

In-Lane Localization in Road Networks Using Curbs Detected in Omnidirectional Height Images

Jörg Stückler, Institut für Informatik, Universität Freiburg,
Hannes Schulz, Institut für Informatik, Universität Freiburg,
Dr. Sven Behnke, Institut für Informatik, Universität Freiburg;

Kurzfassung

Die Verwendung von differentielltem GPS ist für das autonome Befahren vordefinierter GPS Routen nicht ausreichend. Autonome Fahrzeuge müssen ihre Fahrbahn für ein sicheres Befahren zusätzlich wahrnehmen. In diesem Beitrag stellen wir eine Methode zur Schätzung des Positionierungsversatzes innerhalb der Fahrspur bezüglich eines abstrakten Straßenmodells vor, in der Bordsteine mit einem 64-strahligen Laser-Entfernungssensors detektiert werden. Die Entfernungsmessungen werden in omnidirektionale Höhenbilder transformiert, welche in sensorintrinsic Gier- und Nickwinkelkoordinaten repräsentiert werden. In dieser dichten bild-ähnlichen Repräsentation können Bordsteine durch Anwendung einfacher orientierter Kantenfilter detektiert werden. Bordsteinmessungen werden mit Erwartungen registriert, die am Straßenmodell abgeleitet werden. Die resultierenden instantanen Versatzmessungen zwischen Straßenmodell und detektiertem Straßenverlauf werden schließlich gefiltert, um eine Korrektur der Position innerhalb der Fahrspur zu erhalten.

Abstract

For autonomous driving on predefined GPS routes, the use of differential GPS alone is not sufficient. Autonomous cars must perceive the street in order to drive safely. In this paper, we propose to estimate in-lane positioning offsets with regard to a high-level road network model by sensing street curbs with a 64-beam laser-range-finder. The laser measurements are transformed into height images represented in sensor-intrinsic yaw and pitch coordinates. In this dense image-like representation, curbs can be detected with simple oriented edge filters. The curb measurements are matched to expected curbs derived from the road network model. This scan-matching procedure yields measurements of offsets between road model and detected road, which are filtered to obtain a correction of the in-lane position.

1. Introduction

In many autonomous car applications, driving on predefined GPS routes is common practice. For this purpose, differential GPS in combination with inertial and odometry measurements is used to gain position measurements with a relatively high accuracy compared to regular

GPS. To follow predefined routes, road network models are used which typically are based on sequences of GPS waypoints. The route trajectory is intended to keep the car in a safe position within the lane. However, errors in the GPS measurements and inaccuracies of the road network model cause deviations of the car's estimate from its intended position on the lane. Thus, autonomous cars must perceive the road to improve localization within the lane. We propose to correct position measurements by matching a high-level road network description to the measurements of a rotating 64-beam LIDAR (light detection and ranging) sensor. These measurements are processed as height images, represented in the sensor-intrinsic coordinate system of pitch and yaw angles. Preserving the neighborhood of adjacent measurements, removal of high objects, and filling-in of missing measurements yields a dense image-like representation of the road and the curbs, in which curbs can be detected by simple oriented edge filters. The filter responses are matched to expected curbs, which are derived from the road network model. This scan-matching yields an instantaneous estimate of the position and orientation offsets of the measured curbs to the expected curbs. The offset observations are probabilistically filtered using a Kalman filter. The method estimates the position offset of the vehicle to its intended position within the lane. We evaluate the proposed approach using a data set that was captured by one of the autonomous cars that recently participated in the DARPA Urban Challenge 2007.

This paper is structured as follows: In Sec. 2 we discuss related work on the detection of curbs with various sensory equipment. Sec. 3 outlines the LIDAR and position sensing devices that have been used for our experiments. The processing of LIDAR measurements for the detection of curbs is detailed in Sec. 4. We describe the method for estimating the positioning offset within the lane with respect to the road network model in Sec. 5 and Sec. 6. The proposed approach is experimentally evaluated in Sec. 7 on a data set which has been acquired in an urban environment.

