Following Dirt Roads at Night-Time: Sensors and Features for Lane Recognition and Tracking

Sebastian F. X. Bayerl, Thorsten Luettel and Hans-Joachim Wuensche

***Abstract*— The robust perception of roads is a major prerequisite in many Advanced Driver Assistant Systems such as Lane Departure Warning and Lane Keeping Assistant Systems. While road detection at day-time is a well-known topic in literature, few publications provide a detailed description about handling the lack of day-light.**

In this paper we present multiple sensors and features for perceiving roads at day and night. The presented features are evaluated according to their quality for road detection. We generated a large number of labeled sample data and extracted the quality of the features from their probability distributions. The challenge of tracking an unmarked road under bad lighting conditions is demonstrated by comparing receiver operating characteristics (ROC) of the features at day and night-time.

Based on these results we present a road tracking system capable of tracking unmarked roads of lower order regardless of illumination conditions. Practical tests prove the robustness up to unmarked dirt roads under different weather conditions.

I. Introduction

In the last decades autonomous robots became more and more a focus of interest. Both scientific research facilities and car companies are investing considerable man-power into this field. On the way from Advanced Driver Assistant Systems (ADAS) to a completely self-driving vehicle a lot of challenging tasks have to be solved robustly to enable the robot to participate in traffic. One of those issues is environment perception. In order to keep an autonomous vehicle driving on a lane or to give a lane departure warning to the driver, one has to robustly detect the lane's geometry and position. The most common approach are vision systems, but due to the varying appearance of roads and its strong dependency to illumination conditions this is still a challenging task ([1], [2], [3]).

Moreover, little work has been done on tracking roads at night. In this paper we present characteristics of rural roads and sensors in order to perform a robust tracking at day and night. We are especially interested in tracking roads without any kind of boundary markings. To this end, we utilize and extend our road tracking approach shown in [4].

This paper is structured as follows: In the following section some related work on road tracking methods is described focusing on their performance at night. In Section III we show our autonomous vehicle, the sensors it is equipped with and how the sensor measurements are fused into a multilayer terrain map. The features extracted from the sensor data are shown in Section IV, followed by an evaluation in Section V. A tracking system utilizing these features is presented in chapter VI. Finally, we conclude with a short summary and an outlook to future work.

All authors are with the Institute for Autonomous Systems Technology (TAS), University of the Bundeswehr Munich, 85577 Neubiberg , Germany, Contact author email: sebastian.bayerl@unibw.de

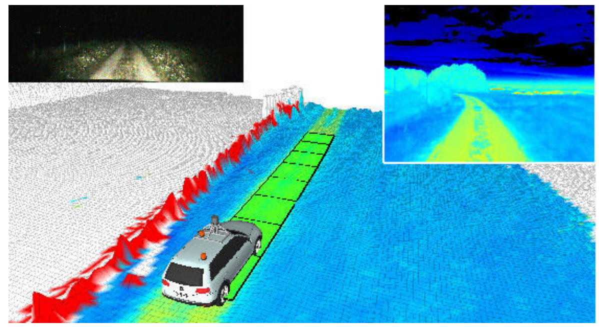


Fig. 1. Tracking a rural road at night using a multi layer terrain map (thermal layer is shown). Red cells indicate obstacle cells. Upper left: Color Night Vision camera, Upper right: thermal camera

Sensor Fusion → Feature Calculation → Particle Filter

Fig. 2. Global Architecture of the Tracking System

II. Related Work

In the past, a lot of different lane recognition systems have been developed. Today perception of high- and freeways work quite well and is already implemented in modern cars for lane departure warning systems. Most of them are based on vision systems and are developed for daytime application. A commonly used recursive filtering technique allows estimating the parameters of road models ([5], [6], [7]) such as clothoids or splines.

