Technical Disclosure Commons

Defensive Publications Series

April 2020

GENERATION OF HIGH-PRECISION MAPS BY MEANS OF VEHICLE SENSOR DATA AND GROUND TRUTH MAPS

Verena Blunder

Bertrandt Ingenieurbüro GmbH

GENERATION OF HIGH-PRECISION MAPS BY MEANS OF VEHICLE SENSOR DATA AND GROUND TRUTH MAPS

Technical task:
By matching vehicle sensor data and ground truth maps, highly accurate maps can be generated using a machine learning model.

Initial situation:
Today, common methods such as SLAM, filtering and optimization techniques are used to derive high-precision maps from vehicle environmental observations. Environmental observations of vehicles can be partially obscured (by structural objects or vehicles driving in front of them) or not detected (sensor technology is not designed for such detections). This results in gaps or errors in the generated map model, which leads to inferior map quality.
To solve this problem patent applications exist, which process electronic maps by neural networks, create high-precision ground truth maps by machine learning and image processing of satellite cameras or create 3D map segments by machine learning using sensor data from the vehicle. The following invention, however, uses the sensor data from the vehicle and the highly accurate ground truth maps to predict ground truth maps from the sensor data by machine learning of the deviations.

Solution:
Today's surveying vehicles can produce highly accurate map models (hereinafter also referred to as GT or ground truth maps), which define a high degree of detail and absolute accuracy. In addition, details and objects can be surveyed which are not yet detected by current or future serial sensor technology of vehicles. However, such high accuracy maps are expensive, require a high degree of manual post-processing and are only available in the regions where a contract has been awarded to the respective surveying company. By driving in observed environment information of a vehicle fleet, this information can be related to already high-precision surveyed map information or ground truth maps. Consequently, a machine learning procedure can be applied to predetermine the output data (ground truth maps) with the input data (sensor information). Thus, the statistical dispersion of the sensor information can be used to draw direct conclusions about the true value.

Advantages:
The machine learning model (ML model) can be taught skills, e.g. to ensure the automatic closing of gaps caused by incomplete observation processes. Furthermore, the addition of undetected observations is possible (e.g. automatic placement of arrows of road markings). An overarching advantage is that training the ML model provides an understanding of country-specific road construction behaviour. As a result, it is possible to determine automatically how, for example, the road width is defined in certain sections, double crossed out lines are located, etc. In addition, the necessary number of passes by traffic areas can be reduced compared to common mapping methods (SLAM, filter or optimization methods). Furthermore, complex road situations (e.g. intersections or large roundabouts) can be created more accurately, since the ML model has an understanding through training by means of high-precision surveyed road areas.

Possible application:
The starting position is described by the two data sources sensor information from the vehicle and its environmental observations as well as ground truth map material (measured with high precision). Figure 1 shows sensor information with an overlaid grid and Figure 2 shows a ground truth map with overlaid grid.

The grid is used for the assignment of the observations and the measured reality (ground truth). Possible approaches for the setting of a grid would be e.g. the Mercator projection, quadkeys etc. Besides quadrangles, hexagons etc. can also be used. A sector of the grid is henceforth called a tile. Around such a tile again an n-neighbourhood can be described. In the previously shown figures 1 and 2 an 8-neighbourhood is drawn.

The grid thus limits the amount of information of the input. However, the size of the tiles can be chosen arbitrarily. The amount of information includes the geometric descriptions of environmental observations, but also, for example, the type of geometries, time stamps of the observations and much more.

Consequently, by defining a grid, an assignment between input and output can be determined, as shown in Figure 3.

