

LaneLoc: Lane Marking based Localization using Highly Accurate Maps

Markus Schreiber, Carsten Knöppel and Uwe Franke

Mobile Perception Systems
FZI Research Center for Information Technology
76131 Karlsruhe, Germany
schreiber@fzi.de

Environment Perception Group
Daimler AG Group Research & Advanced Engineering
71059 Sindelfingen, Germany
{carsten.knoepfel,uwe.franke}@daimler.com

Abstract—Precise and robust localization in real-world traffic scenarios is a new challenge arising in the context of autonomous driving and future driver assistance systems. The required precision is in the range of a few centimeters. In urban areas this precision cannot be achieved by standard global navigation satellite systems (GNSS). Our novel approach achieves this requirement using a stereo camera system and a highly accurate map containing curbs and road markings. The maps are created beforehand using an extended sensor setup. GNSS position is used for initialization only and is not required during the localization process. In the paper we present the localization process and provide an evaluation on a test track under known conditions as well as a long term evaluation on approximately 50 km of rural roads, where a precision in centimeter-range is achieved.

I. INTRODUCTION

Autonomous driving is becoming a main focus in driver assistance research since it offers increased comfort and safety at the same time. In contrast to early approaches of the 90th that were purely based on vehicle sensors (e.g. in the Prometheus project [1]), recently proposed solutions explicitly make use of map data. This simplifies the task, since one can follow a pre-computed path as long as it is not covered by obstacles. In this paper we will ignore sensing tasks like obstacle detection, traffic light recognition etc. but concentrate on a precise localization relative to the given map that is necessary for the path following.

In open space (air, sea, desert) high precision Inertial Navigation Systems (INS) combining an Inertial Measurement Unit (IMU) and a global navigation satellite system (GNSS) reach the required accuracies in the centimeter range. In these scenarios any path can be defined in WGS84 coordinates and the current position of the autonomous system can be measured by a commercially available INS. Unfortunately, in typical traffic scenes several disturbing effects, i.e. occlusions by vegetation and buildings or multi-path effects due to reflections, render this approach as not feasible.

Therefore the goal is to realize a localization system that is independent of satellite systems. Usually a “map” is generated that contains sufficient information to determine the current position of the autonomous system relative to this map. In their path breaking work, Google [2] builds a bird’s eye image of laser reflections on the road surface and determines the current position by matching the current

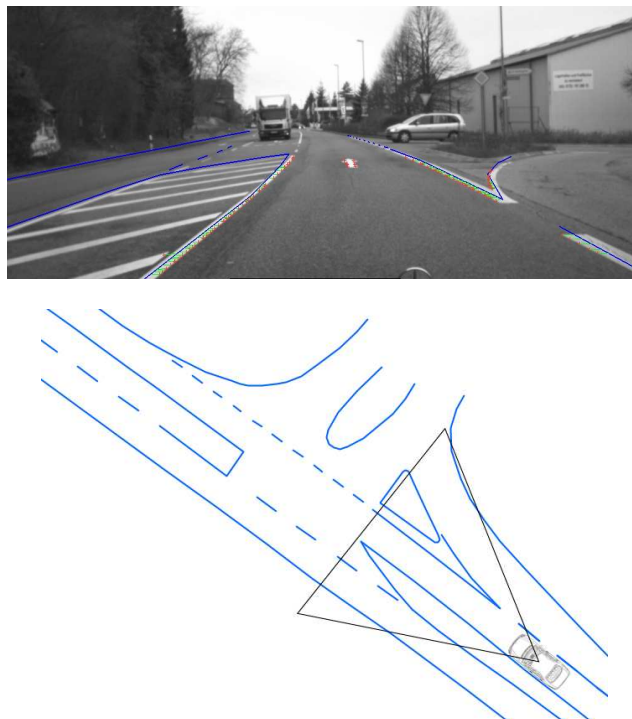


Fig. 1. Example of a map and the reprojection in the camera image.

sensor readings with the recorded and filtered data. The 360° Velodyne HDL 64 laser scanner also used by German research groups in Hannover and Berlin provides a large field of view and is independent of illumination. Unfortunately, commonly used laser scanners can only be a research solution since both, costs and design make them unattractive for the use in consumer vehicles.