In vision-based approaches features like extracted lane-markings, color and/or texture information is used. For example a particle filter approach investigated by *Franke et al.* [8] treats the road recognition as a maximum-a-posteriori estimation task that optimizes the parameters of a road model given an image sequence. For weighting the large number of particles they calculate a joint probability of each hypothesis by considering features such as color saturation, texture and edges. One benefit of this kind of tracking is the easy way to add or remove features as they are treated independently. Expanding on this idea *Manz et al.* [9] developed a hybrid estimation approach that was able to even follow dirt roads. Parameterless systems have also been presented. In [3] the problem of detecting the road was interpreted as finding a minimum-cost path from the lower to the upper image part.

The costs are defined by extracted lane markers, the grass verge (a color based classification) and the free space that was detected by stereo vision.

Although all system proved their strength in difficult scenarios, they all have weakness, too. The vision-only approaches strongly suffer from the disadvantage of their single sensor. Even the best camera does not reach the quality of a human eye, and thus, challenging lighting conditions or less colored environments are limiting the power of this approach to road recognition. Multi-sensor systems combine the benefits of different sensors to achieve a higher level of robustness. In [4] sensor data from color-camera and a LiDAR is fused, accumulated and used for tracking road networks.

One aim of this paper is to show features that indicate the existence and the position of the road at night. An example application of road tracking at night is introduced in [10]. *Serfling et. al* developed a particle filter which is able to estimate position errors inside a digital map by comparing the map-based road information with the sensor data. Each particle represents a potential road course with a different width, robot position and orientation. They suggest measuring road's boundary gradient and its orientation by edge operations on a night-vision image. Due to the different reflectivities of the road and the non-road surface, the road boundary is also visible without lane-markings. Additionally, a lighting independent imaging radar sensor is used for weighting the particles with information about road area and gradient. Another night-vision approach for map-matching is shown in [11]. *Schüle et. al* propose a sensor fusion system that employs digital map information in combination with radar and camera sensors to estimate the 3D road course. This information is fused with the results of an optical lane recognition system. The result is an accurate road course. Road curvature estimation is performed in [12] by using a far infrared camera, a near infrared camera and an imaging radar sensor. *Harmann et. al* are training a Convolutional Neural Network in order to distinguish between three different road curvatures. This kind of detection system is not able to keep a robot on a road, since only a few discrete statements about the lane are made. For our application we need a detailed description of the lane that has to be followed.

III. Sensors

A. Robot and Hardware

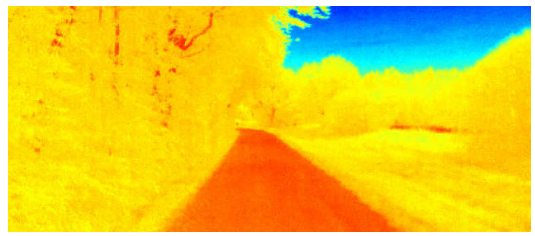
The software we develop is applied to our robot platform MuCAR-3 (Munich Cognitive Robot Car - 3rd Generation), a stock VW Touareg equipped with full drive-by-wire capability. The vehicle motion is estimated by a Kalman filter that fuses data from an inertial navigation system with vehicle odometry to provide jump-free estimates in an inertial integration space.

MuCAR-3 is equipped with different sensors that are beneficial for road detection. The main sensors are listed below:

- A Velodyne HDL-64 laser scanner provides 1.3 million measured points per second. This 3D point cloud provides a lot of information about the terrain profile up to 120 meters in distance. Since the laser-scanner is an active sensor that is not affected by lighting conditions, it is ideally suited for night-time application. Obstacles like bushes, trees or any kind of terrain slopes indicate the boundary of a road. In a flat-world scenario however, one is not able to detect a road using this sensor only, so we suggest using additional types of sensors.
- The second type of sensors are default color cameras. This very common sensor is used in a lot of previous work (e.g. [8], [9]) for road detection, but suffers from its sensor-typical limitations: At low light conditions the camera needs a lot of time for image exposure. High integration times leads to smearing in high dynamic scenarios. As a result we limit the integration time and take less illuminated images into account.
- The application of a color night vision (CNV) camera provides more information even in less illuminated areas. Thanks to larger pixel size and other low-light technologies this camera is better adapted for tracking roads at night.
- The last sensor is a thermal camera utilizing an uncooled microbolometer. The thermal camera is able to measure the different surface temperatures of the terrain independently of the lighting conditions. Due to the different materials of the surface the corresponding temperature is different. Figure 3 (c) gives a visual impression of the surface temperature of a dirt road. The temperature of the road is higher than that of the non-road area.

