QUALITY ASSESSMENT OF LANDMARK BASED POSITIONING USING STEREO CAMERAS

S. Hofmann, M. J. Schulze, M. Sester, C. Brenner

Institute of Cartography and Geoinformatics, Leibniz Universität Hannover, Appelstraße 9a, 30167 Hannover, Germany – (sabine.hofmann, maltejan.schulze, monika.sester, claus.brenner)@ikg.uni-hannover.de

Working Groups I/2, III/1, III/4, III/5

ABSTRACT:

Driving autonomously requires highly accurate positioning. Therefore, alternative positioning systems to GPS are required especially to increase the accuracy, and to have a complementary data source in areas where GPS is not available. As more and more on-board sensors are used for safety reasons, information gathered about their environment can be used for positioning based on relative measurements to landmarks along the road. This paper investigates the accuracy potential of positioning using a stereo camera system and landmark maps. Therefore, we simulated several stereo camera systems with variable opening angle and base length to compute the positioning accuracy in a test area. In the first step, localization was calculated based on single positions, in the second step we used a Kalman filter additionally. While positioning in the first case was not successful along the entire trajectory, the Kalman filter led to far better results.

1. INTRODUCTION

Driving autonomously or driver assistance applications make highly accurate positioning much more important than in today's navigational devices. Therefore, alternative positioning systems to GPS are required to increase the accuracy, and to have a complementary data source in areas where GPS fails, for example in street canyons.

On-board sensors, such as cameras, laser scanner and radar, which gather information about their environment for active safety systems, for example environment detection to avoid collisions with pedestrians or recognition of traffic signs, can be used for positioning based on relative measurements to objects along the road. To use this data for positioning a highly accurate representation of the environment is required. Iconic representations, for example occupancy grids or symbolic representations, for example landmark based maps (Burgard and Hebert, 2008) can be used. Using landmarks such as poles of traffic signs and traffic lights for positioning has been investigated earlier, e.g. by Weiss et al., 2005, Brenner, 2010.

To investigate, how accurate positioning is possible using a stereo camera system and landmark maps, we simulated several stereo camera systems with variable opening angle and base length. In addition to calculating the accuracy of single positions, we simulated the knowledge of two different types of inertial measuring units (IMU), a precise and an automotive grade device.

2. DATA

Data basis for the simulation is a 21.7 km long trajectory which runs through densely built-up regions as well as along highway-like roads in the area of Hannover, Germany. The data was acquired by the Streetmapper mobile mapping system (Kremer and Hunter, 2007) to obtain a dense laser scan. From this point cloud we extracted 2658 pole-like objects, for example sign posts, street lights or tree trunks, fully automatically (Brenner, 2009). These extracted objects build the two-dimensional landmark map which is used in the simulation. The accuracy of position for every object is in the order of 12 cm (Brenner and Hofmann, 2010).

Figure 1 shows the trajectory (red line) together with the extracted objects (green dots). The objects are not distributed equally. Along highway-like roads (left side) there are only very few poles, whereas at inner-city junctions there are normally many usable objects, for example sign posts and traffic lights.
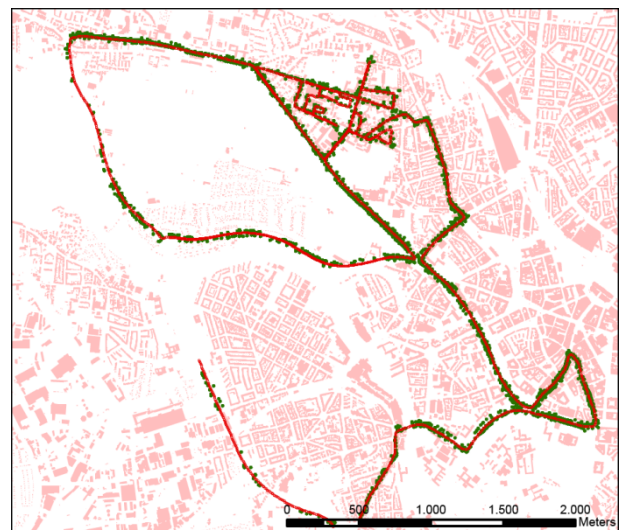

Figure 1. Trajectory (red) and extracted pole-like objects (green) which are used as reference map.