The input is thus routed to the input layer of an ML model (e.g. approaches such as neural networks, CNN, DCNN, GAN, etc.). In the training phase, the model is trained to ensure that the output generated thereby corresponds as closely as possible to the assigned ground truth section. Figure 4 illustrates the entire processing chain. Sensor information and ground truth maps are converted into uniform n-dimensional feature spaces. In the prediction phase an output is predicted from further sensor information, which corresponds optimally to a ground truth card. The clamping of a feature space is realized programmatically in an n-dimensional matrix. This requires at least the two dimensions that describe

the geometries of the map. As further dimensions the height of the geometries, but also properties such as type of geometry, temporal aspects of the observation etc. can be included. Figure 5 shows how to span a feature space.
The representation of the geometries in feature space is critical. One possible variant would be to convert the geometries into an imaging process (see Figure 3 with possible image conversion and definition of the pixel interpretation against a possible unit length of millimetres), where in turn the individual pixels are held in a 2-dimensional matrix. This matrix can be extended by further feature dimensions. Compared to this variant, an alternative variant can describe the geometries using mathematical systems of equations, which are also programmatically recorded in a matrix, which in turn is expanded with the necessary dimensions.
To prevent possible drifting of the geometry transitions at the tile boundaries, the described neighborhood approach can be transferred to feature space. This ensures consistent transitions.

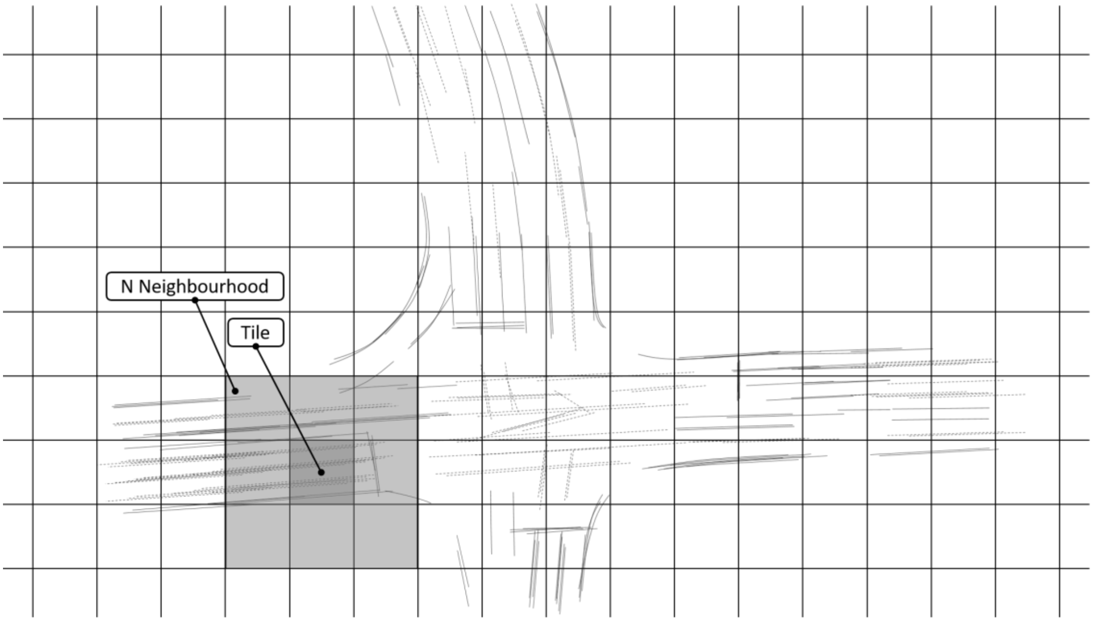

Figure 1: Possible sensor information (e.g. road markings) drawn in a coordinate system and a superimposed grid. There are gaps in the observations (compared to the ground truth map in Figure 2). Additionally, a tile is shown with an n=8 neighbourhood.