(a) Color camera with 100 ms integration time

(b) CNV camera with 50 ms integration time



(c) Thermal image scaled from 9°C (blue) to 13°C (red)

Fig. 3. Camera images showing a dirt road. Headlights are enabled.

B. Sensor Fusion

In a fusion step the high amount of sensor data is combined in low level manner. We suggest to use the fusion algorithm according to [13] and [4] and extend them by a temperature and a NIR layer. We do not want to go into detail, but the main ideas can be described as follows. A local terrain map containing multiple layers of information is introduced and built as a Cartesian grid with a cell size of $0.2\,\mathrm{m}$. (A multi layer map including color, height and obstacle information is shown in Figure 4 (a).) The aim of this terrain mapping is to produce a dense local representation of the environment, making use of all data the sensors provide. Hence, we accumulate the measured data from several scans and overcome the limitation of the limited field-of-view (FOV) of the cameras and the vertical resolution of the LiDAR. We consider occluded areas, which are not visible to the camera, as well as overhanging structures, which can be interpreted as non-obstacles. The fusion step is passing several maps (such as obstacle, height, height-difference, color and thermal map) to the feature extraction.

IV. Features

In order to detect and track a road several features are defined. This chapter includes their detailed description, whereas their evaluation is part of the following chapter V. According to the nature of their main sensors they can be split up in three groups: color, 3D and thermal.

A. Color

The first group of features is based on the color information of the terrain map.

- The first features we can extract from the color map are edges, since we expect a change in color at the lane boundaries. Therefore we apply one edge operator for each direction (x and y) in order to get the edge intensity $v_{\mathrm{EI}} = \sqrt{v_{\mathrm{EI,x}}^2 + v_{\mathrm{EI,y}}^2}$. The change in color should be perpendicular to the direction of the road. As a result we use the corresponding edge phase, that can be calculated consequently with $v_{\mathrm{EP}} = \operatorname{atan}\left(\frac{v_{\mathrm{EI,y}}}{v_{\mathrm{EI,x}}}\right)$.
- The green-ratio feature assumes that the road area has a low ratio of green color compared to the other color channels ($\frac{g}{r+g+b}$). Since slightly vegetated forest roads do have a high green ratio, we also consider the area below the robot. There is no additional sensor to observe this area, but the cells are taken from the accumulated map (Section III-B). Our algorithm adapts a reference value online and calculates the cell's green-ratio feature accordingly.
- According to [9] the color of frequently-driven roads tends to have a low saturation value in the HSV color space. Because the saturation channel of the HSV space is independent of the illumination intensity, the feature depends less on shadows.
- Assuming that the vehicle drives on a road, we compare all color information inside the terrain map with the cells below the robot. The accumulation of sensor data (see Section III-B). One method to do the comparison is to setup a histogram of RGB-color values below the vehicle and perform a histogram back projection to the complete terrain map. This color comparison can be executed also by using another color space. We suggest utilizing the YUV color space, by calculating the mean and standard deviation of its components inside the area below the robot. The feature value of a cell results from the Mahalanobis distance to the color values below the vehicle. Since the Y component represents the illumination, which possibly might have negative influence at scenarios with challenging lighting conditions, it can be neglected.
  These features should never be used alone, since the robot will not able to rejoin a road after having left once.

B. 3D Features

The second feature group refers to the 3D texture of the surface. Their main source of information is the LiDAR, but it can be driven with any system capable of providing 3D data.