Stereo camera systems already used in driver assistance systems in upper class vehicles like Mercedes E-class and S-class are an attractive alternative. Recently, vision based localization has become a popular research topic. Badino [3] presented an extremely robust and fast integrative approach using low-cost cameras for a rough localization that could be used for initialization. More common are feature based schemes that build 3D-feature maps using bundle adjustment, e.g. [4] and use those maps for later localization. This approach requires a sufficiently structured environment

and stable features to obtain reliable results. Unfortunately, rural areas often suffer from unique structures. Moreover, despite the progress in feature detection and description the robustness of generic feature with respect to illumination and weather is still unsatisfactory. These facts limit this approach to urban areas and cooperative conditions if robustness is mandatory.

On most roads where we intend to drive autonomously, the lanes are limited by markings and curbs. It is well known from commercially available lane keeping systems how this information can be robustly extracted even under adverse weather and lightning conditions. Pink [5] proposed to extract markings and their world position from air plane images and to match this data to the current image. His scheme was not real-time capable and suffered from occlusions by trees and other objects. A more general approach was already presented by Franke in 2003 [6] who determined the position of a car relative to an arbitrary marking configuration in order to guide a bus into its parking position.

In our approach described in this paper we generate - in a separate drive - a map that contains markings and curbs. While driving, the map is matched to the current images. Like in common lane recognition systems, a Kalman Filter is used to determine the position of our vehicle relative to the markings, both in lateral and longitudinal direction. The localization algorithm runs in real-time on our test vehicle and has been extensively tested on urban and rural roads. Due to missing ground truth data we cannot compute the precision directly, but perform experiments that let us claim an accuracy in the range of approximately 10 cm.

The paper is organized as follows: Section II describes the mapping process, in Section III the online localization is explained. The experimental evaluation follows in Section IV. Finally, conclusions are drawn in Section V.

II. MAPPING

Common maps, i.e. in commercially navigation systems or Open Street Map, usually represent roads as connected lines with additional attributes. However, exact road geometry is not encoded. Therefore, we generate our own maps that contain all visible lane markings and curbs on the road. To achieve high reliability in the map and guarantee outlier free data, the chosen representation of the map data can be reviewed and controlled manually.

Formally a map consists of a set of line segments with different attributes. Each line segment l_i is defined by a starting point $p_{s,i}$, an end point $p_{e,i}$ and a describing attribute a_i , where $p_{s,i}, p_{e,i} \in \mathbb{R}^3$ (latitude, longitude, altitude) and $a_i \in \{\text{solid}, \text{dashed}, \text{curb}, \text{stopline}\}$. In the case of dashed lines, each p_i specifies beginning and end of a marked segment on the road. Stoplines usually occur perpendicular to the driving direction.

Mapping and online localization are completely separated. For mapping purposes a vehicle with an extended sensor setup is used. Position data is obtained by a high precision GNSS unit with recording of GNSS raw data for postprocessing. A velodyne laser scanner delivers a high range

360 degree environment. Thus the complete road geometry, including parts of diverging roads, can be captured.

An additional camera setup with downwards facing cameras and wide angle lenses is used to obtain high resolution images of the road surface without occlusion by other vehicles. This setup also restrict the influence of changing roll and pitch angles while driving. From both sensors a bird's eye view is generated (Figure 4). On the camera images a lane detection algorithm (see Section III-C.1) is applied to automatically extract lane marking information. However, due to the orientation of the cameras other lanes on the road are not covered. For further lane markings or road boundaries on the opposite side, the images generated with the laserscanner are used. The low spatial resolution and contrast of these images limit the use of automatically lane detection algorithms, so the marking information has to be labeled manually (Figure 5). While we assume an accuracy of approx. 2 cm in automatically generated maps, lane marking positions errors increase in manually labeled parts up to 10 cm.

A complete overview of the mapping process is given in Figure 2. Figure 3 shows the vehicle for mapping.

