Article

Towards HD Maps from Aerial Imagery: Robust Lane Marking Segmentation Using Country-Scale Imagery

Peter Fischer [1,*], Seyed Majid Azimi[2], Robert Roschlaub[3] and Thomas Krauß[2]

[1] AUDI AG, 85045 Ingolstadt, Germany

[2] German Aerospace Center,82234 Oberpfaffenhofen, Germany; seyedmajid.azimi@dlr.de (S.M.A.);
   Thomas.Krauss@dlr.de (T.K.)

[3] Bavarian Agency for Digitisation, High Speed Internet and Surveying, 80538 Munich, Germany;
   Robert.Roschlaub@ldbv.bayern.de

[*] Correspondence: peter2.fischer@audi.de

Abstract: The upraise of autonomous driving technologies asks for maps characterized by a broad range of features and quality parameters, in contrast to traditional navigation maps which in most cases are enriched graph-based models. This paper tackles several uncertainties within the domain of HD Maps. The authors give an overview about the current state in extracting road features from aerial imagery for creating HD maps, before shifting the focus of the paper towards remote sensing technology. Possible data sources and their relevant parameters are listed. A random forest classifier is used, showing how these data can deliver HD Maps on a country-scale, meeting specific quality parameters.

Keywords: autonomous driving; HD maps; aerial imagery

1. Introduction

Within the last decade, the research on fully autonomous vehicles (AV) has progressed so much that what started as a vision is quickly becoming a reality. Nowadays it is recognized that the increasing usage of AV will have beneficial influence on a broad range of socio-economic and transportation factors in our society, for instance energy consumption and air pollution, land use, and car ownership. The increasing usage of AV will therefore be beneficial not just for the single owners of the cars, but also for society as a whole [1]. A recent overview about challenges and opportunities concerning AV is given in [2]. Besides of Mobility on Demand (MoD) and car-sharing, AVs are probably the key challenge for future Intelligent Transportation Systems (ITS).

Nowadays cars, both Commercial-of-the-Shelf (COTS) and research platforms, are equipped with a broad range of different sensors, e.g., camera-based systems, Lidar, Radar, Ultrasonic, Infrared, Odometry sensors and Global Navigation Satellite System (GNSS) receivers. Advanced Driver Assistance Systems (ADAS) make extensive use of the information acquired from these sensors to accurately estimate the vehicles position and pose, and to actively map its surroundings. The combination and harmonization of the permanently recorded heterogeneous data stream is a key factor for achieving a reliable and consistent view of the vehicles in a close-range environment. The perception of the environment remains the biggest challenge to solve, in order to achieve a reliable, smooth and safe driving [1]. Weather conditions like rain and fog, rapidly changing illumination caused by shadowing, or a corrupted GNSS positioning signal caused by multipath effects, are just examples of what has to be taken into account for real world applications. In addition to hardware components, algorithms play a key role for AVs. Publicly available benchmark datasets

have therefore been published in the past to provide a reliable source to evaluate and compare algorithms performance in environment perception [3].

1.1. HD Maps for Ego Positioning

A High Definition Map (HD Map) can contribute to this task by spanning an expectation horizon of what should be seen and at which position objects should be located [4]. Concerning GNSS-based positioning within an urban environment, especially if a car is traveling along an urban canyon, signal reception is affected by strong biases caused by multipath effects and atmospheric propagation delays. Obst et al. [5] therefore propose the usage of a 3D digital map to overcome these issues for vehicle localization. Within this map, potential multipath situations can be predicted and identified through a probabilistic ray-tracing model. Tao and Bonnifait [6] proved that both the tight and loosely coupling of GNSS with 2D lane marking maps may increase the accuracy and integrity of the positioning information, particularly when these are computed autonomously. The same authors also implemented a road invariant Extended Kalman Filter which makes use of Lane Markings for the enhanced estimation of GPS errors. Using images and HD Maps has been proven to be especially reliable at reducing the lateral error in GNSS positioning [7]. A recent example is given in [8], based on the data-set provided by the University Grand Challenge at ITS World Congress 2016 in Melbourne. Bauer et al. [9] used a HD Map derived from georeferenced aerial orthophotos with a spatial resolution of 10 cm, and digital elevation and surface models with resolutions of 2 m and precisions of 20 cm for precise urban vehicle localization. The experiments showed a significant positioning improvement pinpointing out the high potential of HD Maps.

In situations where no GNSS signal is available, the impact of HD Maps for positioning raises significantly. Schreiber et al. [10] proposed an approach which is completely independent of GNSS and solely based on highly accurate maps. This focuses on the precise localization relative to the given map that is necessary for path following, with experiments showing that localization accuracy is in the range of the map accuracy: therefore, they conclude that future work should focus on the mapping process. Burgard et al. [11] pointed out that for an AV the localization requirements are in the order of decimeters, which makes GNSS alone insufficient to meet these requirements. The experiments reported therein using a HD Map resulted in a relative lateral localization error to the HD Map which was mostly within 5 cm.

1.2. HD Maps for Scene Understanding

Besides the contribution to a highly accurate vehicle positioning with and without GNSS, HD Maps can help in holistic 3D scene understanding. Wang et al. [12] generated 3D HD Maps using volunteered geographic information from the OpenStreetMap project. Their 3D object detection algorithm was validated on the KITTI dataset [3], outperforming the baseline in all metrics and tasks.

As a result of the unpredictable activity of pedestrians, bicyclists and traffic movement, travelling urban environments with AVs is probably the most challenging scenario. The DARPA urban challenge 2007 can be marked as a key event addressing such a scenario. Several research groups from around the world participated in this challenge, with the team AnnieWAY reporting about their experiences [13]. Whilst no prior HD Map was available, their system combined the spatial information from the sensors to generate a 2D map of the environment on the fly. More specifically, this map was centered at the vehicles position and moving with it. For environmental perception an occupancy grid with a spatial resolution of 15 cm was generated.

