Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Behavior-Based Relative Self-Localization in Intersection Scenarios

Benedict Flade, Edoardo Casapietra, Christian Goerick, Julian Eggert

2017

Preprint:

This is an accepted article published in 20th International IEEE Conference on Intelligent Transportation Systems. The final authenticated version is available online at: https://doi.org/10.1109/ITSC.2017.8317772

Behavior-Based Relative Self-Localization in Intersection Scenarios

Benedict Flade[1*], Edoardo Casapietra[2*], Christian Goerick[1], Julian Eggert[1]

***Abstract*— Accurate map-relative localization is crucial when dealing with navigation or other driver assistance tasks. Global Navigation Satellite Systems (GNSS), such as GPS, GLONASS or Galileo, meet the requirements demanded from road-level navigation systems, but with lane-level assistance being the next step, more accurate localization technologies are needed. A lot of effort has been put in improving localization relative to lanes in scenarios with simple road geometry, whereas more complex scenarios such as intersections have received less attention, even though being a key factor in navigational tasks.**

In this paper, we focus on self-localization in the vicinity of intersections by using the trajectories of other cars as spatial cue. The geometry of the considered intersections as well as the ideal routes of cars traversing the intersection are derived from enhancements of freely available and crowdsourced map data. The trajectory of each car is compared to possible intersection routes in order to find the route which most likely corresponds to the actual trajectory. Subsequently, a geometrical transformation is computed to align the trajectory to the corresponding route and to estimate the position of the ego-vehicle with respect to the intersection.

The feasibility of the approach is demonstrated on multiple scenes from the KITTI Benchmark Dataset [1]. We show that the presented approach is suitable for self-localization in intersection scenarios.

I. Introduction

According to [2], intersection-related accidents present $40 - 50\%$ of all road crashes in most countries. Thus, the topic of localization in the vicinity of intersections is essential for the improvement of Advanced Driver Assistance Systems (ADAS) and for reducing the amount of traffic fatalities as is the goal of Vision Zero [3].

In this context, state-of-the-art ADAS have to cope with various challenges, one of which being reliable localization of the ego vehicle within a digital map. ADAS usually make use of Global Navigation Satellite Systems (GNSS), which have become an integral factor in the task of localizing any kind of vehicle, such as aircrafts, ships, trains and cars.

If we assume highly accurate localization, then the ego-car, as well as sensed surrounding traffic participants, can be located within a map. Based on this localization, lane-level map information can be used to assign lanes to each individual traffic participant, which then allows to estimate possible routes, to infer future behavior and to predict possible risk. In reality, both, the measured ego-position as well as the map itself, can contain significant errors of up to a few meters, which prevents lane-level localization as well as a wide range of prediction related ADAS tasks.

In this paper, we therefore focus on map-relative self-localization in the vicinity of intersections. This is done by using the trajectories of other cars as spatial cue in combination with a rough prior absolute position estimation, provided by GNSS such as GPS, GLONASS, GALILEO or similar systems. The geometry of the considered intersections as well as the ideal routes of cars traversing the intersection are derived from enhancements of freely available OpenStreetMap (OSM) data [4]. The trajectory of each car is compared to possible intersection routes in order to find the route which most likely corresponds to the actual trajectory. Subsequently, a geometrical transformation is computed to align the trajectory to the corresponding route and to estimate the position of the ego-vehicle with respect to the intersection.

The remainder of this work is organized as follows: In section II, we present related work and research. The explanation of our concept is divided into section III-A on map data representation and section III-B on the matching and alignment algorithm. Section IV demonstrates the approach on multiple scenes from the KITTI Benchmark Dataset. Finally, in section V, we conclude the paper with a summary and discussion on future steps.

II. Related Work

Self-Localization is provided by global satellite navigation systems but standard off-the shelf GNSS receivers alone do not meet the accuracy requirements for future ADAS technologies. Localization errors arise from distortions due to atmospheric interference as well as the time discrepancy between GNSS satellite and receiver. Reflection and diffraction of signals may be additional sources of error [5]. The latter may prevent successful localization and reduce the availability, especially in areas in which objects such as buildings and urban canyons, trees or mountains may block the direct line of sight. Owing to those errors, additional concepts providing a sufficient accuracy, integrity and reliability are required.

By using solutions such as differential GPS fused with vehicle sensors as presented in [6], the localization accuracy can be improved. Nevertheless, map-related errors due to map offsets or caused by the conversion from geographic to Cartesian coordinates still exist.

Toledo-Moreo et al. tackle this type of error by combining GNSS with dead reckoning and map data [7], [8]. In the proposed approach, data from a GNSS receiver, an odometer and a gyroscope is hybridized with road information stored in

* The first two authors contributed equally to this work

[1]Benedict Flade, Christian Goerick and Julian Eggert are with the Honda Research Institute (HRI) Europe, Carl-Legien-Str. 30, 63073 Offenbach, Germany {benedict.flade, christian.goerick, julian.eggert}@honda-ri.de

[2]Edoardo Casapietra is with the CoR-Lab, University of Bielefeld, Universitaetsstr. 25, 33615 Bielefeld, Germany eddy1988@hotmail.it

©2017 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC.2017.8317772

a digital map by a particle-filter-based algorithm. The authors show the general feasibility of the approach for lane-level applications. However, it occurred that outliers lead to an error in the position estimation but are found feasible by the map observations and are therefore not detected by their integrity indicator. For those cases, it is recommended to develop efficient means for GNSS fault detection.