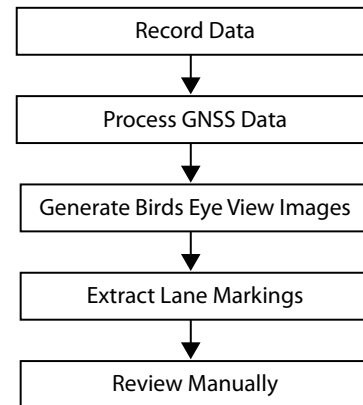


Fig. 2. Overview of the mapping process: Record data, process GNSS data, generate birds eye view images, extract lane markings, review manually.



Fig. 3. Vehicle for recording mapping data: Equipped with an high precision GNSS unit, Velodyne laser scanner and cameras.

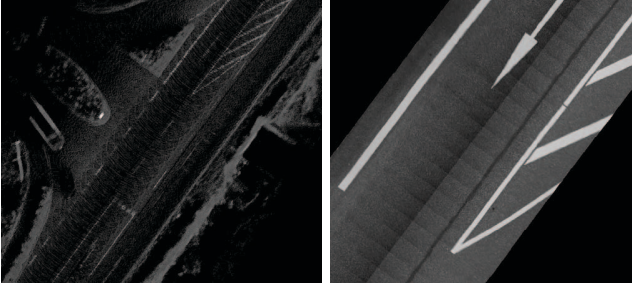


Fig. 4. Birds eye views generated by Velodyne laser scanner with long range (left) and downwards facing cameras with high resolution for relevant lane markings and curbs (right).

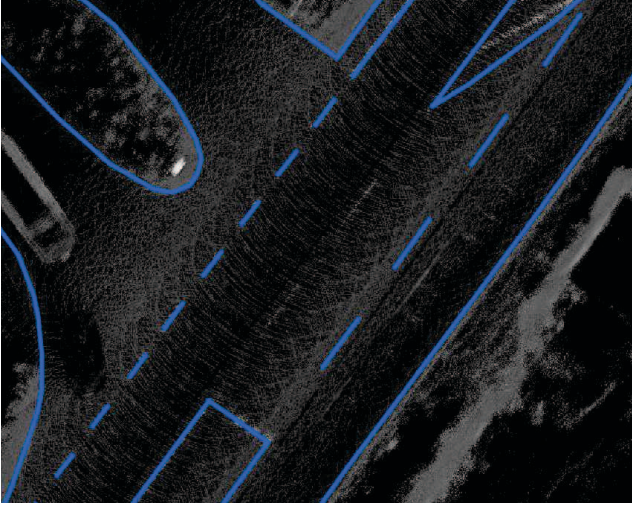


Fig. 5. Map of an intersection with manually labeled lane markings based on Velodyne data.

III. ONLINE LOCALIZATION

Lane marking based localization proposed by Pink [5] is based on an iterative solver which matches lane marking patches from a satellite images generated map into the camera image. This localization shows good results if there is a sufficient number of marking patches. To increase the accuracy and the robustness with respect to the localization efficiency in terms of computing time we consider the problem from the perspective of lane recognition. Here, the use of a Kalman Filter guarantees a high degree of robustness and the marking detection can be performed by means of well tested methods. Similar to standard lane recognition systems, our localization system uses a forward facing camera system and an IMU (see Figure 6). For initialisation an additional GNSS unit is used.

In the following we propose our Kalman Filter based localization model, the map matching technique and the measurement extraction of lane markings and curbs.

A. Localization Model

The ego vehicle is described by means of a point position \vec{P}_{veh} and an orientation φ in an stationary coordinate system (X/Y), shown in Figure 7. Accordingly, the state vector

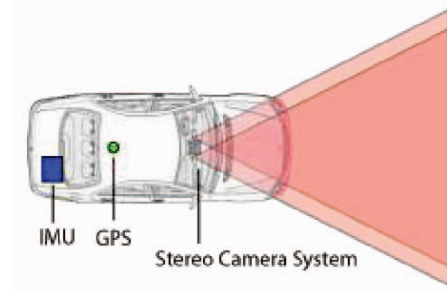


Fig. 6. Sensor setup of the localization vehicle. GNSS is used for initialisation only.