2. Related Work

In urban environments, curbs are dominant features for the course of roads. Also, driving onto curbs must be avoided to prevent from damage to the car and its inmates. Thus, the detection of curbs has gained interest from several researchers. Typical sensors applied for curb detection are cameras (e.g. [1, 2, 3]), radar (e.g. [4, 5]), and LIDAR (e.g. [6, 7]).

An approach that employs LIDAR sensors has been proposed by Wijesoma et al. [6]. They detect curbs with a 2D laser-range-finder (LRF) that is mounted in front of the car and that is tilted downwards to sense the road in a distance between 10-15m. They segment and filter line segments in each 1D scan. A pair of curbs is extracted as line segments that are ap-

proximately parallel to the road and have a distance that corresponds to an assumed road width. Finally, these extracted curbs are tracked using a Kalman filter. While this method establishes correspondence of curbs along the road in subsequent time steps, the robustness of the method could be increased by considering measurements at adjacent scanning pitch angles during the extraction of curbs.

Oniga et al. [3] detect curbs in dense 3D data obtained from stereo cameras. They generate 2D elevation maps storing the maximum height of measurements at discrete positions in a metric grid. In this representation, they detect edges and extract line segments using the Hough transform. Among these line segments at most two are chosen as curbs, if the median heights on both sides of the curb in a local neighborhood constitute a small height variation. As data is sparse in this Cartesian representation, they dilate the height image to generate context for edge detection. In our approach, the sensor-intrinsic coordinate system is preserved yielding a dense image-like representation. Grid cells without height occur, if the measurements in the cell have failed, if they are much higher than the road, or if they have been reflected by the car itself. For such cells, height is estimated from local context.

3. Sensor Description

A Velodyne HD LIDAR provides 3D range scans by rotating a 64-beam LRF array around its vertical axis at high rates (10 Hz in our application). In the horizontal direction, the LRF provides 360° field-of-view with an angular resolution of approximately 0.09°. Vertically, the pitch angles range from -24.8° to +2° with an angular resolution of ca. 0.4°. Its range measurement accuracy typically is within 5 cm. Highly reflective obstacles can be measured in distances of up to 120 m, whereas the measurable distance to objects with lower reflectance like the pavement approximately is 50 m. The sensor is mounted on top of the car providing range scans with full field-of-view in horizontal direction, while parts of the car and sensory equipment are constantly detected. Fig. 1 shows the LIDAR and its placement on the car. In addition, the car is equipped with an Applanix LV220 system that fuses differential GPS, inertial measurements, and odometry to a robust position estimate.

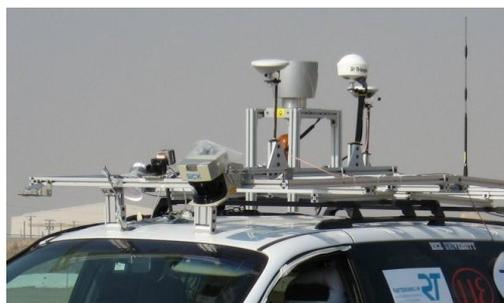


Fig. 1. Left: Velodyne HD Lidar sensor. Right: Placement of the sensor.

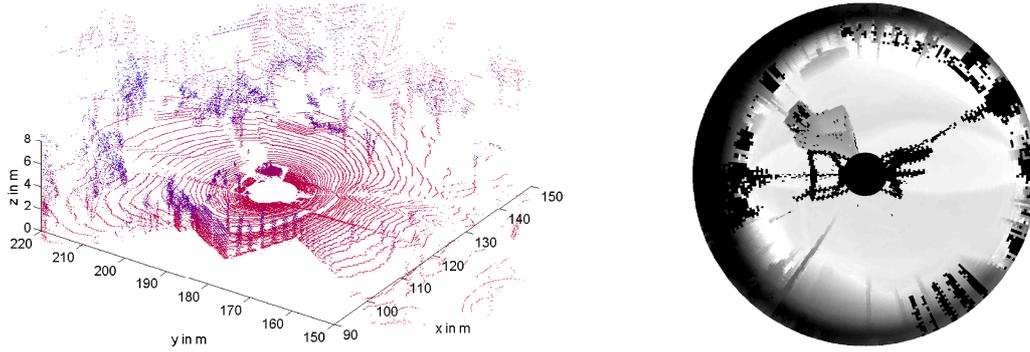


Fig. 2. Left: 3D Measurements of a complete LIDAR sweep. Right: Omnidirectional height image without minimum and maximum range readings (black cells).