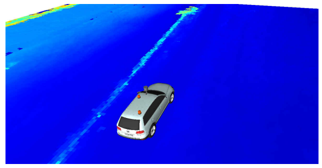
- We expect roads to be free of static obstacles. This implies that occupied cells are definitely non-road cells and areas of high obstacle density tend not to be roads.
- The obstacle detection does not interpret a cell as an obstacle if the cell is located inside a low terrain slope. But even small slopes indicate the position of a road. Areas with a high variation in slope have a low road probability.
- The terrain height can be introduced in a similar way to the slope. The cell's information about its height is an indicator of non passable obstacles or areas that are not nice to drive on. The higher the difference between the cell's height and the robot plane, the more likely the cell is to be non-road.
- We use a classification of vegetation such as grass or bushes by analyzing the frequency of the distance signal of each laser scanner diode. This method allows us to distinguish between road and grass independently of light. This information, however, should not be a hard decision criteria, since a lot of dirt roads or forest roads are slightly covered with grass.

C. Thermal

The last feature group uses information provided by the thermal camera. In most cases the base material of the road is different to the non-road material. This causes a different heating of the surface at day-time and thus we expect differences in the temperature profile of road and non-road area. Our practical tests have confirmed this and we were even able to observe thermal differences between the white markings of a road and the road itself. See Figure 3. In order to use this we introduce the following two features:

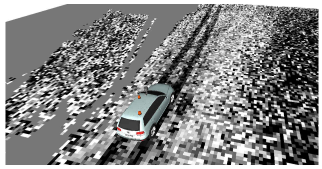
- Similar to the edges in color, we expect edges in temperature located at the road boundaries. For this reason we apply edge operators to the thermal layer

(a) 3D color and obstacles (red)



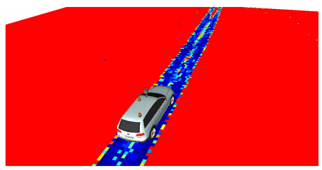
(b) Heights

(c) Edge intensity



(d) Road saturation

(e) Thermal edge



(f) Thermal backprojection

Fig. 4. Features of road scene shown in Figure 1

of the terrain map and get measurements of the edge's intensity and direction.

- The last important observation is the homogeneous temperature distribution along the road. Based on this we calculate the temperature statistics under the robot and compare it to all cells of the terrain map. If we assume that the robot is already located on the road, we can extract cells with similar temperature and the road itself.

V. Feature Evaluation

In the following chapter we analyze the quality of the presented features and state which features are beneficial under challenging night-time lighting.

A. Data Generation

The generation of feature data is performed by applying the road tracker from Section VI to several recordings of sensor measurements. Given the tracking result, we can generate pairs of positive and negative road samples. The positive is defined by the road tracking result itself and the negative is built by falsifying the road geometry (clothoid parameters, such as displacement or curvature) of the positive one. Given the road geometry we are able to extract the sample's features - e.g. edges at the road boundary. For each sample all the feature values are finally stored with a label (positive or negative) in memory.

The different sensor recordings include different seasons (summer, winter, ...) and types of roads (paved, unpaved, dirt, forest). Each track was driven at day- and at night, so we can evaluate which features are still beneficial at less illumination. We were able to generate at least 10000 night samples and the corresponding 10000 day samples.

B. Comparison

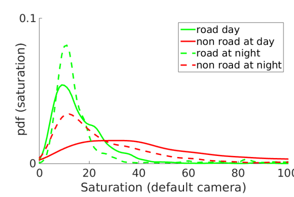
In order to perform this comparison, we define criteria which represent the feature's quality. We suggest taking a look at the distribution of the labeled feature values and the Receiver Operating Characteristic (ROC) of the features. The influence of day and night-time causes differences in the ROC of a feature.

In general the ROC's curve is created by plotting a feature's true positive rate against the feature's false positive rate at various threshold settings. High true positive rates at low false positive rates mark the most powerful features.

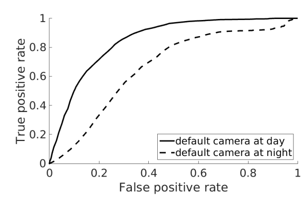
The ROC of one single feature was generated as follows: the feature's values of the collected data are compared to a threshold and classified as road or non-road. Since the data is labeled, we can identify false positives, false negatives, true negatives or true positives. We sample the threshold and get different numbers of the false positives and true positives for each threshold.