\vec{x}_{veh} is given by:

$$\vec{x}_{veh} = \begin{bmatrix} \vec{P}_{veh} \\ \varphi \end{bmatrix} = \begin{bmatrix} X_{veh} \\ Y_{veh} \\ \varphi \end{bmatrix} \quad (1)$$

Our proposed non-linear discrete system model $f(\vec{x}, \vec{u})$ is given by:

$$\vec{x}_{veh,k+1} = f(\vec{x}_{veh,k}, \vec{u}_k) = \begin{bmatrix} \vec{P}_{veh,k+1} \\ \varphi_{k+1} \end{bmatrix} \quad (2)$$

$$= \begin{bmatrix} X_{veh,k} + v \cdot \Delta t \cdot \cos(\varphi_k + \Delta t \cdot \dot{\psi}_{veh}) \\ Y_{veh,k} + v \cdot \Delta t \cdot \sin(\varphi_k + \Delta t \cdot \dot{\psi}_{veh}) \\ \varphi_k + \Delta t \cdot \dot{\psi}_{veh} \end{bmatrix} \quad (3)$$

In this model the control vector \vec{u} consists of the yaw rate $\dot{\psi}_{veh}$ of the ego vehicle and its velocity v . Δt is the cycle time between to consecutive innovation steps.

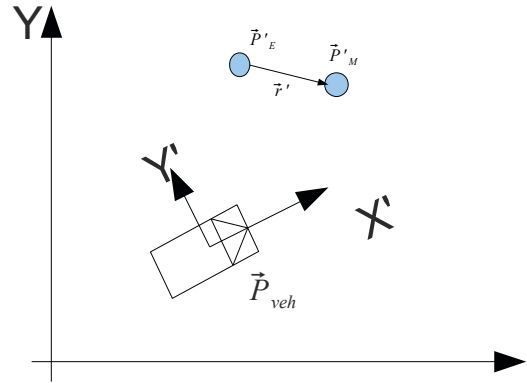


Fig. 7. Measurement \vec{P}_M in the vehicle koordinates $X' Y'$; Map Point \vec{P}_E transformed in the vehicle koordinates $X' Y'$; Residium \vec{r} in the vehicle system $X' Y'$; Ego Position \vec{P}_{veh} in the inertial coordinate system XY

The residual \vec{r} between a map point \vec{P}'_E and a measuring point \vec{P}'_M in the vehicle coordinate system (X'/Y') is shown in Figure 7.

$$\vec{r}' = \vec{P}'_M - \vec{P}'_E \quad (4)$$

A map point \vec{P}'_E is determined in the stationary coordinate system. To use \vec{P}'_E in Equation 4, \vec{P}'_E has to be transformed into the vehicle coordinate system (X'/Y') as follows:

$$\vec{P}'_E = \underline{R}_{veh} \cdot (\vec{P}_E - \vec{P}_{veh}) \quad (5)$$

\underline{R}_{veh} specifies the rotation from the stationary to the vehicle orientated coordinate system.

Our measurement model describes the residual between the expected position of all measurement points $\vec{P}'_{e,i}$, $i \in \{1, \dots, n\}$ as a function $h(\vec{x}_{veh})$ depending on the state vector and all measurements $\vec{P}'_{m,i}$, $i \in \{1, \dots, k\}$ as \vec{y} in the vehicle coordinate system.

$$\vec{r}' = \vec{P}'_M - \vec{P}'_E = \vec{y} - h(\vec{x}_{veh}) \quad (6)$$

The variance of measurement noise consists of noise in the map data σ_{map}^2 and noise in the back-projection σ_{cam}^2 , which depends on the camera height and the orientation to the road.

$$\sigma_{meas}^2 = \sigma_{map}^2 + \sigma_{cam}^2 \quad (7)$$

Equation 7 allows us to describe both sources of measurement noise separately. With our mapping process and localization system we assume $\sigma_{map}^2 = 10\text{cm}$ and $\sigma_{cam}^2 = 3\text{px}$, respectively.