1.3. Experiments on Public Roads Using HD Maps

Almost all car manufacturers are nowadays carrying out experiments with modified vehicles on public roads demonstrating their current abilities. One of the most remarkable experiments was carried out by Daimler in 2013, in which a modified Mercedes Benz S 500 traveled the so-called Bertha Benz Memorial Route, a 103 km long route in the southwestern part of Germany connecting

the cities of Mannheim and Pforzheim. Unlike the majority of other research vehicles, the car comes with no costly LiDAR. The HD Map played an important role to ensure localization. The subsystem Localization comprises the modules (a) Lane-Marking-Based Localization, (b) Feature-Based Localization and (c) Localization Fusion, where (a) and (b) are complementary vision algorithms. The used HD Map contained significantly more information than typical navigation maps. All static environmental properties which are necessary for driving but cannot be reliably detected by sensors were stored, e.g., the layout of drivable lanes [14].

The BMW Group also reported about experiences, results and lessons learned from automated driving on Germany's Highways, where the main routes are around Munich and the highway between Munich and Nuremberg, better known as DigitalesTestfeld Autobahn. The used HD Map was generated from a small number of vehicles equipped with the same sensors as the later used AV. It was pointed out that the data had to be harmonized and that the HD Map comprises two layers: a semantic geometric layer and a localization layer. For more details concerning BMW's experiences, we refer to the prospective published in [15].

There is a broad agreement on HD Maps being a key technology for AVs [4]. Both commercial players, Daimler and BMW, pointed out that the generation and provision of HD Maps is currently one of the main challenges when it comes to autonomous driving. Aeberhard et al. [15] explicitly asked for validation and certification of large-scale digital maps. Ziegler et al. [14] stated that for the generation of digital maps not only the technical performance, but also the scalability of the chosen solutions in terms of a commercial roll-out is pivotal.

1.4. Descriptive Parameters, Metrics and Content of HD Maps

Here, we introduce needed definitions and clarifications. Whilst all the aforementioned papers and experiments deal with HD Maps, the content of such maps and the meanings of the given accuracy statements are quite inhomogeneous. Concerning the contents, the lowest common denominator is that all HD maps featured a lane markings detection step. Focusing on the term accuracy, there's a rich literature dealing with its specification for mapping products, and giving guidelines on how these quantitative parameters should be derived, which perfectly works also nowadays. Macfarlane and Stroila [16] from HERE addressed the wording Uncertainty in the domain of autonomous driving, as from a mapper's prospective especially (a) Uncertainties in Object Detection and Localization and (b) Temporal Map Uncertainties are of interest. For (a) we have to distinguish between relative and absolute accuracy, where both quantities can be given. In most papers it is stated that a relative accuracy of 10 to 20 cm was used, whereas almost no statements are given concerning the absolute accuracy. This may be reasonable from a practical point of view where the most important issue is to ensure a stable relative positioning of the AV with respect to its surroundings. Nevertheless, the absolute accuracy is the most meaningful quality measure when it comes to spatial applications, therefore it should not be simply neglected. For (b) a mapping product relies always on a given date of data acquisition. As the road network is recognized as a highly vital environment, where changes can appear everywhere at any time, the update rate and the time stamp stating the time an object was detected play an important role.

To the best of our knowledge, most map makers currently use data recorded by probes or a fleet of vehicles. Massow et al. [17] presented an approach on how to handle, maintain and exploit the big amount of data recorded by probes. Furthermore, three different data layers were identified: (a) the road geometry, (b) the road furniture, and (c) dynamic data. The road geometry and its surrounding furniture can be considered as base map where an absolute accuracy statement is useful, whereas dynamic data such as information about accidents do not need to be included. The authors furthermore developed a Map Evaluation Tool (AET), which performs a metric evaluation on the absolute positioning accuracy of the map. Dabeer et al. [18] presented recent results concerning an end-to-end system for crowdsourced 3D maps for autonomous driving. Their error metrics also include absolute positioning accuracy.

1.5. HD Maps and Aerial/Satellite Imagery, Literature Review

Aerial images have been used for a long time for the extraction and description of transportation infrastructure. Kim et al. [19] proposed the extraction of road feature information by combining several image processing algorithms into a sophisticated system. The authors proposed transfer learning approach using fully convolutional neural networks (FCNNs) to extract the roads firstly and afterwards with the segmentation of the lane marking in the masked image using the first step. This approach has been common in the majority of the current methods to extract lane-markings in remote sensing. The main contribution of the authors is the usage of FCNNs instead of the traditional methods.

In the algorithm proposed by Jin et al. [20] the roads are extracted firstly and afterwards Gabor filters are applied in order to highlight for the lane markings. This step is followed by the thresholding algorithm of Otsu to achieve binary segmentation. The outputs are then produced by either morphological operations or support vector machines (SVMs).

Despite obtaining good results in highways, this approach yields some false positives in the objects represented by white linear features, such as the boundaries of bridges over highways. Also lines belonging to vehicles might be misclassified as they are inside the road regions. In addition, lane-marking extraction was not investigated into detail and only one output image was provided. In their work, objects like trees over highways and roads or not well-visible lane markings on the surface of roads were mentioned to be a main reason leading to a perturbation of the final results. To address the mentioned problem, Jin et al. [21] proposed a newer approach consisting of the centerline of the road extraction step, the road surface detection and in the end pavement markings extraction. However, this approach is also heavily dependent on the road-mask components. Despite having superior performance over previous methods, this approach exhibits low accuracy with regards to lane-marking detection without road mask.