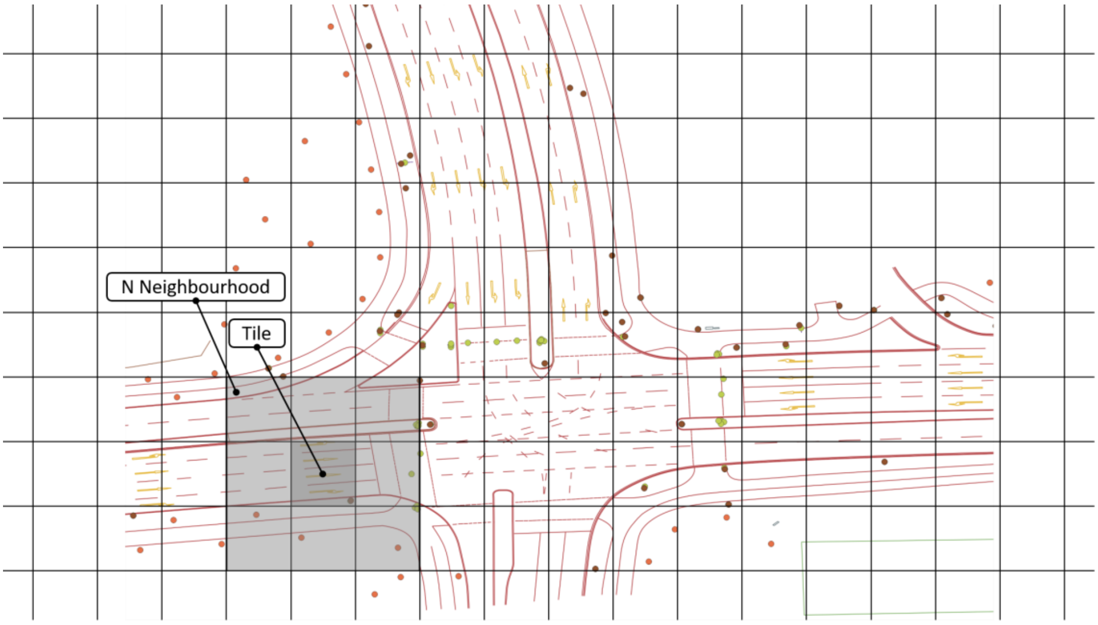

Figure 2: Associated ground truth map and a superimposed grid additionally, a tile is shown with an n=8 neighborhood. Furthermore, road arrows, trees, posts, traffic signs and much more can be found in the ground truth map.

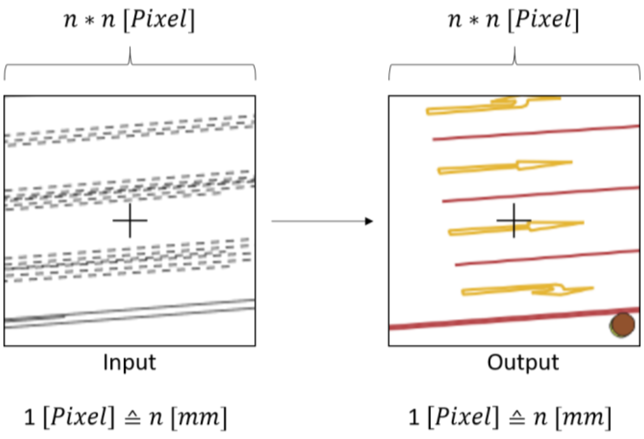

Figure 3: Sections of line observations observed by vehicle sensors on the left and the corresponding ground truth map section on the right. The sections define a tile assignment (e.g. by quadkeys). The cross defines the center of the tile. The center of the tile defines the center of the tile. The center of the tile is used to determine the drawn geometries absolutely in the world coordinate system by a relative coordinate transformation.