With cameras, laser scanners and similar technologies finding their way into production vehicles, more and more research focuses on fusing visual data and object shape information with GNSS and map data. Schindler et al. present a self-localization approach in which environment information is detected by a monocular camera and a laser scanner [9]. The idea is to match sensed objects (lanes, road markings and landmarks) with corresponding elements in of high-precision digital map. Schreiber et al. use stereo camera data in combination with pre-recorded highly accurate map containing road markings and curb information [10]. In [11], the main focus is put on vehicle localization and SLAM in areas with poor GPS coverage. Levinson et al. combine maps of urban environments with GPS, IMU, wheel odometry, and LIDAR data. They create own map data in a first drive and perform a localization in a second drive. One major difficulty are dynamic objects, which complicate the localization process as they have to be tracked and removed. Common to those vision based approaches is the requirement of highly precise map data. Map makers are recording highly accurate 3D map data (see e.g. [12]) but such maps are unlikely to be globally available in the near future. Therefore, solutions employing crowdsourced resources should still be investigated.

A solution without the restraint of requiring HD Maps is shown in [13]. Cao et al. use basic road geometry derived from a publicly available map database. This geometry is automatically enhanced by inferring lane junctions and full road geometry based on lane width and type assumptions. Being based on orientation features instead of an explicit line-, or respectively lane-matching, Cao et al. claim that their approach is robust towards small local geometric transformations. A further localization approach using freely available map data is presented in [14]. Floros et al. introduce an approach for global vehicle localization by combining visual odometry with OSM data. Their goal is to use map data to compensate for the drift that visual odometry accumulates over time in order to improve the overall localization quality. The authors show that the proposed system provides automatic global localization at low infrastructural cost by only using simple map information.

Lane identification, also referred to as lane assignment, presents a sub-field of self-localization with the aim of estimating the discrete ego lane number. A visual lane assignment approach is presented in [15]. Kühnl et al. use a classifier on SPatial RAY (SPRAY) features, representing the spatial layout of the road. Being a purely visual based approach, the authors limit the evaluation on situations with low traffic density. Furthermore, they recommend to use a complementary system with surrounding vehicle detections from stereo, radar, or object recognition which is suitable for high traffic conditions as e.g. proposed in [16].

Rabe et al. present an algorithm to assign the ego lane using a particle filtering approach based on GPS, camera data, inertial sensors, radar and a road map [17]. Visual lane-marking detection is used in combination with radar measurements to find the position relative to road boundaries. The authors use a particle filter, considering static objects, moving objects detected by radar as well as visually detected lane markings. A set of particles, forming a cloud, is spawned around a GPS position estimate which is moving according to odometry and yaw rate measurements. The approach is showing promising results while approaching the intersection but the authors suggest to disable the lane-level localization while crossing an intersection since a significant portion of false lane-assignments could be observed in those cases. Therefore, situations such as turns on intersections or roundabouts were excluded. In subsequent work, the authors propose to use least-squares optimization on potential path hypotheses, weighting the distance to the hypothesis, the heading difference as well as the distance to lane markings [18]. Assuming that other vehicles drive basically in the center of a lane, their relative position is used to build a map of the lanes closest to the ego-vehicle, providing one further cue for the hypothesis evaluation. Rabe et al. show that this approach allows reliable lane assignment especially after intersections. Nevertheless, the approaches rely on visual lane marking detection and therefore reach their limits if lane markings are hardly visible or not existing at all.

With lane identification covering lateral direction, additional solutions are required to solve the task of 2D localization. In recent research, intersections have been used to perform localization also in longitudinal direction [19], [20] and [21].

Our approach targets map-relative localization at intersections based on the behavior of surrounding traffic participants. Here, the focus lies on calculating a position correction offset, thus allowing map-relative localization with its strength in longitudinal direction. Concerning the underlying map, our approach aims at using freely available road-level map data. This data is enhanced based on several assumptions, leading to lane-level road geometry. By doing so, lower accuracy of map geometry and lower level of detail have to be accepted, however, map data is available for practically any location and due to crowdsourcing, map data is updated frequently in return.

In the following section, we describe the basic concept of our road network representation framework which builds the basis of the behavior-based localization approach.

III. Methods

A. Enhanced Lane-Level Information for Digital Maps

Fig. 1 shows the general pipeline of the behavior-based self-localization approach. GNSS receiver output is used in combination with ego-vehicle sensor data to obtain other vehicle's trajectories in a global reference system. This data is then matched and aligned with enhanced geometric map

©2017 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC.2017.8317772

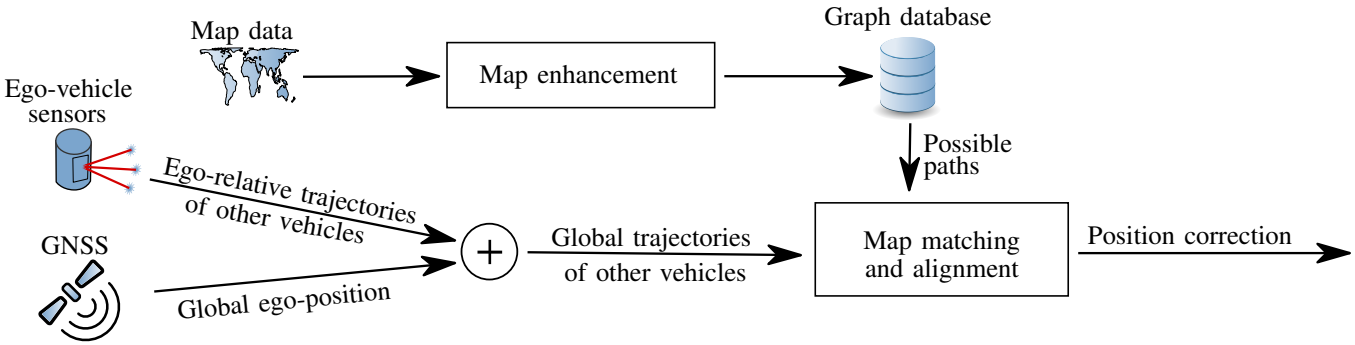


Fig. 1: Behavior-based relative self-localization approach. First, fusion of GNSS receiver output and ego-vehicle sensor data to obtain other vehicle's trajectories in a global reference system. Then, matching and alignment of this data with enhanced geometric map information, leading to a position correction.