Jin et al. [22] used hierarchical image analysis based on color component analysis and wavelet transformation to detect the position of roads in a first step, then applied Gabor filtering to extract road lane markings. The authors used an unsupervised algorithm to segment the road area in the first. Afterwards, a contrast measurement was deployed in order to improve the quality of the lane-marking extraction assuming that the difference between the pixel intensity of lane-markings and the surface of road is noticeable. In the next step, morphological features were applied to smooth the edges in the shadow areas. Finally, the extracted lane-markings were improved using a modified Wang-Zangen algorithm as well as a line fitting step using least square regression. Despite an extension of lane-marking to rural areas, without road area masking step, this method would suffer from a considerable amount of false positives.

Unlike the two aforementioned works, Hinz and Baumgartner [23] proposed a methodology for urban areas using multi-view imagery as well as the usage of context information. The author used the detected thin lines as a sign for the existence of a road. Despite of having very good performance, the proposed method requires several images from different view angles of an area of interest. Mnih and Hinton [24] also focused on urban areas, but used unsupervised classification for feature extraction and neural networks to finally detect roads in a reliable and robust manner. In addition to methods relying only on aerial imagery, the usage of additional images collected by vehicles gave promising results. Mattyus et al. [25] gave an example of the combined parsing of both ground and aerial images to generate HD Maps and to extract road topology from aerial imagery [26]. Their methods is based on Markov Random Fields as well as the combination of parsed aerial and ground imagery data in order to produced detailed results. Even though, these road masks could be utilized for lane marking localization, they can't be deployed directly to improve the lane marking extraction. More complete overviews about the extraction of roads and road features from aerial imagery are given in [27, 28].

The main advantage of airborne remote sensing is its ability of covering large areas in short time within a reasonable budget. This point addresses the need to provide a scalable map production tool chain which Ziegler et al. [14] asked for. In the case of road networks, special advantages with respect to ground based systems are; (a) the independence concerning actual traffic

situations (e.g., traffic jams, acceleration/braking for following traffic participants), (b) the point of view where occlusions caused by ground-based obstacles are minimized, and (c) the broader field of view where also features not immediately in line-of-sight by vehicles can be detected. Aerial imagery has furthermore proven its ability in serving as a reliable source for retrieving road related features (e.g., lane markings, land marks) which can be used for the generation of HD Maps.

We propose to generate at least static basic HD Maps using aerial imagery giving the absolute accuracy as spatial quality measure. A not yet outdated overview on how such quality measures are generally given for mapping products is given by Greenwalt [29]. These basic HD Maps with a given absolute accuracy can then act as reference point for further dynamic map contents, recorded by fleets of probes. The content of such basic HD Maps are road markings, which are already earlier identified as the smallest common denominator of the HD Maps used in the referenced works. They are easily to identify in aerial imagery given a suitable spatial resolution and are the main describing features of nowadays road networks, where most or even all further information of interest for AVs can be derived from a complete lane marking HD Map.

Following the mentioned works, the approaches by Javanmardi et al. [30] and Huang et al. [31] applied adaptive threshold in aerial imagery. Javanmardi et al. [30] approach consists of several steps which include the processing of digital surface model (DSM), the removal of vehicles by multiple images and finally the utilization of an adaptive thresholding in order to extract lane marking. As it can be seen, lane markings are not extracted directly and the usage of third party data is necessary for the removal of non-lane marking objects.

Tournaire and Paparoditis [32] proposed dashed-line and zebra crossing extraction by using the obtained information of the primitives extracted from the images together with a reconstruction process. They only addressed the rectangular shape lane markings and their geometric properties were studied to extract them.

In order to grasp the hidden pattern in the features of the lane marking via an end-to-end feature learning approach such as deep learning ones, Azimi et al. [33] proposed Aerial LaneNet approach to extract lane marking by learning them. They introduced the first lane-marking segmentation dataset in the remote sensing area called AerialLanes18. They utilized a modified fully convolutional neural network and modified it to be able to extract tiny lane-markings from aerial images which are hard to detect for convolutional neural network given the fact that these methods suffer to detect small objects and the lane-markings appear to be very small. They enhanced the accuracy of their approach by using wavelet transforms combined with convolutional neural network and used a custom loss function to address the issue of unbalanced ratio between lane markings and non-lane markings in their dataset. They cropped high-resolution aerial images into patches to be fed into their methods and in the end; they are stitched together based on the specific policy. In our work, we did not compare our results with the results of this method.

Lane marking extraction has also been studied in in-situ imagery. Lee et al. [34] proposed a convolutional neural network (CNN) based with multi-task purpose for joint lane-marking extraction and classification given the fact that their images contained different weather and lightning conditions. Gurghian et al. [35] proposed a CNN-based method for lane-marking extraction on right and left sides of a car. Even though, these methods achieve a very high accuracy, but their main focus is not on the small size of lane-markings. As mentioned, lane-markings have a tiny size in aerial imagery and should be given a special attention.

1.6. Aim of This Paper

This contribution tackles the overlapping topics of HD Maps, Computer Vision and Remote Sensing. The whole process from orthorectified image data towards the final HD Map is described, with a special focus on image segmentation and pattern recognition for identifying the relevant features describing the transportation system and its nearby environment.

The possibility of using aerial and very high resolution satellite imagery as data source for generating HD Maps is discussed. Section two describes shortly the used data. A methodological

view by means of image classification is given in Section 3. The effectiveness of the proposed algorithm is tested, results are presented and discussed in Section 4. We conclude in Section 5.

2. Materials/Image Data

In contrast to imagery collected by vehicles, orthorectified aerial and satellite imagery comes with one or several descriptive parameters, easing the mapping process based on such images. The Ground Sampling Distance (GSD) gives a strict connection between pixel spacing and real world distances. Some camera systems offer the possibility of collecting imagery with a side-looking view, and the viewing angle is given. Another interesting parameter given for most aerial imagery products is the acquisition date/time. In combination with a surface model which is in most cases available for optical flight campaigns, shadow masks which are extremely useful for several subsequent image processing steps can be computed.