For the simulation of the positioning accuracy we used 2141 positions at a regularly spaced interval of 10 m with known coordinates and heading along the Streetmapper trajectory.

3. SIMULATION

For each of the 2141 positions along the trajectory the possible accuracy was calculated for varying opening angles (from 50° to 100° in steps of 10°) and base lengths (from 0.25 m to 2.25m in steps of 25 cm) of a simulated camera system.

The simulation is separated in three parts. The first part contains the retrieval of all visible objects within the database for each position mainly depending on the distance of the objects to the sensor to preserve a certain minimum width and the opening angle of the cameras.
The second part is the analysis of accuracies for the vehicle's position and heading, which is based on a least squares adjustment.
The third part is the analysis of accuracies using the results from the first and second part of the simulation and an additional Kalman filter.

3.1 Retrieval of Visible Objects

In the first part of the simulation all visible poles have to be determined for each position. An object is visible if its size in the image is not below a minimum size n_{px} and it lies within the overlapping field of view given by the opening angles of both cameras.
The size of the objects in the image depends on the distance between the sensor and the measured object, the focal length c, the size of the pixels on the sensor d_{px} and size of the object itself d_{obj}. A maximum distance s_{max} can be calculated as follows:

$$s_{max} = \frac{c \cdot d_{obj}}{n_{px} \cdot d_{px}} \tag{1}$$

with $c = 0.006$ m
         $d_{obj} = 0.3$ m
         $d_{px} = 5.5 \cdot 10^{-6}$ m.

With a given minimum size of 7 pixels only objects within a maximum distance of 46.75 m can be detected.

In addition, objects have to lie within the overlapping field of view of both cameras (Figure 2), which depends on the opening angle γ, heading κ and the camera positions $K(X_{0j}, Y_{0j})$, $j = 1,2$ number of camera, which are given as

$$\begin{bmatrix} X_{0j} \\ Y_{0j} \end{bmatrix} = \begin{bmatrix} X_0 \pm \frac{b}{2} \cdot \cos\kappa \\ Y_0 \pm \frac{b}{2} \cdot \sin\kappa \end{bmatrix} \tag{2}$$

where X_0, Y_0 = coordinates of centre of base line
         b = base length.

For both projection centers the direction t_p for each object have to be calculated. Therefore, visible objects, which are then used

for the further examination, have to match the following condition:

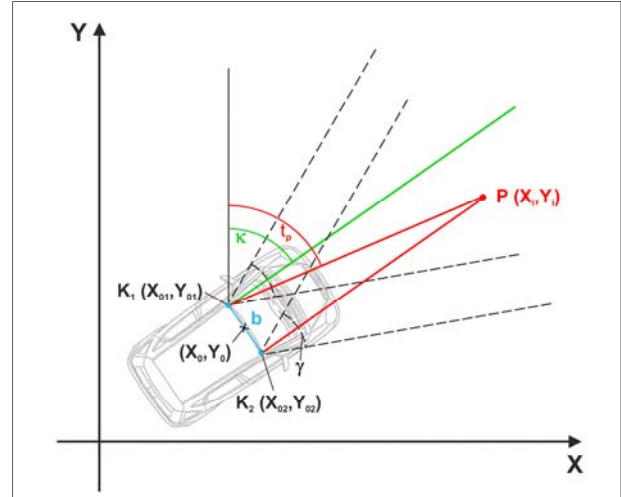$$|\kappa - t_p| \le \frac{\gamma}{2}. \tag{3}$$

Figure 2. Visibility requirements.

3.2 Analysis of Accuracies

To determine the accuracies of the positioning (position and heading) based on stereo cameras, we used a least squares adjustment. At least two visible objects are required to calculate a position.
The functional model of the adjustment which describes the relationship between the measurements (image und corresponding object coordinates) and the unknown parameters (position and heading) are the collinearity equations. As we used a two dimensional map, the collinearity equations can be simplified and reduced to

$$x_i' = c \frac{r_{11}(X_i - X_0) + r_{21}(Y_i - Y_0)}{r_{12}(X_i - X_0) + r_{22}(Y_i - Y_0)} \tag{4}$$

where x_i' = image coordinate
         $P(X_i, Y_i)$ = object coordinates for i-th object in the reference map
         r_{ij} = elements of rotational matrix R.