4. Bottom-Up Curb Detection in Omnidirectional Height Images

We transform the range and bearing measurements of a complete LIDAR sweep into sensor-centric Cartesian coordinates. The height components of the measurements are incorporated into an angular grid $l(\theta, \phi)$ that is discretized in yaw (θ) and pitch (ϕ) coordinates. We choose 360 grid cells in yaw and 64 cells in pitch direction. For cells containing multiple measurements, we average their heights and relative 2D position to the car. In this omnidirectional height image (s. Fig. 2) curbs are characterized by small, elongated height steps. Our aim is to detect curbs at those positions where small height differences occur in a local context. If each cell has such context, convolution operators for smoothing and derivation can be applied to obtain an edge image.

Naturally, not every grid cell contains valid range measurements. When multiple measurements fall into one cell, we determine the difference between minimum and maximum heights. A large difference indicates steep measurement changes, such that averaged height and relative position are presumably invalid. Also, height steps and heights above the average road height which are caused by objects higher than curbs must be removed. By this it is possible to detect curbs close to high objects. Finally, we remove measurements of the car.

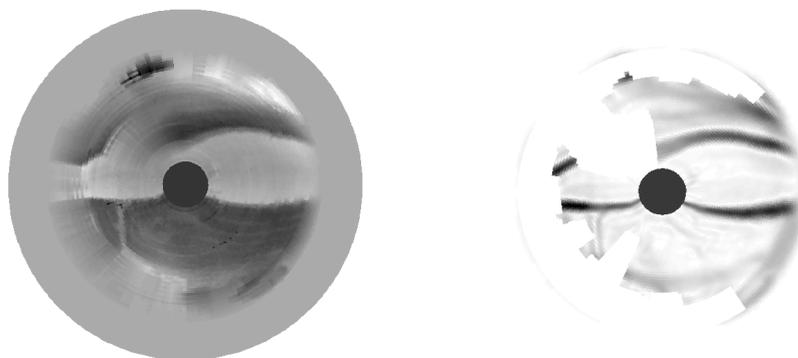


Fig. 3. Left: Omnidirectional height image after removal of invalid measurements filling-in with adjacent heights. Right: Edge response in valid cells (darker is stronger).

To still provide context, cells without measurements are filled-in with linear interpolations of adjacent heights. This yields a dense and image-like representation of the height measurements (s. Fig. 3), in contrast to the sparse Cartesian 3D representation.

To measure the height of objects with respect to the road, we estimate the average road height. In each yaw direction in the grid we track along the pitch direction starting at the lowest angle (i.e. close to the car in the height image) with the attachment height of the sensor. For each cell, a linear recursive filter is applied to estimate the road height considering its measured height and the height of its preceding cells. If the measured height is not compatible to the tracked height, it is not updated with the cell's measurement and the procedure is resumed with the next cell. By this, the road height estimate incorporates gradual changes. In this representation we can use simple image processing operations to detect edge strength and orientation: We convolve the image with smoothing and first-order derivative convolution kernels c_ϕ and c_θ in pitch and yaw direction separately. We account for the differing angular grid resolution in pitch and yaw direction by applying kernels of differing width:

$$c(k) := \left(\frac{1}{16} [1, 4, 6, 4, 1] \right)^k * \left[-\frac{1}{2}, 0, \frac{1}{2} \right]$$

For the pitch direction kernel we use two binomial five-kernels for smoothing, i.e. $c_\phi := c(2)$, whereas for the yaw direction $c_\theta := \alpha c(6)$ we apply six. To be able to estimate orientations from both edge responses, we normalize the edge responses such that the equivalent convolution kernels have equal two-norms, i.e. $\alpha := \frac{\|c(2)\|_2}{\|c(6)\|_2}$.