Some countries distribute aerial imagery acquired by their governmental institutions under a free-and-open data policy. For instance, the United States Geological Survey (USGS) offers aerial imagery of the whole United States with GSD's down to the decimeter level. European entities, for instance the German Federal Mapping Agencies, now tend to follow this trend: the federal states of Hamburg, Berlin and North-Rhine Westfalia run open-data web portals were orthophotos with a spatial resolution of 20 cm are available and distributed under a free-and-open data policy. We used digital orthophotos with a GSD of 20 cm provided by the Bavarian Agency for Digitisation, High-Speed Internet and Surveying. The data were derived from the Bavarian Flight Program with an image overlap of 80% along track and 50% cross track. Since 2017, the actualization was decreased from a 3-year- to a 2-year-cycle. In the current Bavarian Flight Program 2018, the southern part of Bavaria (appr. 40.000 sqkm) will be covered with a GSD of 20 cm. The contractors use common digital high resolution sensors e.g., UC Eagle or DMC II/III. The parameters of the spectral resolution are 5 channels (PAN, R, G, B, NIR) and a color bit depth of 16 bit.

In 2018 the generation of orthophotos will change to true-orthophoto [36]. The basis for the generation of true-orthophotos is the orthogonal projection with a Digital Surface Model (DSM) from aerial imagery in a grid spacing of 20 cm. True-orthophotos avoid occluded areas or displacements from e.g., buildings—an important advantage in inner cities.

The imagery comes in a tiled fashion where the tile size is 5000 × 5000 pixels. In the experiments seven tiles where used for training (train/validation equals 70%/30%) and three for testing. The total length of the motorway depicted in the imagery is greater than 10 km. As the two driving directions are divided into two roads, with both having a minimum of three lanes, the total lane is longer than 120 km. The imagery depicts a part of the motorway A9 between Ingolstadt and Munich, which is publicly well known as Digital Motorway Test Bed, operated by the German Federal Ministry of Transport and Digital Infrastructure.

3. Methodology

Most algorithms dealing with image segmentation and pattern recognition for HD Map creation consist of two steps. On the first stage, a raw division between the road network and its surroundings is done. By doing so, the search space for the second stage is in most cases shrinked by several order of magnitudes. The second stage deals with the determination of lanes, road markings and other descriptive parameters of the transportation systems.

3.1. Raw Image Segmentation

Depending on the environment, the road determination can be undertaken using simple or sophisticated algorithms, with the latter integrating assumptions about the color and shape of roads. Recent contributions show promising results concerning the extraction of road topology from aerial imagery [24]. Given the impact and completeness of volunteered geographic open data projects such as OpenStreetMap (OSM), road network masks can also be generated by combining the georeferenced images with such data bases. The creation of buffer nets based on the OSM road

network where the buffer size depends on the road classes, the rasterization of the result and their combination with the aerial images leads in most cases to an accurate search space fitting.

Figure 1 depicts such a buffer net, on which we apply an algorithm for the classification of lane markings, where the orientation angle θ of the road is an obligatory input parameter. As roads are represented as polylines within the OSM data model, the approximated angle can be retrieved using Equation(1), where x_1, y_1 and x_2, y_2 are the points between which the angle is computed.

$$\theta = \tan^{-1}\frac{y_2 - y_1}{x_2 - x_1} \tag{1}$$

Figure 1. OpenStreetMap mask showing color-encoded road direction angle.

3.2. Image Classification—Lane Marking Determination

The determination of the lane model is achieved by classifying the lane markings. In our experiments, we used a Random Forest classifier which is well-known for being robust to overfitting due to its stochastic nature. All of the employed features highlight the unique radiometric and spatial appearance of lane markings. Five features are used, namely the three components of the RGB image converted to the HSV (Hue, Saturation, and Value) space, a special intensity-based image filter and a Gabor filter.

The HSV representation is preferred, as unicolor pixels (low Saturation) with high intensity (high Value) can be identified faster than in raw RGB color space, where the dependency between the three channel values has to be taken into account. Figure 2 depicts the HSV transformation where Fig. 2 (a) is the RGB input image and Fig. 2 (b, c, d) are the HSV components.
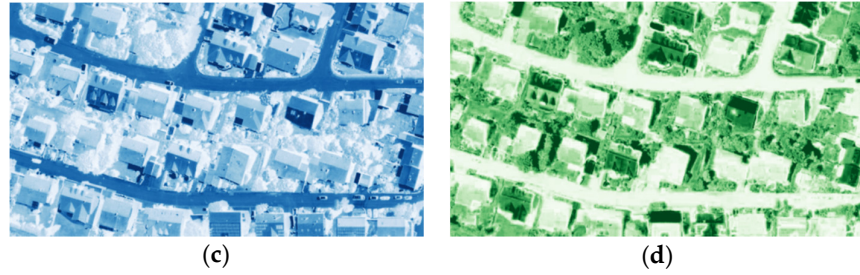

(**a**)　　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 2. RGB to HSV components: (**a**) RGB, (**b**) Hue, (**c**) Saturation and (**d**) Value.

The specific offset in color appearance between lane marking and road is emphasized by introducing a fourth feature, which is described in Equation(2). This is computed as the brightness at position z_0 minus the mean brightness $\bar{z}$ in a circular surrounding (inner radius 0.5 m, outer radius 2 m), divided by the standard deviation within the circular surrounding. Figure 3 shows the behavior of this filter.