In addition, we have to take into account, that the projection centers of both cameras do not lie in X_0, Y_0 but are shifted by $b/2$. Therefore, the observation equations for camera $j = 1, 2$ are as followed:

$$x_{i,j}' + v_{x_{i,j}'} = c \frac{\cos\kappa (X_i - X_0) + \sin\kappa (Y_i - Y_0) \pm \frac{b}{2}}{-\sin\kappa (X_i - X_0) + \cos\kappa (Y_i - Y_0)}. \tag{5}$$

The stochastic model contains information about the uncertainties of the observations. In the simulation object coordinates of the reference objects (standard deviation 0.1 m) and image coordinates of the measured objects (standard deviation 1/3 Pixel) are used as measurements. They are considered as fully uncorrelated, base length and focal length are considered to be accurate.

As a result, the cofactor matrix Q_{xx} of the positioning parameters contains the stochastic information of the vehicle's position:

$$Q_{xx} = (A^T \cdot P \cdot A)^{-1} \qquad (6)$$

where A = design matrix
            P = weight matrix.

The design matrix A provides information about the vehicle position, heading and object coordinates (number and distribution, given by reference map), thus the geometry of the setup. The weight matrix P contains the stochastic information, which is the accuracy of the image and the object coordinates.

3.3 Simulation using Filtered Positions

In a real scenario, not only individual positions would be calculated – as described before – but the positions would be filtered along the trajectories. Therefore, we simulated a second scenario, using two different types of inertial measurement units (IMU), a high precision and an automotive grade device. For the filtering, we used a standard Kalman filter based on a simplified car motion model.

The used car model is as follows

$$
\begin{aligned}
x_i &= x_{i-1} + dt \cdot v_{i-1} \cdot \sin \kappa_{i-1} \\
y_i &= y_{i-1} + dt \cdot v_{i-1} \cdot \cos \kappa_{i-1} \\
v_i &= v_{i-1} \\
\kappa_i &= \kappa_{i-1}
\end{aligned} \qquad (7)
$$

where dt = time step between time i and time i-1
            v_i = vehicle speed.

The transition matrix Φ is given by

$$
\Phi = \begin{bmatrix}
1 & 0 & dt \cdot \sin \kappa_{i-1} & 0 \\
0 & 1 & dt \cdot \cos \kappa_{i-1} & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}. \qquad (8)
$$

For the simulation we analysed the cofactor matrix Q_{xx} at time step i of the positioning parameters which is given by

$$Q^+_{xx(i)} = Q^-_{xx(i)} - K \cdot A \cdot Q^-_{xx(i)} \qquad (9)$$

where '–' denotes the prediction and '+' denotes the update step of the filter, with

$$
\begin{aligned}
Q^-_{xx(i)} &= \Phi \cdot Q^+_{xx(i-1)} \cdot \Phi^T + Q_{ww} \\
K &= Q^-_{xx(i)} \cdot A^T \cdot (Q_{ll} + A \cdot Q^-_{xx(i)} \cdot A^T)^{-1}
\end{aligned} \qquad (11)
$$

where K = Kalman gain
            Q_{ll} = variance matrix of the measurements
            Q_{ww} = system noise.

To filter the positions, landmark information were only used when they provided a high positioning accuracy (< 0.2 m). In all other cases only information provided by the IMU was used.

4. RESULTS

4.1 Results using Single Positions

In the first case, the simulation is based on single positions, no filtering is used along the driven path. The results show that the most important factor for the positioning accuracy is the number of objects in the field of view (Figure 3). Along the trajectory there are 2.7 to 4.6 visible objects on average depending on the opening angle (50° to 100°, respectively). With less than two visible objects, positioning is not possible. In our test area, the number of cases where we failed to retrieve the position decreases from 38.9 % to 20.6 % with increasing opening angle. Increasing the base length leads to a higher positioning accuracy, in cases where positioning is possible. For camera systems with a small opening angle, a larger base length leads to fewer visible objects based on a smaller overlap of the field of view. On average, for a typical camera system with opening angle of 100° and base length of 0.25 m, accuracies between 0.41 m (three poles visible) and 0.12 m (more than six poles visible) were achieved.
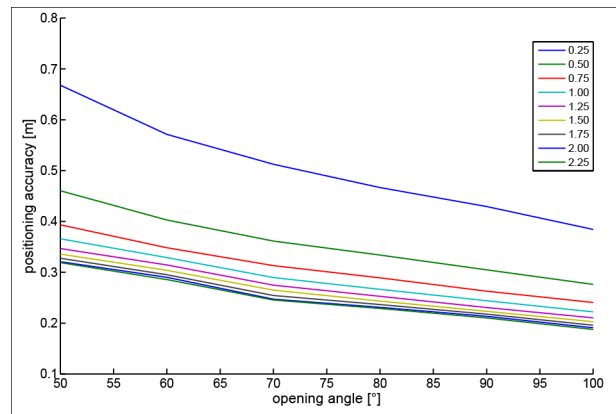