information, derived from a crowdsourced map database. With regard to risk-estimation and lane-level navigation tasks, our goal is the calculation of a position correction offset of the ego-vehicle. This offset then similarly applies for the other traffic participants that are sensed relative to the ego-vehicle. The following sections explain the approach in more detail.

1) Graph Database: To store, recall and manipulate relational map information, we use a graph-database. Segments S and junctions J present the two basic elements types of the graph. From geometrical and property point of view, segments and junctions share the same characteristics. The difference is that junctions can be connected to an intersection node I, a third and structural element of the graph, which connects all junctions of one intersection.

Depending on the requirements, it is possible to store geometrical and topological information on up to three different layers of detail. The most detailed and lowest abstraction layer are lanes. With regard to multi-lane roads, all lanes pointing in the same direction form the so called *half-roads*, presenting the second layer. A *road*, being the third layer, therefore consists of up to two half-roads which itself can consist of an arbitrary number of lanes. This results in a set of seven map-element types $\mathcal{M} = \{S_{\text{road}}, S_{\text{half-road}}, S_{\text{lane}}, J_{\text{road}}, J_{\text{half-road}}, J_{\text{lane}}, I\}$. For each, except for the intersection, we generate and store geometrical left and right delimiter marking polylines, the center polyline as well as further optional information such as the drivable path, covered area or curvature.

All entities are described in geological latitude / longitude coordinates.

To obtain such geometrical and topological mapping information, approaches range from recording new data using extensive sensor equipment [11] to creating map data by manually parsing aerial satellite images [22], [23]. In order to be independent from a specific location, we chose to use data from a publicly available crowdsourced map database.

2) Crowdsourced Map Data: Collaborative mapping projects offer map data of the world with the advantage of their data being editable and updated regularly. In our work, we use OpenStreetMap (OSM) data [4]. OSM maps contain mostly road or half-road level information. Therefore they can be used for navigational and routing purposes on a strategic level but mostly do not directly provide the information needed for next generation navigation or future ADAS technologies. In the following, we separate mapping information into the three categories, namely provided, missing and additional.

Provided Information (PI): Provided is the centerline of the road, referred to as "way" in OSM terminology. A way is defined by several nodes, described in geographic coordinates. In the case of structurally separated roadways as e.g. highways, a way for each direction is modelled. The center of an intersection is thus implicitely defined by a node which is part of two different ways.

Missing Information (MI): With regard to behavior-based self-localization at intersections, lane-level information is needed, respectively lane center polylines as well as lane connection information ($\mathcal{M}_{\text{lane}} = \{S_{\text{lane}}, J_{\text{lane}}\}$). The amount of lanes within a road is added as a tag, but explicit lane-level geometrical information, respectively a polyline, is missing. Explicit information about how individual lanes are connected is provided in few cases in big cities by a tag (left, through and right), but not widely available. Hence, we classify lane-level junction data J_{lane} as missing as well. Furthermore, the width of the road is rarely tagged and thus also classified as missing.

Additional Information (AI): OSM data provides the type of the road as e.g. motorway, trunk, residential or living street. Traffic rules and traffic signs such as speed limits are included as well in most cases. Both can be used as a cue for missing information such as lane width.

3) Trajectory Generation: For our self-localization approach, shown in Fig. 1, lane-level road geometry information (S_{lane} and J_{lane}) is required. Since the OSM map data only provides road centerlines as well as tagged lane number and road type, we enhance the map data and store the resulting lane-level geometry information in our graph database, described in section III-A.1.

The enhancement of the OSM data is illustrated in Fig. 2

Starting from the road centerline (Fig. 2a), we infer lane segment centerlines (Fig. 2b). In this context, we have to assume a lane width. Hall et al. compare the width values of 23 countries [24]. According to their research, the lane

©2017 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC.2017.8317772

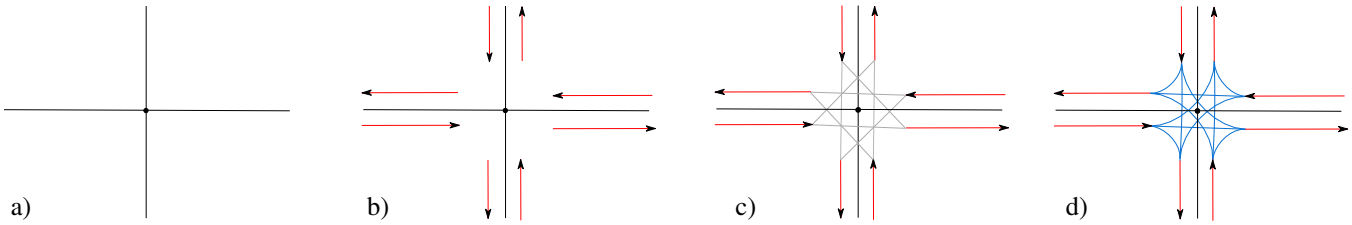


Fig. 2: Enhancement of map data. a) Road-level map information from crowdsourced map database. b) Inference of lane segment centerlines. c) Inference of connections between lane segments. d) Calculation of lane junction geometry.

