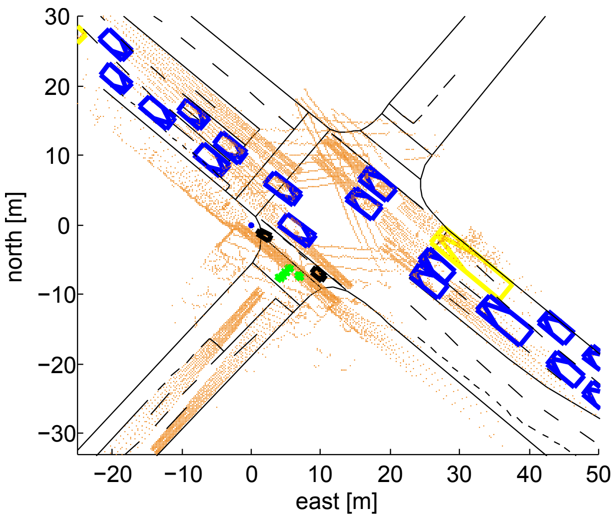


Fig. 7. Laserscanner data and object labels. Blue boxes represent cars, green boxes represent pedestrians, yellow boxes represent trucks, and black boxes represent bikes.

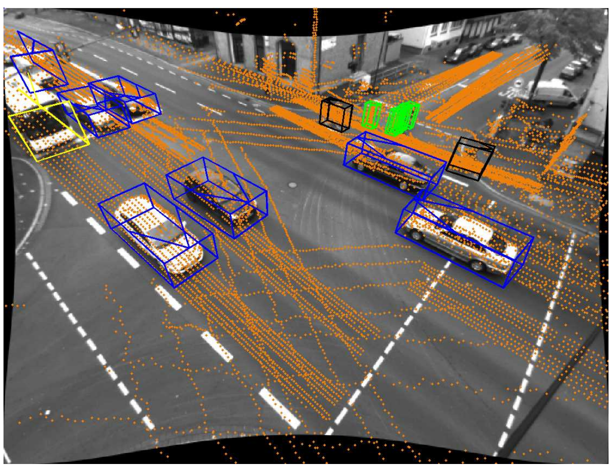


Fig. 8. Image of SK_1, projected laserscanner measurements, and object labels.

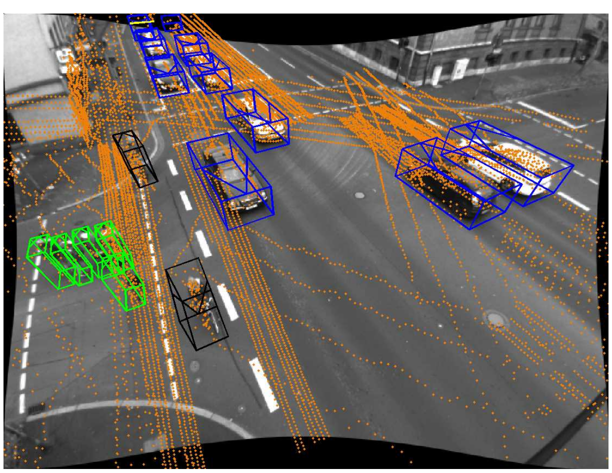


Fig. 9. Image of SK_4, projected laserscanner measurements, and object labels.

TABLE VI

CONTENT OF LABELED SEQUENCE 1

Sequence	Cars	Trucks	Pedestrians	Bikes	Duration (s)
1a	63	1	10	0	96
1b	63	3	13	3	96
1c	81	5	7	3	96
1d	83	3	8	4	97

TABLE VII

CONTENT OF LABEL DATA MEASUREMENT POINTS

```
result.BoxfittingLabels_REF.data{1}
.availableFeatures =
    FEAT_X_POS x position [m]
    FEAT_Y_POS y position [m]
    FEAT_Z_POS z position [m] (set to zero)
    FEAT_WIDTH width of object [m]
    FEAT_LENGTH length of object [m]
    FEAT_HEIGHT height of object [m]
    FEAT_ORIENTATION_ANGLE orientation angle [rad]
    FEAT_ID unique object id
    FEAT_CLASSIFICATION object class id
```

E. Reference Data

The reference data of the vehicles has been acquired by a real-time kinematic global positioning system (RTKGPS) with inertial measurement unit. Providing only reference data which features a highly accurate status and has been validated using the digital map of the intersection, guarantees a position accuracy better 0.15 m. Thus, the reference data is sufficient to evaluate environment perception methods. The reference data is referred to the center point of the cars' front bumper on street level.

The `availableFeatures` are shown in Tab. VIII.