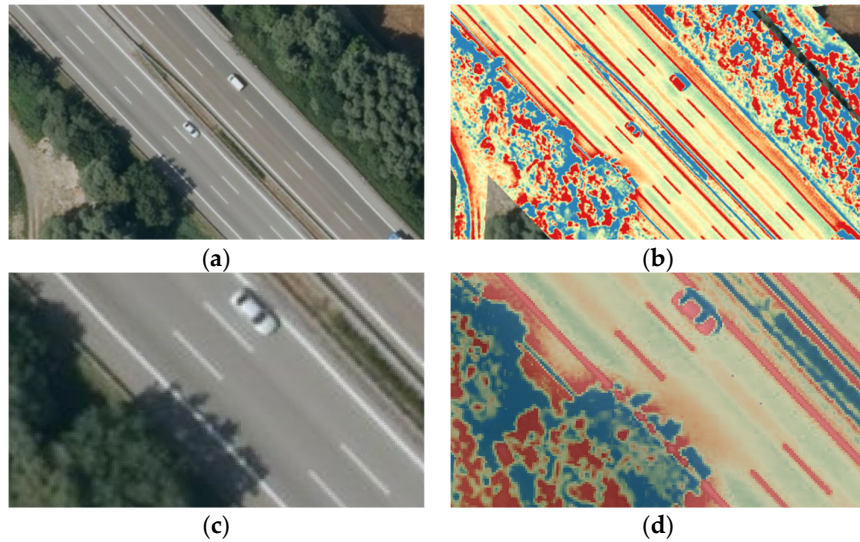$$f(x,y) \; = \; \frac{z_0 - \bar{z}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(z_i - \bar{z})^2}} \tag{2}$$

Figure 3. (**a**) typical situation on motorway, (**b**) result of intensity-based filtering of (**a**), (**c**) subset of (**a**) with shadow, (**d**) result of intensity-based filtering of (**c**).

The spatial characteristics of the road markings are given by a specific length, width and orientation. A directed texture descriptor such as the Gabor filter, where certain parameters such as the wavelength λ, the direction θ and the bandwidth must be set, can be used to highlight lane markings. The angle θ is derived from the pre-introduced OSM mask, while the wavelength λ relates to the specific width of lane markings in our images (here approximately four pixels). The result of the Gabor filtered image consists of a real and imaginary part of a complex number. In our approach, we neglect the imaginary part and use the real part, where the general formulation is given in Equations(3)–(5). Figure 4 visualizes the Gabor filtering.

$$f(x,y,\lambda,\theta,\psi,\sigma,\gamma) \; = \; \exp(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}) \cos(2\pi\frac{x'}{\lambda} + \psi) \tag{3}$$

where

$$x' \; = \; x\cos(\theta) + y\sin(\theta) \tag{4}$$

$$y' = -x\sin(\theta) + y\cos(\theta) \tag{5}$$

(a)　　　　　　　　　　　　　　　　　　　　(b)
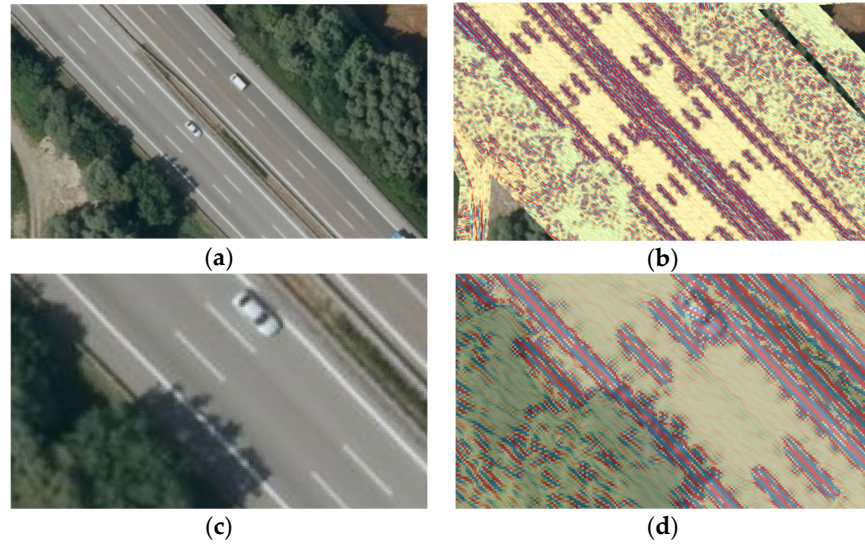
(c)　　　　　　　　　　　　　　　　　　　　(d)

Figure 4. (**a**) RGB image showing motorway, (**b**) Result of Gabor Filtering of pre-masked aerial imagery, (**c**) subset of (**a**), (**d**) Result of Gabor Filtering of (**c**).

Based on this features, a random forest classifier was trained and applied. During the training phase, experiments were carried out for hyper parameter tuning. The classifier consists of 300 trees having a maximum tree depth of 13 and minimum 700 samples per leaf. The used image patch size per sample is 7 × 7 pixels. As the ratio True Positives to True Negatives is quite unbalanced (most of the imagery is background), we used a synthetic ratio of 1:5 during the training for ensuring the classifier to get a good understanding of how a lane marking should look like.

4. Results and Discussion

4.1. Results

Three images were used for evaluation of the trained classifier. As figures of merit we report Accuracy (Equation(5)), Sensitivity (Equation(6)) and Intersection over Union (IoU, Equation(7)). All of the three are ratios considering two or more of the following values: TruePositives, True Negatives, False Positives, and False Negatives. Due to the unbalanced nature of the problem (almost 100% true negatives), we consider accuracy as an inappropriate measure, but include it for completeness. Table 1 gives the metrics on each of the three test images and the ratio for the three test images.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$Sensitivity = \frac{TP}{P} \tag{6}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{7}$$

Table 1. Classification results for three test images.