width ranges globally from 2.75 m to 4.0 m on freeways and arterial roads. If we additionally consider minor or local roads the width ranges from 2.25 m to 4.0 m. For individual countries, the largest variations for freeway and arterial road width can be found in the Netherlands, Yugoslavia and Czech Republic ($\pm$37.5 cm). Considering all road types, the largest range is found in South-Africa ($\pm$72.5 cm). With the aim of lane level accuracy, it is therefore needed to take guidelines for urban road design of the individual country into account in order to narrow the maximum error based on the road type. For the example of Germany, one can then assume a maximum lane width error of $\pm$25 cm.

We further assume the other traffic participants to drive in the center of a lane. Liu et al. study the lateral distance between lane centerline and vehicle track on underground roads with three lanes. According to their research, the mean value of the lateral distance for lanes with a width of 3.25 m can range, depending on the position of the lane (left, middle or right lane), from 3 cm to 11 cm with standard deviations in the range of 35 cm to 41 cm. Both, lane width and driving path assumption can present sources of error and should be kept in mind, when calculating a position correction in later steps.

Based on the lane segment information, we infer possible connections (Fig. 2c) of each lane segment by heuristics. We now have intersection, or more precisely, lane junction information on topological level but no geometrical junction information yet. To obtain the geometry of lane junctions, we interpolate between two related lane segments.

For the interpolation, B-Spline curves [25]

$$B(t) = \sum_{i=0}^{n-1} \boldsymbol{p}_i N_{i,k}(t), \tag{1}$$

are chosen, where the basis function $N_{i,k}(t)$ of degree k and order $k+1$ is defined by

$$N_{i,0}(t) = \begin{cases} 1, & \text{if } t_i \leq t < t_{i+1} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

$$N_{i,k}(t) = \frac{t - t_i}{t_{i+k} - t_i} N_{i,k-1}(t) + \frac{t_{i+k+1} - t}{t_{i+k+1} - t_{i+1}} N_{i+1,k-1}(t), \tag{3}$$

with $\boldsymbol{t} = \{t_0, ..., t_{n+k+1}\}$ being a nondecreasing sequence of real numbers, called the knot vector. In the case of a division by zero, the respective quotient is defined to yield zero.

We use fourth order B-Splines consisting of cubic polynomial functions. The n control points $\{\boldsymbol{p}_0, \boldsymbol{p}_1, ..., \boldsymbol{p}_{n-1}\}$ lie on the two corresponding lane segments and should be chosen close to the end in order to guarantee a smooth transition (Fig. 2d).

Additionally, lane markings can be derived by adding lateral offsets to the lane junction centerlines and are useful for visual consistency checks when superimposed on aerial images.

B. Matching and Alignment

In the case of errors in the measured absolute ego-position or when dealing with offsets regarding the map data, neither the ego-vehicle, nor surrounding traffic participants that are sensed by the ego-vehicle, can be correctly placed on the map. In Fig. 3, the ego-vehicle is placed in the map under usage of absolute position information. Similar to the ego-vehicle, also the other traffic participant is placed on the map with an offset. The actual map-relative position is indicated by the respective semi-transparent vehicle. Therefore, it is required to estimate the errors. The following section presents the matching and alignment step of the approach.

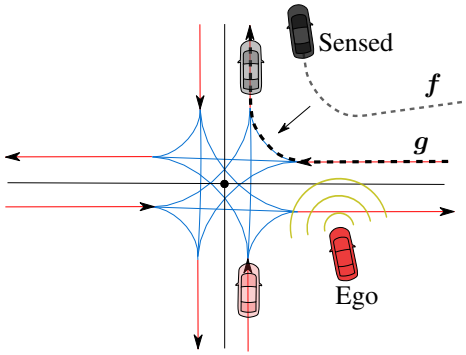


Fig. 3: Intersection scenario with an ego-vehicle wrongly placed outside the road. A sensed traffic participant is placed on the map with similar offset. Actual map-relative position is indicated by semi-transparent vehicle.

Map-relative self-localization is done by comparing the traces of other traffic participants with possible paths, derived from the map data. Since the traces are measured with respect to the ego-car, their global position depends on the ego-car reference frame. Correcting the alignment between traces and map-derived routes allows us to estimate the error of the ego-car position in the map reference frame.

The success of the approach depends on the accuracy and length of the sensed traces. State-of-the-art vehicle sensors that serve to sense the environment include computer vision, lidar and radar. In general, our approach is designed

©2017 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC.2017.8317772

to be independent from the choice of sensing technology. Nevertheless, it influences the expected results and accuracy significantly. Computer vision is the cheapest solution and is particularly suitable for tasks such as texture or scene interpretation, which is not of high importance in our case. Lidar, which uses pulsed laser light to measure distances, has its strengths in accurately detecting objects in a distance of more than 100 m in all directions. Waiting at an intersection, it is therefore possible to record a trace of a length of over 200 m if a vehile is passing from one side to the other, assuming optimal conditions. Compared to radar, the drawback of lidar and computer vision is its vulnerability to occlusions, for example caused by other traffic participants or buildings. A solution to the problem of occlusion is the combination of computer vision and radar data as e.g. presented in [26].