For robustness reasons the measurement noise is determined as a probabilistic function depending on its residuum as proposed by Hartley and Zissermann [7]. In our approach we used a Cauchy Function to limit the influence of outliers.

B. Map Matching

The used map contains line segments to represent the markings or curbs, while our measurements are obtained as points. The goal of map matching is to achieve the best possible matching of a point measurement to a line segment. This means both a lateral association as well as a longitudinal which is not determined by the search for the shortest distance between a measurement point and a line segment (Figure 8). Therefore, each map line segment is sampled to map points (Figure 8 C). For each of these map points the closest measurement point is searched in order to extract the residuals described above. This achieves both, lateral and longitudinal residuals.

C. Measurement Extraction

1) *Lane Marking Measurement*: An oriented matched filter is applied to detect the lane markings, as it is successfully used as a robust lane measurement extraction in common lane detection systems [8]. Nevertheless, there are still problems caused by incorrect marking measurements on objects, i.e. cars, walls etc. To reduce this, we evaluate only image areas in the Freespace [9] delivered by our stereo vision system.

To detect lane markings, the map is projected into the image using the current estimate and search lines are positioned around the expected lane marking position. The

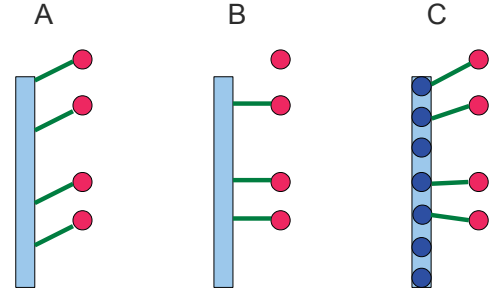


Fig. 8. Three cases of measurement connections at the birds eye view of one dashed lane marking; red points are marking measurements, the connections/residuals are shown as green line. (A) shows is the best case, unfortunately not possible; (B) shows the minimum point line distance match with no longitudinal constraints; (C) shows the connections with sampled map-points from map lines.



Fig. 9. Inner city scene with parking cars along the road. Freespace is shown with green line. Only measurements inside the detected freespace are used (red crosses).

oriented matched filter will recognize within these search line a pattern of a low-high-low grey value according to a marking measurement in the image. With the aid of stereo depth information these detected positions are projected onto the flat road to determine the required measurement in the vehicle coordinate system described in Section III-A.

Since this projection is very sensitive to the pitch angle of the ego vehicle, typical variations of the pitch angle while driving lead to projection errors. To deal with this a pitch angle estimation based on a V-Disparity approach [10] [11] is performed additionally.

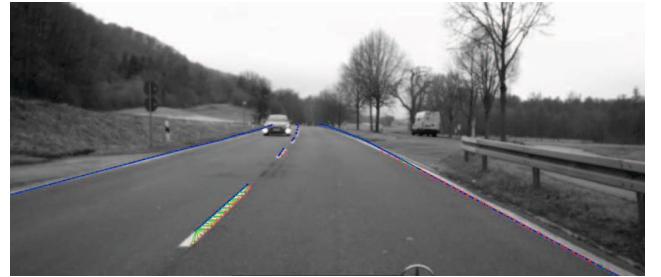


Fig. 10. Marking Detection (red crosses) and Map Matching (green lines) on real world data. The map is drawn in blue lines.

2) *Curb Measurement*: A robust approach to recognize curbs in urban scenarios with high accuracy is proposed in [12]. The intensity image as well as a height profile are

used in the classifier based recognition. Both, position in the image and the probability of existence is obtained as result. Figure 11 shows an example in urban area.



Fig. 11. Curb classifier patches on the right side.

IV. EVALUATION

A. Data set

The localization algorithm is evaluated on two different data sets: A small test track at Daimler's driver assistance proving ground in Sindelfingen and approximately 50 km of rural road around Karlsruhe. The first test track is a round course on a flat testing ground. Road markings are well visible, but it lacks of curbs and other traffic. It is used to prove the basic functionality and precision under good conditions. To achieve best results, the map is completely automatically generated and thus provides highest accuracies. Due to the good visibility of the road markings, no further manually labeling was required (Figure 12).