Using the convolved images $I_\theta := c_\theta * I$ and $I_\phi := c_\phi * I$, we can easily determine edge strength $s(\theta, \phi) := \sqrt{I_\theta(\theta, \phi)^2 + I_\phi(\theta, \phi)^2}$ and orientation $\psi(\theta, \phi) := \text{atan2}(I_\phi(\theta, \phi), I_\theta(\theta, \phi)) + \theta$. To obtain orientations in the car-centric coordinate frame, we must take the yaw direction θ of a cell into account.

Of course, this edge image also contains responses at locations in the image that do not correspond to curbs but to other small height changes in the environment. In the next section, we therefore combine the bottom-up edge detection with top-down information about the expected location of curbs which we derive from our road model.

5. Top-Down Curb Detection in a High-Level Road Network Model

Our abstract road model is given in the DARPA Road Network Description Format (RNDF). In this format, roads are modeled as segments. Each segment consists of multiple lanes,

which are described by average lane width and a sequence of GPS waypoints ordered by the direction of traffic. Transitions between lanes (e.g. at road crossings) are represented by corresponding entry and exit waypoints. As the GPS waypoints are sparsely distributed, the lanes are interpolated by Akima splines [8] to provide the target trajectory for driving. The interpolated lanes are used to derive expected curb positions by sampling the spline in fixed intervals in forward and backward direction and by expecting curbs orthogonal to the spline on each side of the road. We ignore expectations at road crossings, as our assumption that curbs are parallel to the course of roads is violated. The lane width given in the RNDF is only an estimate of the average width along the whole lane. We therefore improve the lane width estimate within a segment recursively by adapting it to the distance of complementary curb measurements along the orthogonal direction to the spline. Given this estimate, we can determine predictions for curb measurements as follows. We sample the lane spline at fixed positions in forward and backward direction. We choose a distance of 0.5m between sample points and 20m in each direction. At the sample points, we predict curbs orthogonally to the spline at the road boundary to each side. The road boundaries are simply determined by half the lane width to the right of the rightmost lane and the rest of the road width to the left. Curb measurements within a range of 5m around the car are not expected. For each expected curb $q \in Q_{left/right}$ a corresponding curb measurement $p \in P_{left/right}$ is determined: First, the curb strengths and orientations in the polar grid are transformed into a Cartesian 2D grid using the average relative position in each polar grid cell (a grid resolution of 0.1m is used). Curbs are characterized by ridges in rising edge strength along the spline direction. Thus, we correlate ridge-like patterns along the direction orthogonal to the spline, where we only account for the strength of an edge in a cell, if it is oriented appropriately. For each expected curb q a measured curb p is detected where the correlation measure attains its maximum in a specific search interval (e.g. $[-1.2m, 2m]$). See Fig. 4 for an example.

6. Recursive Offset Estimation

Scan-matching with a point-to-line metric is used to match curb measurements P to curb expectations Q on each road side in car-centric coordinates. We seek to obtain the instantaneous rotation R and translation t of the car that explains the displacement of the measured curbs to the expected curbs. We apply a variant of the Iterative-Closest-Points-Algorithm [9] (ICP) to determine rotation R and translation t that minimize the error function