The typical trace of a detected car is shorter than the routes extracted from the lane-level map data. For this reason, the matching has to be between a trace and a small section of a route. The trace $\boldsymbol{f}$ is iteratively compared to a section $\boldsymbol{g}$ of each route with the same amount of points N, obtained by a moving window. Each comparison is carried out by computing a score, featuring two different contributions. The first contribution is related to the Euclidean distance $\|\boldsymbol{d}_{\boldsymbol{f},\boldsymbol{g}}\|$ between the average point of the trace and the average point of the route section:

$$\boldsymbol{f}_{\text{av}} = \sum_{i=1}^{N} \boldsymbol{f}_i / N \tag{4}$$

$$\boldsymbol{g}_{\text{av}} = \sum_{i=1}^{N} \boldsymbol{g}_i / N \tag{5}$$

$$\boldsymbol{d}_{\boldsymbol{f},\boldsymbol{g}} = \boldsymbol{f}_{\text{av}} - \boldsymbol{g}_{\text{av}} \tag{6}$$

If the distance is higher than a maximum distance D, the section is discarded and the system moves to the next section. The second contribution represents the similarity between the two shapes. The tangential direction of motion along the two curves is computed on every point, and a histogram is created for each curve. Both histograms are compared by computing the Pearson correlation coefficient $r_{\boldsymbol{f},\boldsymbol{g}}$, which will be used for the overall score. A negative correlation coefficient leads to immediately discarding the section. The score is expressed as the following:

$$S_{\boldsymbol{f},\boldsymbol{g}} = \frac{(D - \|\boldsymbol{d}_{\boldsymbol{f},\boldsymbol{g}}\|)}{D} + r^2_{\boldsymbol{f},\boldsymbol{g}} \quad \text{with } D \neq 0, \tag{7}$$
$$\text{if } \|\boldsymbol{d}_{\boldsymbol{f},\boldsymbol{g}}\| < D \text{ and } r_{\boldsymbol{f},\boldsymbol{g}} > 0, \text{ otherwise skip}$$

This score function represents a trade-off between proximity and similarity between the two curves. Distance needs to be considered in order to ensure that the closest lane is always picked over a more distant one, while the correlation factor makes sure the shape of the lane fits with the trace. In our experiments we have these two contributions normalized between $[0, 1]$, with the correlation factor being squared in order to favor high values over average ones.

The match between a car trace $\boldsymbol{f}$ and an ideal route $\boldsymbol{g}$ is determined based on the highest score, and the alignment proceeds thereafter by subtracting the average point from both curves:

$$\overline{\boldsymbol{f}} = \boldsymbol{f} - \boldsymbol{f}_{\text{av}} \tag{8}$$

$$\overline{\boldsymbol{g}} = \boldsymbol{g} - \boldsymbol{g}_{\text{av}} \tag{9}$$

The covariance matrix $C_{\boldsymbol{f},\boldsymbol{g}}$ of the two new curves is computed, and a singular value decomposition (SVD) is performed:

$$C_{\boldsymbol{f},\boldsymbol{g}} = \text{cov}(\boldsymbol{f}, \boldsymbol{g}) = \overline{\boldsymbol{f}} \cdot \overline{\boldsymbol{g}}^T \tag{10}$$

$$U, \Sigma, V = \text{svd}(C_{\boldsymbol{f},\boldsymbol{g}}) \tag{11}$$

The rotation matrix R necessary to align the two curves, and thus the angle of rotation $\alpha_{\boldsymbol{f},\boldsymbol{g}}$, is then calculated as

$$R = U^T \cdot V \tag{12}$$

$$\alpha_{\boldsymbol{f},\boldsymbol{g}} = \arctan \frac{R_{0,1}}{R_{0,0}} \quad \text{with } R_{0,0} \neq 0, \tag{13}$$

while the translation required for the alignment is simply $\boldsymbol{d}_{\boldsymbol{f},\boldsymbol{g}}$.

It has to be noted that different kinds of trajectories do not carry the same amount of information. In particular, trajectories lying on a straight line provide information only on the orthogonal axis. This happens because the trajectory can be matched to any section of that line, so that the position along the line is not relevant to our application. In order to account for this phenomenon, we consider the *significant offset* of the trajectory, which we define as the translation necessary to align the trajectory to the ideal route, projected onto the axis orthogonal to the trajectory itself. This can be computed from the covariance matrix of trajectory $\boldsymbol{f}$:

$$C_{\boldsymbol{f}} = \text{cov}(\boldsymbol{f}) = \overline{\boldsymbol{f}} \cdot \overline{\boldsymbol{f}}^T \tag{14}$$

$$U, \Sigma, V = \text{svd}(C_{\boldsymbol{f}}) \tag{15}$$

If the singular values in Σ are in decreasing order, the axis we are looking for is the second column of U, as we are interested in the direction of minimum variance. The significant offset $\delta(\boldsymbol{f}, \boldsymbol{g})$ is therefore computed as follows:

$$\delta(\boldsymbol{f}, \boldsymbol{g}) = \boldsymbol{d}_{\boldsymbol{f},\boldsymbol{g}} \cdot U_2 \tag{16}$$

IV. Experiments

In this section, the approach is applied to streams from the KITTI benchmark dataset [1]. This datasat provides, amongst others, position and heading of the ego-vehicle as well as trajectories of other traffic participants, detected by a laserscanner . For all lanes, we assume a width of 3.3 m.

The following experiments aim at demonstrating the potential of our approach by estimating the discrepancy between the map data and the sensed traces of the other traffic participants. This discrepancy can then be used to correct the ego-vehicle position with regard to the underlying map.