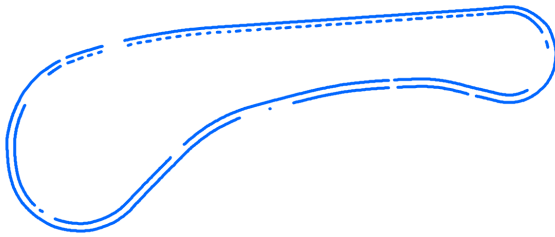


Fig. 12. Automatically generated map of the round course.

The larger data set is recorded on typical suburban areas in Germany. Long parts with rural roads alternate with small towns and urban areas along the road. The route contains typical intersections, roundabouts, underpasses and tunnels. The total length is about 50 km. Typical scenes are shown in Figure 13.

B. Ground Truth

All real data sets have in common, that ground truth data is not available. Measurement errors in the GNSS position due to occlusions, multi-path-effects or atmospheric disturbances exceed the required accuracy of the localization system. Even with postprocessed GNSS data the errors are larger than the expected precision. Therefore the computed position cannot be compared directly to the position of the GNSS unit.

C. Evaluation on Test Track

The aim of the evaluation on the test track is to determine the accuracy of the localization system under good conditions. Since the lane markings are clearly visible, there are no outliers. In addition, no other traffic that could occlude markings is present. Influence of roll and pitch angle is limited by a slow velocity on the flat ground. These conditions also provide an optimal base for an automated and precise mapping process with high map accuracies.

To determine the accuracy of the online localization the mean residual is considered over the entire track course and up to a measurement distance of 15 meters in front of the cameras in each image. In the case of the test track data set, an accuracy of 7.0 cm is achieved. Thus, the result is close to the assumed mapping accuracy.

D. Evaluation of Road Dataset

The larger data set evaluation has been performed under day conditions in a real traffic scenario with a map created for this route beforehand, as presented in Section II. The localization system has been initialized by means of GNSS position at the beginning of the 50 km route. GNSS is no longer used during the rest of the complete drive. It has to be pointed out that a reinitialization was not necessary.

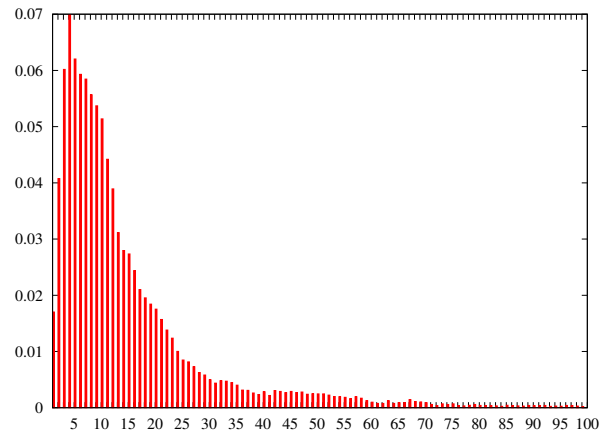


Fig. 14. Histogram of mean residual measurements in cm.

Figure 14 shows the histogram of the measurement residuals over the complete data set. All measurements are regarded up to a distance of 15 meters in order to determine the mean residual given in cm (obtained with stereo reconstruction). The median of all residuals is 11 cm, which corresponds in our system to a median of approximately 11 pixels in the image. While most residuals concentrate on a narrow range around the peak, a reason for higher residuals are parts of the road with only solid lines. In those sections the lateral error increases and large residuals occur when a dashed segment or stopline is reached again.

Figure 15 shows the results considering only the left and right side taken from the point of view of our ego vehicle in driving direction. Generally significantly sharper peaks can be seen, with a median residuum of 19 cm on the left



Fig. 13. Typical scenes of the real traffic data set.

and 10 cm on the right side. The difference in the residuals of the left and right side has two major reasons: First, the measurements are unequally distributed in the image. While on the right side usually a solid line or a curb is visible, the left side often is defined by a dashed line or a curb on the opposite side of the road. Secondly, the high resolution birds eye view for map creation is only available for the driven lane. Curbs or markings outside of these images can only be labeled manually and therefore have a much higher variance in position. Thus the map accuracy is higher for markings close-by the driven trajectory.