Scene	Accuracy	Sensitivity	IoU
1	0.99	0.54	0.5
2	0.99	0.65	0.6
3	0.99	0.62	0.59

Mean	0.99	0.6	0.56

4.2. Discussion

Several aspects can lead to false predictions. One of the more dominant error sources is represented by man-made features sharing similar colors and shapes with lane markings. Figure 5 gives an example about such a situation. Figure 5a depicts the used Ground Truth in cyan, whereas in 5b the prediction in magenta is shown. Figure 5c gives the RGB input image, while Figure 5d represents the Ground Truth (cyan) and prediction (magenta) overlaid on the grayscale version of the input image. Between the two inner guardrails a long solid line is predicted, which is obviously wrong. The two roads are divided by concrete barriers which look like bright lines (almost white) in the imagery (ref. Figure 5c); however, these are no lane markings, and therefore this mis-prediction lowers the quality of the overall results.
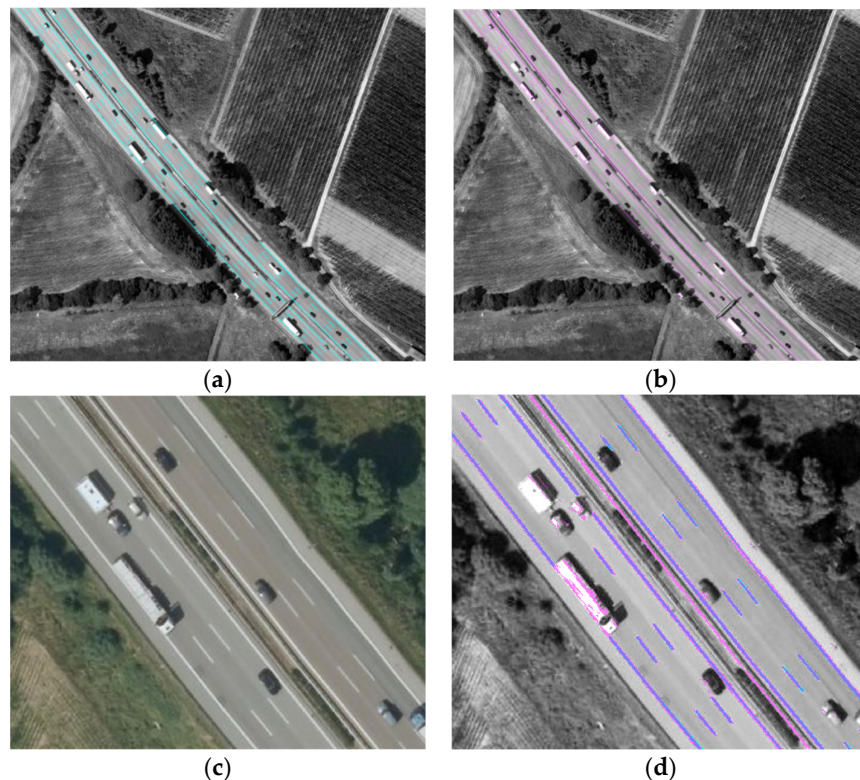

Figure 5. (**a**) Ground Truth, (**b**) Prediction, (**c**) Image, (**d**) Overlay of Ground Truth and Prediction.

Figure 6 gives a closer look at another error source. Due to the harsh environment situation on motorways, the colors and shapes of lane markings may differ. Furthermore, the concrete itself can exhibit several disturbing features, such as small cracks, potholes or simply unbalanced stress on the different lanes, which make spatially limited reparations necessary. Figure 6 illustrates such a situation, where 6a is the RGB input image, 6b is showing the Ground Truth in cyan and 6c is showing an overlay of Ground Truth (cyan) and prediction (magenta) on a grayscale version of the image. We observe that the classifier seems to underpredict the lane markings in situations where the marking color is washed out. The two directions of the road can easily be distinguished by defining them as the bright one (lower one) and the dark one (upper one). In the second case, we observe that the road surface already shows brighter areas bordering the single lanes, even if there is no road marking. This is caused by heavy usage of the road, where the vehicles weight tends to deform the road surface, while areas which are probably characterized by less traffic appear brighter.

Figure 6. (**a**) Image, (**b**) Ground Truth, (**c**) Overlay Ground Truth and Prediction.

In previous works, shadows and changing illumination conditions have been identified as a major error source. The employed algorithm is quite robust against such influences, due to the usage of Gabor features. Figure 3 gives a good example about the strong influence of shadows on intensity-based features. Figure 3c, d highlight how shadows destroy the relevant information as the foreground/background intensity ratio decreases. Figure 4c, d shows the reliability of Gabor features, which have a beneficial influence on the overall classification results.

In our experiments, occlusions by cars had no negative influence on the overall results as all occluded lane markings were not part of the Ground Truth. However, having the goal of generating country-scale HD maps in mind, a preprocessing of the imagery would be necessary. As the images overlap, being the vehicles moving objects, a median filtering would probably be enough to remove most cars from the input orthophoto mosaic.

Our experiments focused on motorways. The presented algorithm would deliver comparable results in urban situations, as shape and color of lane markings would just change marginally. In urban environments the aforementioned error sources, namely shadows and occlusions, will have a much higher impact as a higher number of man-made objects close to the roads and a significantly higher traffic participant intensity are expected.

5. Conclusions

In this paper we analyzed the possibility of creating HD Maps for autonomous driving using aerial imagery. Using a current test site in Germany, we derived a HD Map using a novel image segmentation algorithm, which combines OSM data and aerial imagery from the Bavarian Agency for Digitisation, High-Speed Internet and Surveying. The quality metrics prove that the proposed algorithm is able to produce reliably and efficiently HD Maps for large areas.