$$E(R, t) = \frac{1}{|C|} \sum_{i \in C} \|Rp_i + t - \Pi(Rp_i + t, Q)\|^2$$

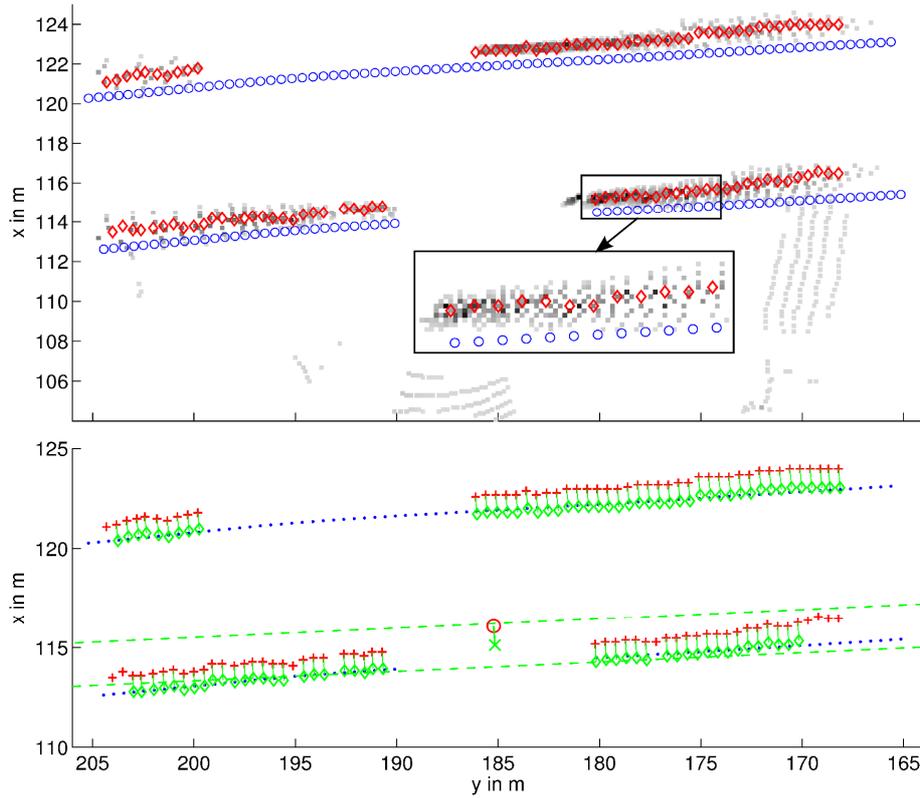


Fig. 4. Top: For each expected curb (blue circles) a corresponding measurement (red diamonds) is detected on the ridges in edge strength in a Cartesian grid (grey squares). Bottom: Scan-matching of curb measurements (red points) with predictions (blue points) using a point-to-line metric yields the instantaneous displacement of the car (green cross with green dashed σ range) from its assumed position (red circle) within the lane. Transformed measurements are shown as green diamonds.

between corresponding points, where $\Pi(p, Q)$ maps p to its projection onto the line between its closest point $q \in Q$ and the predecessor or successor of q , $q' \in Q$. To improve efficiency, we use the expected curb position instead of searching for the closest position among the expected curbs to each measured curb. We therefore ignore correspondences, for which the projection π of p does not lie between q and q' . We regard correspondences as outliers, if the Euclidean distance between π to p is 2 times the average distance of points in P to their projections. We define the set of indices C as maximal subset of $\{1, \dots, K\}$, for which points in P with valid correspondences exist.

The offset can only be usefully measured, if correspondences exist on both sides of the road and to the front as well as to the back of the car. Let n_{left} , n_{right} , n_{front} , and n_{back} be the number of correspondences to the left, the right, the front, and the back, respectively. We consider the correspondences for further processing, if 10% lie on either side, i.e.

$$\frac{\min(n_{left}, n_{right})}{n_{left} + n_{right}} > 0.1 \text{ and } \frac{\min(n_{front}, n_{back})}{n_{front} + n_{back}} > 0.1.$$

To determine the transformation between the matched points and their projections, we first subtract their mean:

$$\bar{p}_c := p_c - \mu_p, \mu_p := \frac{1}{|C|} \sum_{c \in C} p_c \text{ and } \bar{\pi}_c := \pi_c - \mu_\pi, \mu_\pi := \frac{1}{|C|} \sum_{c \in C} \pi_c.$$

Then, we determine rotation R and translation t by singular value decomposition

$$\text{of } W = \sum_{c \in C} \bar{p}_c \bar{\pi}_c^T = USV^T : R = UV^T \text{ and } t = \mu_\pi - R\mu_p.$$

In each time step, the translation obtained from scan-matching is interpreted as offset measurement. On straight roads, curbs are aligned on two parallel lines. In this case, the positioning offset is only observable orthogonal to the spline. On curved roads, all directions can be observed. To obtain a measurement covariance that reflects the high uncertainty in unobservable dimensions, we apply the approximation method proposed in [10]. Let