Considering the mapping process with the resulting map accuracies of approximately 10 cm, the obtained residuals are in the range of the map accuracies.

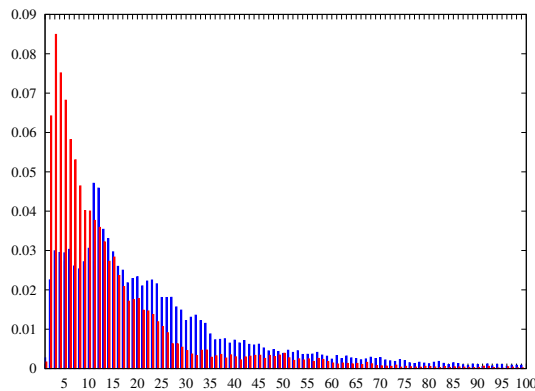


Fig. 15. Histogram of mean residual measurements on the left side (blue) and right side (red) in cm.

V. CONCLUSIONS AND FUTURE WORK

In this proposed approach a novel method for precise and robust localization using highly accurate maps including visible lane markings and curbs is presented. Well known road markings and curb detectors are used to detect markings and curbs online in the vehicle and match them to the map. The localization system uses a stereo camera system and the IMU data of the vehicle only. Thus, modern upper class

vehicles are already equipped with all necessary sensors. The systems works very well in rural conditions. Here a precision of few centimeters can be achieved, which is sufficient to drive autonomously.

Since the localization accuracy is in the range of the map accuracy, future work will concentrate on improvements of the mapping process. With a new camera system images with higher resolution and range are possible and could replace the low resolution Velodyne scans.

REFERENCES

- [1] E. D. Dickmanns, R. Behringer, D. Dickmanns, T. Hildebrandt, M. Maurer, F. Thomanek, and J. Schiehlen, "The seeing passenger car's 'vampers-p'," in *Intelligent Vehicles '94 Symposium, Proceedings of the*. IEEE, 1994, pp. 68–73.
- [2] J. Levinson and S. Thrun, "Robust vehicle localization in urban environments using probabilistic maps," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 4372–4378.
- [3] H. Badino, D. Huber, and T. Kanade, "Visual topometric localization," in *Intelligent Vehicles Symposium (IV), 2011 IEEE*. IEEE, 2011, pp. 794–799.
- [4] H. Lategahn, M. Schreiber, J. Ziegler, and C. Stiller, "Urban localization with camera and inertial measurement unit," in *IEEE Intelligent Vehicles Symposium*, Gold Coast, Australia, 2013.
- [5] O. Pink, "Visual map matching and localization using a global feature map," in *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*. IEEE, 2008, pp. 1–7.
- [6] U. Franke and A. Ismail, "Recognition of bus stops through computer vision," in *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*. IEEE, 2003, pp. 650–655.
- [7] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge Univ Press, 2000, vol. 2.
- [8] E. Dickmanns and A. Zapp, "A curvature-based scheme for improving road vehicle guidance by computer vision," in *Mobile robots: proceedings of SPIE*, vol. 727 (1987), 1987.
- [9] D. Pfeiffer, U. Franke, and A. Daimler, "Towards a global optimal multi-layer stixel representation of dense 3d data," *BMVC, Dundee, Scotland. BMVA Press (August 2011)*, 2011.
- [10] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through," in *Intelligent Vehicle Symposium, 2002. IEEE*, vol. 2. IEEE, 2002, pp. 646–651.
- [11] A. Wedel, U. Franke, H. Badino, and D. Cremers, "B-spline modeling of road surfaces for freespace estimation," in *Intelligent Vehicles Symposium, 2008 IEEE*. IEEE, 2008, pp. 828–833.
- [12] P. Greiner, M. Enzweiler, C. Knoepfel, and U. Franke, "Towards multi-cue urban curb recognition," in *IEEE Intelligent Vehicles Symposium*, Gold Coast, Australia, 2013.