The used imagery is available on a country scale (for all federal states of Germany) and it is frequently updated, therefore the proposed algorithm could be used to produce a base map for all motorways in Germany. Several other European countries (e.g., Austria, France) run comparable programs with slightly differing spatial and temporal resolutions: however, by combining existing image databases, it is possible to create a VHR orthophoto mosaic at a continental scale. The used OSM data for masking and derivation of the road orientation angles could be replaced by more reliable data sources, possibly provided by official national mapping agencies.

The usage of such a base map with its given high geometric stability may be useful for several purposes, for instance: (a) to create an expectation horizon for classifying the objects within the neighborhood of the car, (b) to validate the cars observations and the created live map, and (c) to raise the absolute accuracy of the position estimation. In the future an in-depth analysis has to be done on the possibility of merging aerial with probe-based imagery.

Nowadays car-makers ask for lightweight mapping formats such as OpenDRIVE or Navigation Data Standard (NDS). Due to their shape, size and orientation, the raw binary lane markings can be translated towards an object-specific description. The goal is to achieve a vectorized description of the transportation network on a lane level, which is georeferenced and includes at least a basic topology. The binary lane marking classification could serve as a starting point for the creation of more sophisticated navigation maps.

The validation of existing maps by means of geometry is already an use-case [37] and should be studied in depth in the future. The development of a reliable validation procedure of maps based on aerial imagery may be of interest for a broad range of stakeholders within the domain of AD. The German Aerospace Center recently released a new dataset in order to stimulate new contributions and out-of-the-box solutions for the automotive industry [38].

References

1. Anderson, J.M.;Kalra, N.; Stanley, K.D.; Sorensen, P.; Samaras, C.;Oluwatola, O.A. *Autonomous Vehicle Technology: A Guide for Policymakers*; RAND Corporation:Santa Monica, CA, USA, 2014.
2. Bagloe, S.A.;Tavana, M.;Asadi, M.; Oliver, T. Autonomous vehicles: Challenges, opportunities, and future implications for transportation policies. *J. Mod. Transp.***2016**, *24*, 284–303.
3. Geiger, A. Are we ready for autonomous driving? The KITTI vision benchmark suite. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), CVPR '12, Providence, RI, USA, 16–21 June2012; IEEE Computer Society: Washington, DC, USA, 2012; pp. 3354–3361.
4. Seif, H.G.; Hu, X. Autonomous driving in the ICity—HD maps as a key challenge of the automotive industry. *Engineering***2016**, *2*, 159–162.

5. Obst, M.; Bauer, S.; Reisdorf, P.;Wanielik, G. Multipath detection with 3D digital maps for robust multi-constellation GNSS/INS vehicle localization in urban areas. In Proceedings of the2012IEEE Intelligent Vehicles Symposium, Alcala de Henares, Spain, 3–7 June2012;pp. 184–190.

6. Tao, Z.; Bonnifait, P. Tightly coupling GPS with lanemarkings for autonomous vehicle navigation. In Proceedings of the17th International IEEE Conference on Intelligent TransportationSystems (ITSC),Qingdao, China, 8–11 October 2014;pp. 439–444.

7. Tao, Z.; Bonnifait. P. Road invariant extended Kalman filter for an enhanced estimation of GPS errors using lane markings. In Proceedings of the2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),Hamburg, Germany, 28 September–2 October 2015;pp. 3119–3124.

8. Hosseinyalamdary, S.; Peter, M. Lane level localization; using images and HD maps to mitigate the lateral error. *ISPRSInt. Arch.Photogramm. Remote Sens. Spat. Inf. Sci.***2017**, *XLII-1/W1*, 129–134.

9. Bauer, S.;Alkhorshid, Y.; Wanielik, G. Using high definition maps for precise urban vehicle localization. In Proceedings of the2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Rio de Janeiro, Brazil, 1–4 November2016;pp. 492–497.

10. Schreiber, M.;Knöppel, C.; Franke, U. Laneloc: Lanemarking based localization using highly accurate maps. In Proceedings of the2013 IEEE Intelligent Vehicles Symposium (IV),Gold Coast, Australia, 23–26 June 2013;pp. 449–454.

11. Burgard, W.; Brock, O.; Stachniss. C. *Map-Based Precision Vehicle Localization in Urban Environments*; MIT Press:Cambridge, MA, USA, 2008.

12. Wang, S.;Urtason, S.;Filder, S. Holistic 3d scene understanding from a single monocular image. In Proceedings of the CVPR 201528th IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.

13. Kammel, S.; Ziegler, J.; Pitzer, B.;Werling, M.;Gindele, T.;Jagszent, D.;Schröder, J.; Thuy, M.;Goebl, M.; von Hundelshausen, F.;et al. Team annieway's autonomous system for the DARPA urban challenge 2007. In *The DARPA Urban Challenge,Springer Transactions in Advanced Robotics*; Buehler, M., Iagnemma, K., Singh, S., Eds.; Springer Berlin Heidelberg: 2009;Volume 56, pp. 359–391.

14. Ziegler, J.; Bender, P.; Schreiber, M.;Lategahn, H.; Strauss, T.; Stiller, C.; Dang, T.; Franke, U.;Appenrodt, N.;Keller, C.G.;et al. Making bertha drive—An autonomous journey on a historic route. *IEEE Intell. Transp. Syst. Mag.***2014**, *6*, 8–20.

15. Aeberhard, M.; Rauch, S.; Bahram, M.;Tanzmeister, G.; Thomas, J.;Pilat, Y.;Homm, F.; Huber, W.; Kaempchen, N. Experience, results and lessons learned from automated driving on Germany's highways. *IEEE Intell. Transp. Syst.Mag.***2015**, *7*, 42–57.

16. Macfarlane, J.; Stroila, M. Addressing the uncertainties in autonomous driving. *SIGSPATIAL Spec.***2016**, *8*, 35–40.