©2017 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC.2017.8317772

A. Position Correction

Stream 017 presents a scenario with an ego-vehicle observing a large intersection, at which three cars traverse from left to right on a straight trajectory. The approach has been applied on three vehicles and the results are shown in Table I. Fig. 4a shows the sensed trace as well as its alignment for vehicle #2.

From visual inspection, it appears that in this stream, there appears to be a discrepancy of about $5.5\,\mathrm{deg}$ between the vehicle traces provided by KITTI and the map data we used. The coordinate system of the digital map is oriented with the x-axis facing East and the z-axis facing North. The position offset $[-0.69, -1.21]$ m of the ego-car can be estimated as the average of the traces offset of the three sensed vehicles . However, as all three traces are straight and along the same direction, the offset is reliable only along the orthogonal orientation (as explained in section III-B). Thus, the significant offset δ, which we defined as the absolute value of the required translation, is calculated to be -1.41 m along the direction defined by the vector $[0.46, 0.89]$.

ID #	Maneuver	Position diff. [m]	Direction diff. [deg]
1	Straight	[-0.51, -1.29]	5.46
2	Straight	[-0.98, -1.09]	5.09
3	Straight	[-0.58, -1.26]	6.10

TABLE I: Stream 017 discrepancy estimation results

Stream 057 presents a similar situation of a large intersection, but with a higher number of vehicles. Fig. 4b shows the corresponding map including the trace of Vehicle #6. Most of the vehicles traverse the intersection on a straight trajectory, while a few perform turning maneuvers. Table II shows the position and direction difference of trajectories of 13 vehicles: 11 with a straight trajectory and 2 with a turning maneuver.

It can be seen that the vehicles with a straight trajectory exhibit a consistent discrepancy (with the exception of one clear outlier, #4, possibly due to an uncommon driving line) of approximately $-1\,\mathrm{deg}$ and an offset in position of around 1 m to the south with respect to our map. It is important to notice that the intersection is oriented almost exactly aligned with the four cardinal points. Furthermore, all straight trajectories refer to vehicles that are passing the intersection either from East to West or from West to East. This in turn means that all straight trajectories are oriented along the axis $[1, 0]$, and thus, the offset is reliable along their N-S component only. Excluding the two turning maneuvers, the significant offset δ is calculated to -1.36 m along the $[0, 1]$ direction.

B. Robustness Consideration

A further experiment aims at investigating the robustness by introducing an artificial error. The error is supposed to be corrected by the matching and alignment. This evaluation may serve as a first hint towards robustness. However, prior to an implementation in a vehicle, additional analyses have to be performed.

As a test, we apply an additional bias of $[-0.5, -0.5]$ m to the traces of the other vehicles. In the following, the bias is always applied to the position of the ego and oriented in x and z direction. Resulting estimated offsets show mixed performances (see Table II), as 3 out of 11 straight trajectories are matched with a different lane, due to the proximity and similarity of parallel lanes. All other 8 correct matches show a precise translation of 0.5 m along the z axis (South to North). Since all straight trajectories run from East to West or vice versa, the discrepancy along the x axis, which runs parallel, is unreliable. The significant offset obtained by averaging the 8 well-behaving traces is -1.83 m. Relating to the result of -1.36 m from section IV-A, the new significant offset is consistent with the -0.5 m bias.

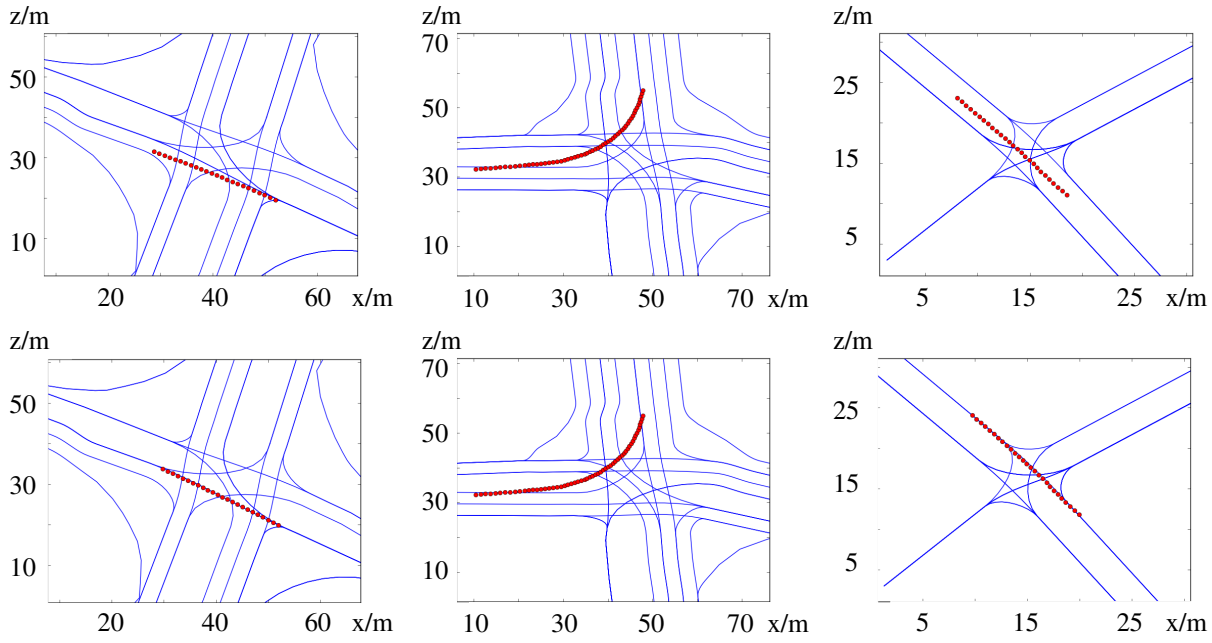
The reliability of our approach decreases for traces of turning cars. This effect is caused by multiple factors. First of all, the lane-level map data is generated based on several assumptions and overly simplified. Paths do not match the real behavior of a car which is causing potential errors in the matching step. Secondly, having only a short trace, which usually lacks the start and finish of the maneuver, causes potential ambiguity in the matching phase, especially when the offset is high.