TABLE VIII

CONTENT OF REFERENCE DATA MEASUREMENT POINTS

```
result.LS_1.data{1}.availableFeatures =
    FEAT_X_POS x position [m]
    FEAT_Y_POS y position [m]
    FEAT_Z_POS z position [m] (set to zero)
    FEAT_X_VEL velocity in x direction [m/s]
    FEAT_Y_VEL velocity in y direction [m/s]
    FEAT_Z_VEL velocity in z direction [m/s]
    FEAT_X_ACC acceleration in x direction [m/s^2]
    FEAT_Y_ACC acceleration in y direction [m/s^2]
    FEAT_Z_ACC acceleration in z direction [m/s^2]
    FEAT_WIDTH width of car without mirrors [m]
    FEAT_LENGTH length of car [m]
    FEAT_HEIGHT height of car [m]
    FEAT_YAW yaw angle [rad]
    FEAT_YAWRATE yaw rate [rad/s]
```

V. CONCLUSION

Within the german joint porject Ko-PER, a complex intersection perception system with 14 laserscanners and eight monochrome CCD cameras has been designed and installed. In order to provide access to sensor data of the unique this intersection, we prepared a dataset of three sequences of laserscanner and camera data. For two of the sequences highly accurate reference data of one car in each case is available. This two sequences include cars performing a right turn and a straight ahead maneuver. One sequence comes with object labels of several hundred road users including differenct object classes. The sequences are highly suitable to develop and evaluate tracking algorithms. With this dataset, we aim to stimulate further research in this area.

ACKNOWLEDGMENT

This work results from the joint project Ko-PER - which is part of the project initiative Ko-FAS - and has been funded by the German Bundesministerium für Wirtschaft und Technologie (Federal Ministry of Economics and Technology) under grant number 19S9022G.

REFERENCES

[1] "research initiative Ko-FAS," http://www.ko-fas.de, 7 2012.

[2] E. Strigel, D. Meissner, and K. Dietmayer, "Vehicle detection and tracking at intersections by fusing multiple camera views," in *IEEE Intelligent Vehicles Symposium (IV)*, 6 2013, pp. 882–887.

[3] D. Meissner, S. Reuter, E. Strigel, and K. Dietmayer, "Intersection-based road user tracking using a classifying multiple-model phd filter," *Intelligent Transportation Systems Magazine, IEEE*, vol. 6, no. 2, pp. 21–33, Summer 2014.

[4] K. Granström and U. Orguner, "A PHD filter for tracking multiple extended targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5657–5671, 2012.

[5] M. Feldmann and D. Franken, "Tracking of extended objects and group targets using random matrices - a new approach," in *International Conference on Information Fusion*, June 2008, pp. 1–8.

[6] J. W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, Issue 3, pp. 1042–1059, July 2008.

[7] A. Petrovskaya, "Towards dependable robotic perception," Ph.D. dissertation, Computer Science Department, Stanford University, 6 2011.

[8] A. Scheel, K. Granström, D. Meissner, S. Reuter, and K. Dietmayer, "Tracking and data segmentation using a ggiw filter with mixture clustering," in *Proceedings of the 17th International Conference on Information Fusion*, 7 2014.

[9] A. Gorjestani, A. Menon, P.-M. Cheng, B. Newstrom, C. Shankwitz, and M. Donath, "Macroscopic review of driver gap acceptance and rejection behavior at rural thru-stop intersections in the us-data collection results in eight states: Cicas-ssa report# 3," 2010.

[10] L. Alexander, P.-M. Cheng, A. Gorjestani, A. Menon, B. Newstrom, C. Shankwitz, and M. Donath, "The minnesota mobile intersection surveillance system," in *Intelligent Transportation Systems Conference, 2006. ITSC'06. IEEE*. IEEE, 2006, pp. 139–144.

[11] J. Ehrlich and et al, "Safespot sp2 - infrasens - sp infrastructure platform, d2.5.2 final report: Results on test and validation," Tech. Rep., 2009.

[12] K. Fuerstenberg and et al, "Intersafe2 - cooperative intersection safety - d1.2 final report," SICK AG, Tech. Rep., 2011.

[13] D. Meissner, S. Reuter, and K. Dietmayer, "Combining the 2d and 3d word: A new approach for point cloud based object detection," in *Intelligent Signal Processing Conference*, 12 2013.

[14] S. Kohler, M. Goldhammer, S. Bauer, K. Doll, U. Brunsmann, and K. Dietmayer, "Early detection of the pedestrian's intention to cross the street," in *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on*, Sept 2012, pp. 1759–1764.

[15] F. Seeliger, G. Weidl, D. Petrich, F. Naujoks, G. Breuel, A. Neukum, and K. Dietmayer, "Advisory warnings based on cooperative perception," in *Intelligent Vehicles Symposium (IV), 2014 IEEE*, 2014.

[16] M. Goldhammer, E. Strigel, D. Meissner, U. Brunsmann, K. Doll, and K. Dietmayer, "Cooperative multi sensor network for traffic safety applications at intersections," in *15th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 9 2012, pp. 1178–1183.

[17] D. Meissner and K. Dietmayer, "Simulation and calibration of infrastructure based laser scanner networks at intersections," in *Intelligent Vehicles Symposium (IV)*, 2010, pp. 670–675.