17. Massow, K.;Kwella, B.; Pfeifer, N.;Husler, F.;Pontow, J.;Radusch, I.;Hipp, J.;Dlitzscher, F.; Haueis, M. Deriving HD maps for highly automated driving from vehicular probedata. In Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Rio de Janeiro, Brazil, 1–4 November2016;pp. 1745–1752.

18. Dabeer, O.;Gowaiker, R.;Grzechnik, S.K.;Lakshman, M.J.;Reitmayr, G.; Somasundaram, K.; Sukhavasi, R.T.; Wu, X. An end-to-end system for crowd sourced 3d maps for autonomous vehicles: The mapping component. *arXiv***2017**,arXiv:1703.10193.

19. Kim, J.G.;Han, D.Y.;Yu, K.Y.;Kim, Y.I.;Rhee, S.M. Efficient extraction of road information for car navigation applications using road pavement markings obtained from aerial images. *Can. J. Civil Eng.***2006**, *33*, 1320–1331.

20. Jin, H.;Miska, M.; Chung, E.; Li, M.; Feng, Y. Road feature extraction from high resolution aerial images upon rural regions based on multi-resolution image analysis and Gabor filters. In *Remote Sensing-Advanced Techniques and Platforms*; IntechOpen Rijeka; 2012.

21. Jin, H.; Feng, Y. Automated road pavement marking detection from high resolution aerial images based on multi-resolution image analysis and anisotropic Gaussian filtering. In Proceedings of the 2010 2nd International Conference onSignal Processing Systems (ICSPS),Dalian, China, 5–7 July 2010; Volume 1, pp. 337–341.

22. Jin, H.; Feng,Y.; Li, M. Towards an automatic system for road lane marking extraction in large-scale aerial images acquired over rural areas by hierarchical image analysis and Gabor filter. *Int. J. Remote Sens.***2012**, *33*, 2747–2769.

23. Hinz, S.; Baumgartner, A. Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS J.Photogramm. Remote Sens.***2003**, *58*, 83–98.

24. Mnih, V.; Hinton, G.E.*Learning to Detect Roads in High Resolution Aerial Images*; Springer: Berlin/Heidelberg, Germany, 2010;pp. 210–223.

25. Mattyus, G.; Wang, S.; Fidler, S.; Urtasun, R. HD maps: Fine-grained road segmentation by parsing ground and aerial images. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June2016;pp. 3611–3619.

26. Gellert, M.; Luo, W.;Urtasun, R. DeepRoadMapper: Extracting Road Topology from Aerial Images. In Proceedings of the International Conference on Computer Vision (CVPR), Honolulu, Hawaii, 21–26 July 2017.

27. Mayer, H.; Hinz, S.; Bacher, U.; Baltsavias, E. A test of automatic road extraction approaches. *Int. Arch.Photogramm. Remote Sens. Spat. Inf. Sci.***2006**, *36*,209–214.

28. Wang, W.; Yang, N.; Zhang, Y.; Wang, F.; Cao, T.; Eklund, P. A review of road extraction from remote sensing images. *J. Traffic Transp. Eng. (Engl. Ed.)***2016**, *3*, 271–282.

29. Greenwalt, C.; Shultz, M. *Principles of Error Theory and Cartographic Applications*; ACIC Technical Report; Aeronautical Chart and Information Center: Springfield, VA, USA, 1965.

30. Javanmardi, M.;Javanmardi, E.; Gu, Y.; Kamijo, S. Towards high-definition 3D urban mapping: Road feature-based registration of mobile mapping systems and aerial imagery.*Remote Sens.***2017**, *9*, 975.

31. Huang, J.; Liang, H.; Wang, Z.; Song, Y.; Deng, Y. December. Lane marking detection based on adaptive threshold segmentation and road classification. In Proceedings of the 2014 IEEE International Conference onRobotics and Biomimetics (ROBIO),Bali, Indonesia, 5–10 December2014;pp. 291–296.

32. Tournaire, O.; Paparoditis, N.; Lafarge, F. Rectangular road marking detection with marked point processes. In Proceedings of theConference on Photogrammetric Image Analysis, Munich, Germany, 19–21 September 2007; Volume 3.

33. Azimi, S.M.; Fischer, P.; Körner, M.; Reinartz, P. Aerial LaneNet: Lane Marking Semantic Segmentation in Aerial Imagery using Wavelet-Enhanced Cost-sensitive Symmetric Fully Convolutional Neural Networks.*arXiv***2018**, arXiv:1803.06904.

34. Lee, S.; Kim, J.; Yoon, J.S.; Shin, S.; Bailo, O.; Kim, N.; Lee, T.H.; Hong, H.S.; Han, S.H.; Kweon, I.S. October. Vpgnet: Vanishing point guided network for lane and road marking detection and recognition. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV*)*, Venice, Italy, 22–29 October 2017; pp. 1965–1973.

35. Gurghian, A.; Koduri, T.; Bailur, S.V.; Carey, K.J.; Murali, V.N.Deeplanes: End-to-end lane position estimation using deep neural networksa. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 38–45.

36. Baltrusch, S. TrueDOP—A new quality step for official orthophotos. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *XLI-B4*,619–624.

37. Fischer, P.; Plaß, B.; Kurz, F.; Krauß, T.; Runge, H. Validation of HD Maps for autonomous driving. In Proceedings of the International Conference on Intelligent Transportation Systems in Theory and Practice mobil.TUM, Munich, Germany, 4–6 July 2017.

38. Kurz, F.; Waigand, D.; Pekezou-Fouopi, P.; Vig, E.; Corentin, H.; Merkle, N.; Rosenbaum, D.; Gstaiger, V.; Azimi, S.M.; Auer, S.; et al. DLRAD—A first look on the new vision and mapping benchmark dataset. In Proceedings of the ISPRS TC1 Symposium—Accepted Contribution, Hannover, Germany, 10–12 October 2018.