The following specifies the different Probability Density Functions (PDF) and ROC curves for day and night time. We do not describe all the features in detail, but we select the most representative ones.

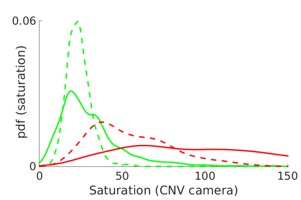
- Evaluating the color features of the road, we noticed - as expected - a reduction of quality. In all ROC curves the degradation of the features based on the default color camera is much higher than to the CNV camera. In some cases (default camera's saturation or green ratio, Figure 5 (b) and (f)) the ROC curve approximates to the $y = x$ axis, which represents a maximum uncertainty.



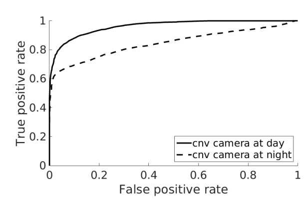
(a) Default saturation distribution



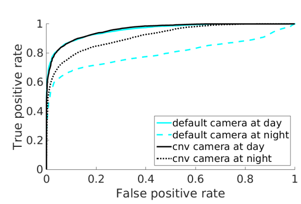
(b) Default saturation ROC



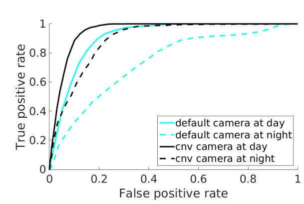
(c) CNV saturation distribution



(d) CNV saturation ROC



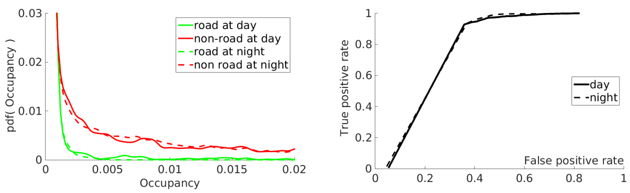
(e) Edge intensity ROCs



(f) Green Ratio ROCs

Fig. 5. PDF of color features

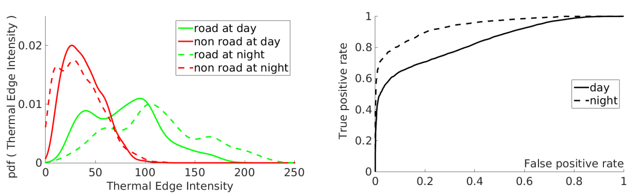
- The active LiDAR measurements allow the the 3D features to be independent of illumination. The distribution of the features is almost the same. For example see the occupancy feature at Figure 6. Also the ROCs do not change from day to night. Thus, they can be used unrestricted for tracking at night. Unfortunately a road cannot be detected by 3D features only in general. The low true positive rates at low false positive rates can be interpreted as follows: The 3D features contain the implicit information about non-road area, but no direct information about roads.



(a) Obstacle feature distribution (b) Obstacle feature ROC

Fig. 6. PDF of 3D features

- The last group to evaluate are the thermal features. Transitions at the road boundary are observable at each time of the day. This group of features has even a better performance at night because there is no direct influence of the sun. Shadows generate different surface temperatures on one surface material.



(a) Thermal edge distribution (b) Thermal edge ROC

Fig. 7. PDF of thermal features

The thermal features depend on weather conditions. Rain and fog cool down the surface and make the road area and the non-road area having less difference in temperature. In Figure 8 the ROC of the thermal edge direction is presented for dry and wet scenarios. The temperature ROC has also a strong dependency to the quality of the road. This dependency is even stronger than the influence of weather. A example of poor structured road can be seen in Figure 10.

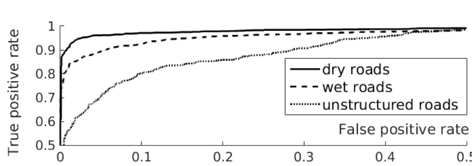


Fig. 8. Evaluation of thermal edge direction feature with different weather conditions and road quality

VI. Tracking System

In this paper we use an extended version of the tracking method suggested by [4]. This method is based on a particle filter, which projects its road hypotheses into the local terrain map and weights them according to a set of road features, similar to the presented ones. The main extension of our method is the increased number of road features and the usage of further sensors. This increases the robustness of tracking and we are able to perceive road networks even at night.