The last experiment has been carried out on stream 046. This stream shows a small intersection, with only one vehicle traversing, on a straight trajectory, from West to East (Fig. 4c). The significant offset computed from this trace is -1.75 m along the direction $[0.75, 0.66]$. Since there are no multiple lanes with same direction in this scene, we can apply a stronger artificial error to the trace of that car, without the risk of it being matched with the wrong lane. In order to show the behavior of the significant offset, we apply two different biases: $[-1, -2]$ m and $[+1, -2]$ m, whose projections onto the direction of significant offset are respectively -2.01 m and -0.56 m. Those are the changes in significant offset we expect after applying the bias. The results, provided by Table III, show that the new significant offsets are consistent with the expected ones.

V. Conclusion and Outlook

In this paper, we present a novel method to improve the self-localization of a vehicle in the vicinity of an intersection. This is done by exploiting the behavior, respectively the trajectories, of other cars detected in the scene. The result of our approach does not depend on the behavior of the ego car, but on the detected surrounding vehicles. Therefore it makes no difference if the ego-vehicle is approaching, in the middle of, or leaving an intersection, assuming an omnidirectional sensor setup. We could show promising results of our methods by applying our approach on selected streams of the KITTI benchmark dataset. In the evaluation of our approach, the trace data of the other traffic participants is provided by the dataset. In general, the approach is suitable for arbitrary methods of sensing the others cars, nevertheless

©2017 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC.2017.8317772



(a) Vehicle #2 of Stream 017 (b) Vehicle #6 of Stream 057 (c) Vehicle #1 of Stream 046

Fig. 4: Three examples of car traces before (top) and after (bottom) the alignment, one for each stream.

ID #	Maneuver Straight / Turning	Position diff. [m]	Direction diff. [deg]	Position diff. after bias [m]	Direction diff. after bias [deg]	Comments
1	Straight	[-0.18, -1.34]	0.15	[0.03, 1.42]	-0.19	Wrong lane matched
4	Straight	[-0.32, -2.52]	3.35	[-0.82, -3.02]	3.35	Meets expectation
5	Turning	[1.40, -0.95]	2.10	[1.35, -1.04]	0.13	Alignment failed
6	Turning	[0.43, -1.55]	-1.93	[0.51, -1.70]	-0.06	Alignment failed
12	Straight	[-0.12, -0.85]	-1.25	[0.36, -1.35]	-1.34	Meets expectation
13	Straight	[-0.30, -1.94]	-1.13	[-0.80, -2.44]	-1.13	Meets expectation
14	Straight	[0.11, -0.76]	-0.33	[-0.39, -1.26]	-0.33	Meets expectation
15	Straight	[-0.43, -1.10]	-0.80	[0.06, -1.60]	-0.92	Meets expectation
16	Straight	[0.40, -1.11]	-1.89	[-0.10, -1.61]	-1.89	Meets expectation
17	Straight	[-0.13, -1.23]	-1.08	[0.36, -1.71]	-1.15	Meets expectation
20	Straight	[0.37, -1.18]	-1.65	[-0.13, -1.68]	-1.65	Meets expectation
21	Straight	[-0.38, -1.00]	-0.91	[-0.16, 1.90]	-0.44	Wrong lane matched
22	Straight	[0.25, -1.58]	-1.36	[0.65, 1.19]	-1.57	Wrong lane matched

TABLE II: Stream 057 robustness consideration. Bias of $[-0.5, -0.5]$ m added to the traces.

ID #	Initial position diff. [m]	Initial direction diff. [deg]	Bias [m]	Position diff. due to bias [m]	Direction diff. due to bias [deg]	$\delta(\boldsymbol{f}, \boldsymbol{g})$ (Expected)	$\delta(\boldsymbol{f}, \boldsymbol{g})$ (Measured)
1	[-1.55, -0.88]	1.39	[-1, -2]	[-3.07, -2.24]	1.76	-3.76 m	-3.79 m
			[+1, -2]	[-1.59, -1.61]	2.14	-2.31 m	-2.25 m

TABLE III: Stream 046 robustness consideration. Bias of $[-1, -2]$ and $[+1, -2]$ added to the trace of the sensed traffic participant.

it remains to be evaluated to which extent different detection methods like computer vision, radar, lidar affect the results.

Currently, there are a few potential sources of error that will be topic of further research: First, our underlying assumption of drivers following the centerline, interpolated using B-Splines, is overly simplistic and therefore a more sophisticated behavior model will be investigated.

Second, both the trajectories of the other traffic participants as well as the road polylines are sampled. In our evaluation, we assume mostly equally sampled polylines. This can be realised by resampling either the vehicle trajectories or the map data. The influence of irregularly sampled polylines will be investigated in subsequent work.

Finally, it has to be noted that the accuracy of the

©2017 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC.2017.8317772

OpenStreetMap data varies depending on the location and that the inference of lane segments and junctions is not always reliable. In order to prove its suitability for production vehicles, on-line evaluation of the approach by using a mobile platform on real world scenarios has yet to be done.