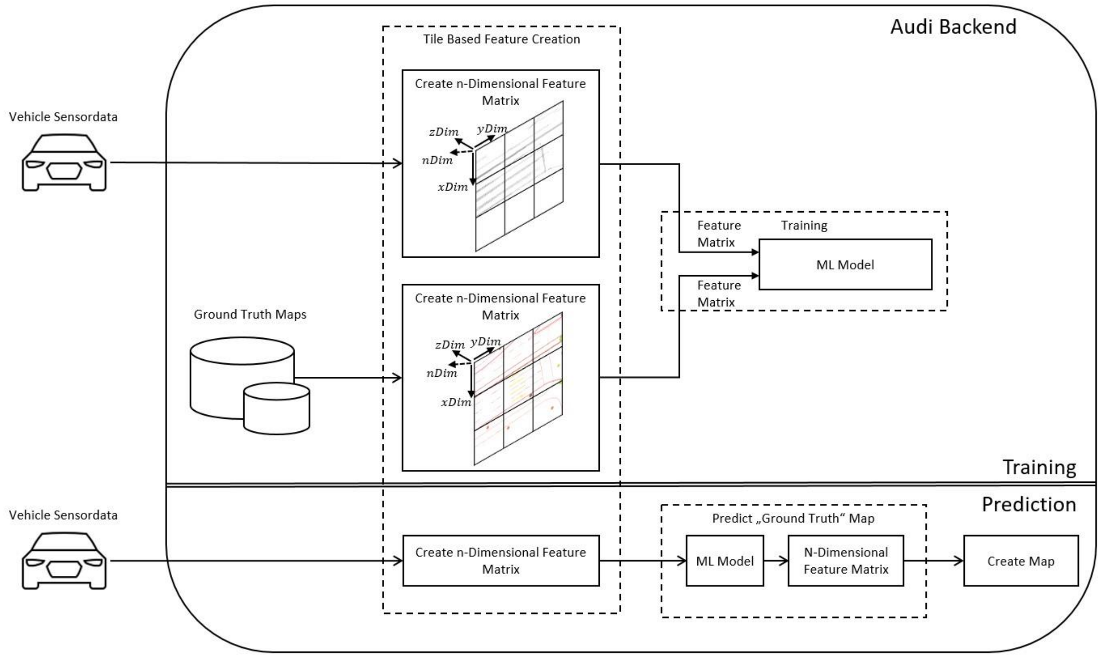

Figure 4: Process representation of the information flow regarding sensor information and ground truth maps. The sources are converted into uniform n-dimensional feature spaces. The spaces are bounded by the previously described grid. In the training phase, the ML model is trained so that the generated output is as close as possible to the ground truth map. In the prediction phase, output is predicted from further sensor information, which corresponds optimally to a ground truth map.

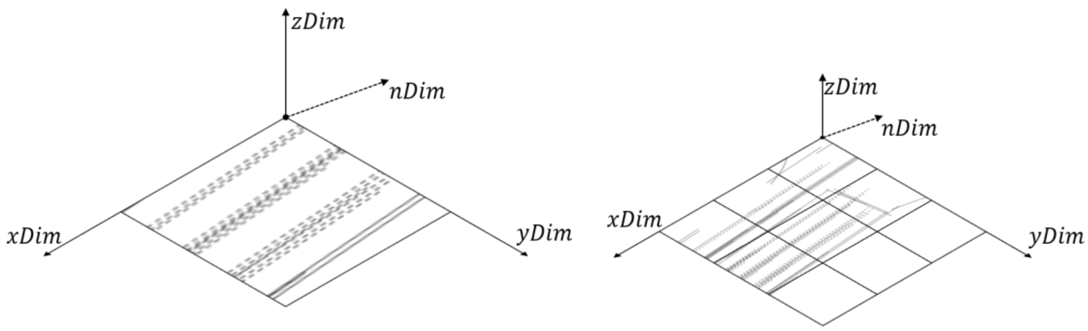

Figure 5: Spanning a feature space (left: tile-based; right: nneighbourhood). Programmatically, this feature space is implemented in an n-dimensional matrix. This requires at least the two dimensions that describe the geometry of the

map (here xDim and yDim). As a further dimension the height of the geometries can be transferred (here zDim), which defines the absolute height above normal zero in an absolute world coordinate system. In addition, further dimensions can be inserted (here nDim), which describe e.g. properties of the type of the geometries, but also temporal aspects could be included in the description space.