A. Particle Filter

Similar to [4] we use two different road geometries to model the road network. A clothoid with a fixed width is used as a road model. The intersections are constructed by multiple clothoids starting at a common origin. The algorithm switches between the different models with the help of a digital road map. All coordinates of the system state are kept relative to the robot's ego coordinate system. The filter can be split up in two main parts:

- A prediction step uses the estimation of the robot's motion in order to update the relative coordinates of the road and intersection (see [9]).
- In the following measurement step we rebuild the probability distribution of the state vector $\mathbf{x}_p$ by assigning a specific weight to each predicted particle. The filter projects the geometric models of all particles into the local terrain map and measures their road quality. This quality q_p is estimated by a Naive Bayes classifier, that compares n measured feature values to a trained feature distribution of roads or uses a manual generated heuristic which again depends on the measured feature values. In a last step the state vector and its covariance are generated as the weighted mean of the best particles.

The calculation time of the algorithm is small enough to run on an ordinary machine (Intel i7 2600K) in sensor rate, which is given by the Velodyne LiDAR as slowest sensor at 10 Hz. Figure 1 is showing the estimation of the road tracker.

B. Effects to Tracking Quality

In this section we want to demonstrate the benefit of the additional night sensors and features and compare it to the previous method introduced in [4].

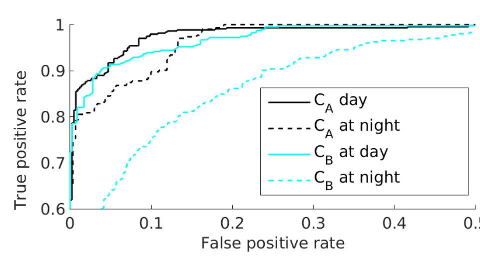
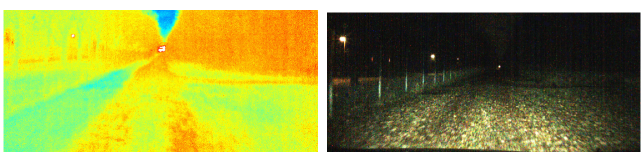


Fig. 9. ROCs of classifier with full and reduced feature capability

Two different Naive Bayes classifiers are trained: one with full feature capability (classifier C_A) and a second with the reduced number of features (classifier C_B). The second reduced classifier represents the classification power of the method presented in [4] and is only trained with default color camera features and without any information of night sensors. At day the performance of C_A is only slightly better than C_B. At night the classifier show a different degradation of their performances: C_A is influenced only a bit because it supports all the presented features. The ROC of C_B is influenced very much since it is limited to the default color camera.

C. System Limitations

A bad classification result indicates that the system is operating near the limit of its perception. One key role for our robust tracking system is the clear separation of road and non-road cells by a couple of features. The system has reached its limits if the sensor data does not indicate any existence of a road. For example in a flat environment the 3D point cloud of a laser scanner cannot be used to determine the position of the road. A similar case is given by the thermal camera and a homogeneous temperated surface. Here, the sensor does not provide useful information. Also a color camera is not a perfect sensor. One is not able to visually extract the road in a uniformly colored terrain. One example of a challenging scenario is show in Figure 10, where the road is covered by leaves. Since the particle filter uses all of the above mentioned features, the system is able to compensate sensors, which do not provide beneficial information. At least one significant feature is necessary to make tracking valid.



(a) Thermal camera (b) CNV camera

Fig. 10. Challenging road scene

VII. Conclusion and Future Work

A. Conclusion

In this paper we describe several sensors and their relevance for road tracking at night-time. In a first step color cameras, a thermal camera and a LiDAR are fused into a multi layer local terrain map. Based on the information of this grid map, we generate a variety of features and evaluate them according to their relevance for road recognition. Obviously, the active sensors are not influenced by lighting conditions. Also the features which are based on thermal data are very informative. The data provided by the color camera gets worse with less light, but is still partially usable. Finally we presented a particle filter based tracking algorithm that utilizes the described features for a robust recognition of road networks.