$E(x) := E(R(\theta), t)$ be the error function in terms of $x := (t_x, t_y, \theta)^T$ and let

$z := (p_1, \dots, p_K, q_1, \dots, q_K)^T$. The covariance Σ_x of x is approximated in closed form by

$$\Sigma_x \approx \left(\frac{\partial^2 E}{\partial x^2} \right)^{-1} \frac{\partial^2 E}{\partial z \partial x} \Sigma_z \frac{\partial^2 E^T}{\partial z \partial x} \left(\frac{\partial^2 E}{\partial x^2} \right)^{-1}.$$

As we assume stochastic independence between point observations, the covariance Σ_z of z is diagonal and calculation of Σ_x is simplified. For the individual measurement covariances we model higher uncertainty the larger the distance of the measurement is to the car. The covariance Σ_t of t is obtained by marginalization, i.e. by extracting the Cartesian 2x2-submatrix from Σ_x . Fig. 4 illustrates this scan-matching procedure.

Finally, we recursively estimate the translational offset from the offset measurements with a linear Kalman Filter (KF) [11]. An innovation test is used to reject outlier measurements. The positioning error within the lane is obtained by projection of the offset on the normal to the car heading. Also, the offset is used to initialize the top-down search for curb measurements.

7. Experimental results

The proposed approach is evaluated on a data set acquired in an urban environment. It contains moving and static vehicles on and near the road. Fig. 5 shows qualitative results of our approach. The left figure demonstrates that the approach captures positioning offsets within the lane. The right figure shows that gross insufficiencies of the road model (e.g. missing

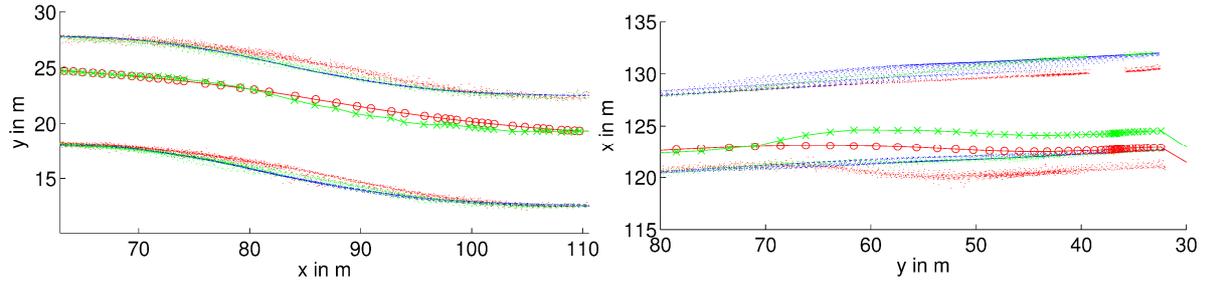


Fig. 5. In-lane localization results (colors/markers as in Fig. 4). Left: The method estimates in-lane positioning errors caused by inaccuracies of the road model. Right: Gross insufficiencies of the road model (e.g. missing lanes) cause the method to estimate a wrong offset.