ACKNOWLEDGMENTS

Map data copyrighted OpenStreetMap contributors and available from http://www.openstreetmap.org. This work has been supported by the European Union's Horizon 2020 project *inLane*, under the grant agreement number 687458.

REFERENCES

[1] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.

[2] S. Lefèvre, C. Laugier, and J. Ibañez-Guzmán, "Risk assessment at road intersections: Comparing intention and expectation," in *IEEE Intelligent Vehicles Symposium*, 2012, pp. 165–171.

[3] M.-Å. Belin, P. Tillgren, and E. Vedung, "Vision Zero a road safety policy innovation," *International Journal of Injury Control and Safety Promotion*, vol. 19, no. 2, pp. 171–179, jun 2012.

[4] OpenStreetMapContributers, "www.openstreetmap.org."

[5] S. Schon and F. DilBner, "Challenges for GNSS-based high precision positioning - some geodetic aspects," in *2007 4th Workshop on Positioning, Navigation and Communication*. IEEE, mar 2007, pp. 229–237.

[6] S. Rezaei and R. Sengupta, "Kalman Filter-Based Integration of DGPS and Vehicle Sensors for Localization," *IEEE Transactions on Control Systems Technology*, vol. 15, no. 6, pp. 1080–1088, nov 2007.

[7] R. Toledo-Moreo, D. Betaille, F. Peyret, and J. Laneurit, "Fusing GNSS, Dead-Reckoning, and Enhanced Maps for Road Vehicle Lane-Level Navigation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 5, pp. 798–809, oct 2009.

[8] R. Toledo-Moreo, D. Betaille, and F. Peyret, "Lane-Level Integrity Provision for Navigation and Map Matching With GNSS, Dead Reckoning, and Enhanced Maps," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 1, pp. 100–112, mar 2010.

[9] A. Schindler, "Vehicle self-localization with high-precision digital maps," in *2013 IEEE Intelligent Vehicles Symposium Workshops (IV Workshops)*. IEEE, jun 2013, pp. 134–139.

[10] M. Schreiber, C. Knoppel, and U. Franke, "LaneLoc: Lane marking based localization using highly accurate maps," in *2013 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, jun 2013, pp. 449–454.

[11] J. Levinson, M. Montemerlo, and S. Thrun, "Map-Based Precision Vehicle Localization in Urban Environments," in *Robotics: Science and Systems III*. Atlanta, GA, USA: Robotics: Science and Systems Foundation, jun 2007.

[12] TomTom, "TomTom launches RoadDNA: Precise localisation technology for automated driving," accessed 15-June-2016. [Online]. Available: http://corporate.tomtom.com/releasedetail.cfm?ReleaseID=931327

[13] G. Cao, F. Damerow, B. Flade, M. Helmling, and J. Eggert, "Camera to map alignment for accurate low-cost lane-level scene interpretation," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, nov 2016, pp. 498–504.

[14] G. Floros, B. van der Zander, and B. Leibe, "OpenStreetSLAM: Global vehicle localization using OpenStreetMaps," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, may 2013, pp. 1054–1059.

[15] T. Kühnl, F. Kummert, and J. Fritsch, "Visual ego-vehicle lane assignment using Spatial Ray features," in *2013 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, jun 2013, pp. 1101–1106.

[16] D. Svensson and J. Sorstedt, "Ego lane estimation using vehicle observations and map information," in *2016 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, jun 2016, pp. 909–914.

[17] J. Rabe, M. Necker, and C. Stiller, "Ego-lane estimation for lane-level navigation in urban scenarios," in *2016 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, jun 2016, pp. 896–901.

[18] J. Rabe, M. Meinke, M. Necker, and C. Stiller, "Lane-level map-matching based on optimization," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, nov 2016, pp. 1155–1160.

[19] F. Kuhnt, R. Kohlhaas, R. Jordan, T. Gussner, T. Gumpp, T. Schamm, and J. M. Zollner, "Particle filter map matching and trajectory prediction using a spline based intersection model," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, oct 2014, pp. 1892–1893.

[20] V. Popescu, M. Bace, and S. Nedevschi, "Lane identification and ego-vehicle accurate global positioning in intersections," in *2011 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, jun 2011, pp. 870–875.

[21] A. Barth, J. Siegemund, and J. Schwehr, "Fast and precise localization at stop intersections," in *2013 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, jun 2013, pp. 82–87.

[22] N. Mattern and G. Wanielik, "Camera-based vehicle localization at intersections using detailed digital maps," in *IEEE/ION Position, Location and Navigation Symposium*. IEEE, may 2010, pp. 1100–1107.

[23] G. Mattyus, S. Wang, S. Fidler, and R. Urtasun, "Enhancing Road Maps by Parsing Aerial Images Around the World," in *2015 IEEE International Conference on Computer Vision (ICCV)*. IEEE, dec 2015, pp. 1689–1697.

[24] L. Hall, R. Powers, D. Turner, W. Brilon, and J. Hal, "Overview of cross section design elements," in *International Symposium on Highway Geometric Design Practices*, 1998, pp. 12:1–12.

[25] L. Piegl and W. Tiller, *The NURBS Book*, ser. Monographs in Visual Communications. Springer Berlin Heidelberg, 1995.

[26] J. Zhao and D. Yao, "Occlusion adaptive object tracking based on video image and radar data," in *2006 6th International Conference on ITS Telecommunications*. IEEE, jun 2006, pp. 943–946.

©2017 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC.2017.8317772