Some visual impressions of the features and the presented tracking system at night can be seen at www.mucar3.de/iros2015-roadtracking.

B. Future Work

The modular weighting step of the particle filter can be extended easily by new features. The basis of a possible feature extension is given by the measured reflectivity values of the laser scanner. Since our focus is in less structured regions without any artificial boundary markings, we expect different surface reflectivities between road and (for example) grass, but no extreme difference as between paved road and its white lane markings.

Acknowledgments

The authors gratefully acknowledge funding by the Federal Office of Bundeswehr Equipment, Information Technology and In-Service Support (BAAINBw).

References

[1] T. Kühnl, F. Kummert, and J. Fritsch, "Spatial ray features for real-time ego-lane extraction," in *Proceedings of IEEE Intelligent Transportation Systems (ITSC)*, Anchorage, AK, USA, Sept. 2012, pp. 288 – 293.

[2] G. B. Vitor, A. C. Victorino, and J. V. Ferreira, "A probabilistic distribution approach for the classification of urban roads in complex environments," in *Workshop on IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, China, May 2014, pp. 1–6.

[3] J. Beck and C. Stiller, "Non-parametric lane estimation in urban environments," in *Proceedings of IEEE Intelligent Vehicles Symposium (IV)*, Dearborn, MI, USA, June 2014, pp. 43 – 48.

[4] M. Manz, M. Himmelsbach, T. Luettel, and H.-J. Wuensche, "Detection and Tracking of Road Networks in Rural Terrain By Fusing Vision and LIDAR," in *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, San Francisco, CA, USA, Sept. 2011, pp. 4562–4568.

[5] E. Dickmanns and A. Zapp, "A Curvature-based Scheme for Improving Road Vehicle Guidance by Computer Vision," in *SPIE Conference on Mobile Robots*, Cambridge, MA, USA, Oct. 1986.

[6] C. Jung and C. Kelber, "Lane following and lane departure using a linear-parabolic model," in *Proceedings of IEEE Intelligent Vehicles Symposium*, Nov. 2005, pp. 1192 – 1202.

[7] H.Loose and U.Franke, "B-Spline Based Road Model for 3D Lane Recognition," in *Proceedings of IEEE Intelligent Transportation Systems Conference*, Madeira, Spain, Sept. 2010.

[8] U. Franke, H. Loose, and C. Knoppel, "Lane Recognition on Country Roads," in *Proceedings of IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, June 2007.

[9] M. Manz, M. Himmelsbach, T. Luettel, and H.-J. Wuensche, "Fusing LIDAR and Vision for Autonomous Dirt Road Following – Incorporating a visual feature into the tentacles approach," in *Proceedings of 21. Fachgespraech Autonome Mobile Systeme (AMS)*, Karlsruhe, Germany, Dec. 2009.

[10] M. Serfling, R. Schweiger, and W. Ritter, "Road course estimation in a night vision application using a digital map, a camera sensor and a prototypical imaging radar system," in *Proceedings of IEEE Intelligent Vehicles Symposium*, Eindhoven, The Netherlands, Nov. 2008.

[11] F. Schüle, R. Schweiger, and K. Dietmayer, "Augmenting Night Vision Video Images with Longer Distance Road Course Information," in *Proceedings of IEEE Intelligent Vehicles Symposium*, Gold Coast, Australia, June 2013.

[12] O. Hartmann, R. Schweiger, R. Wagner, F. Schüle, M. Gabb, and K. Dietmayer, "Night Time Road Curvature Estimation based on Convolutional Neural Networks," in *Proceedings of IEEE Intelligent Vehicles Symposium*, Gold Coast, Australia, June 2013.

[13] S. F. X. Bayerl and H.-J. Wuensche, "Detection and Tracking of Rural Crossroads Combining Vision and LiDAR Measurements," in *Proceedings of IEEE Intelligent Transportation Systems Conference*, Qingdao, China, Oct. 2014.