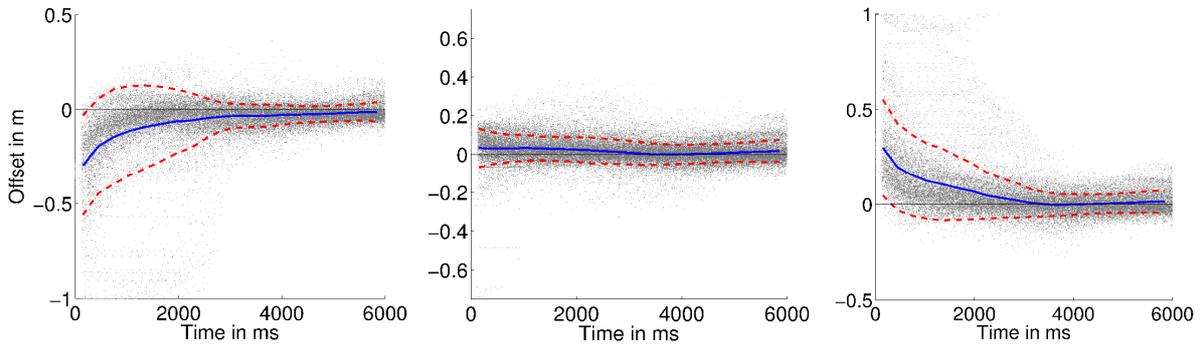


Fig. 6. Deviations (grey points) from the estimated positioning offset within the lane after inducing artificial errors of -1m (left), no error (middle), and 1m (right). Mean (blue solid line) and standard variance (red dashed line) show, that the positioning error is quickly captured by the algorithm within 4 seconds.

lanes) cause the method to estimate a wrong offset.

To assess the performance of our approach quantitatively despite missing ground truth, we compared the estimated offset to the offset obtained, when artificial errors are induced on the measured relative 2D position of curbs. Ideally, the difference between both offsets vanishes up to the artificial error and the standard deviation in this difference is low. Multiple runs on the data set are evaluated each lasting 6 seconds, where the spline is approximately linear, and where no crossings and insufficiencies occur in the road model. At the beginning of each run, the KF offset estimate is reset and an artificial offset orthogonal to the spline direction is induced. Fig. 6 depicts the results for artificial offsets of -1m (left), 1m (right), and no offset (middle) in approx. 1000 runs each. Mean and standard deviation are determined by grouping the data into time intervals of 0.3 sec. After 5 secs, the mean is close to 0m and the standard deviations are below 0.056m, 0.0567m, and 0.0594m for -1m offset, no offset, and 1m offset, respectively. In this time interval, outliers lie within 0.2429m, 0.2386m, and 0.2410m, respectively. The results show, that our approach captures offsets of 1m with a good accuracy.

8. Conclusions

In this paper, we propose an approach to estimate the errors present in differential GPS and road network models by the use of an omnidirectional 3D LIDAR device. The approach improves in-lane localization for applications like autonomous car driving. In our real-time method, curbs are detected in a dense, image-like representation of the 3D range scans, which contributes an efficient method to process the large amount of LIDAR data. The positioning offset within the lane is determined by filtering instantaneous displacements of curb measurements to expected curbs which are derived from a high-level road network model. Our experiments indicate that our approach estimates the in-lane positioning offset with a good accuracy. Its limitations are the reliance on a sufficient road model and the necessity of detectable curbs. To further improve accuracy and reliability when our assumptions are violated the approach could be combined with other sensory cues.

Acknowledgements

The authors would like to thank Team Berlin from Freie Universität Berlin for providing access to the hardware and software of their autonomous car.

- [1] D. Pomerleau and T. Jochem. Rapidly adapting machine vision for automated vehicle steering. In *IEEE Expert*, 1996.
- [2] E. D. Dickmanns and B. D. Mysliwetz. Recursive 3-D road and relative ego-state recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992.
- [3] F. Oniga, S. Nedeveschi, and M. M. Meinecke. Curb detection based on elevation maps from dense stereo. In *IEEE Int. Conf. on Intelligent Computer Communication and Processing*, 2007.
- [4] K. Kaliyaperumal, S. Lakshmanan, and K. Kluge. An algorithm for detecting roads and obstacles in radar images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996.
- [5] M. Nikolova and A. Hero. Segmentation of a road from a vehicle-mounted radar and accuracy of the estimation. In *IEEE Intelligent Vehicles Symposium*, 2000.
- [6] W.S. Wijesoma, K.R.S. Kodagoda, and A.P. Balasuriya. Road-boundary detection and tracking using ladar sensing. *IEEE Transactions on Robotics and Automation*, 2004.
- [7] H. Cramer and G. Wanielik. Road border detection and tracking in non-cooperative areas with a laser radar system. In *German Radar Symposium*, 2002.
- [8] H. Akima. A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points. *ACM Transactions on Mathematical Software*, 1978.
- [9] P. Besl and N. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992.
- [10] A. Censi. An accurate closed-form estimate of ICP's covariance. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2007.
- [11] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME, Journal of Basic Engineering*, 1960.