Figure 3. Positioning accuracy as a function of opening angle and base length (between 0.25 m and 2.25 m, see legend in upper right corner).

In the following examples the camera systems may have different opening angles but the same base length of 0.25 m.

In Figure 4 the positioning accuracy for two simulated camera systems are presented. Green ellipses show the positioning accuracy with 1 σ standard deviation for 100° opening angle, orange ellipses for 60° opening angle. Due to larger opening angles the overall positioning accuracy for the first case (green ellipses) is higher. For evenly distributed objects the accuracy

perpendicular to the driving direction is higher due to the model of the stereo camera. With few unfavorably distributed objects (e.g. only on one side of the road) the accuracy perpendicular to the driving direction is lower (maximum, green: 2.09 m, orange: 2.97 m). Very high accuracies (0.1 m) are achieved in both cases e.g. along an avenue lined with trees (Figure 4, upper right side).
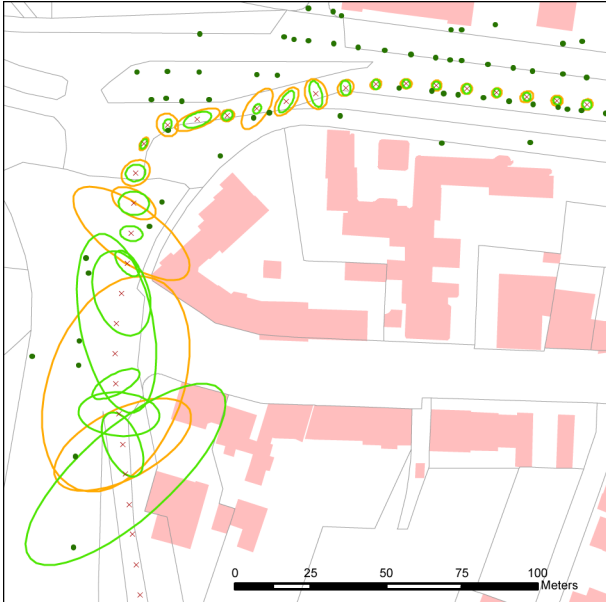

Figure 4. Positioning accuracy for single positions drawn as error ellipses (green: opening angle 100°, orange: opening angle 60°). All ellipses scaled by a factor of 20. (Red crosses: positions on trajectory, green dots: objects in reference map.)

4.2 Results using Filtered Positions

Compared to positioning based on measurements at single positions, the number of positions where positioning failed was decreased significantly when using the Kalman filter with IMU data (Figure 5 and Figure 6). With a camera system with 100° opening angle, positioning based on single positions failed in 20.6 % of the cases. Using a precise IMU (defined by $\sigma_v = 0.02$ m/s (standard deviation velocity), $\sigma_\phi = 0.01°$ (std. dev. heading), $\sigma_{\phi'} = 0.001°$/s (std. dev. angular velocity)) the number of failures decreased to 0.2 %, where failure is defined as positioning accuracy > 20 m. Furthermore, the accuracy of calculated positions increased using a precise IMU.

Figure 5 compares the results using single positions and filtered positions based on a precise IMU. Along the highway (lower left side), very few poles lead to a high number of failed positioning, when not using a filter (green ellipses). Using a precise IMU bridges areas without a sufficient number or poorly distributed landmark objects (blue ellipses).
Using an automotive IMU (defined by $\sigma_v = 0.1$ m/s, $\sigma_\phi = 0. 1°$, $\sigma_{\phi'} = 0.005°$/s), the number of failures also decreased. Comparing the results (Figure 6), an automotive IMU also helps to bridge areas without landmark objects. However, for some areas with very few landmarks over a long distance the positioning accuracy was in the range of 1 m.
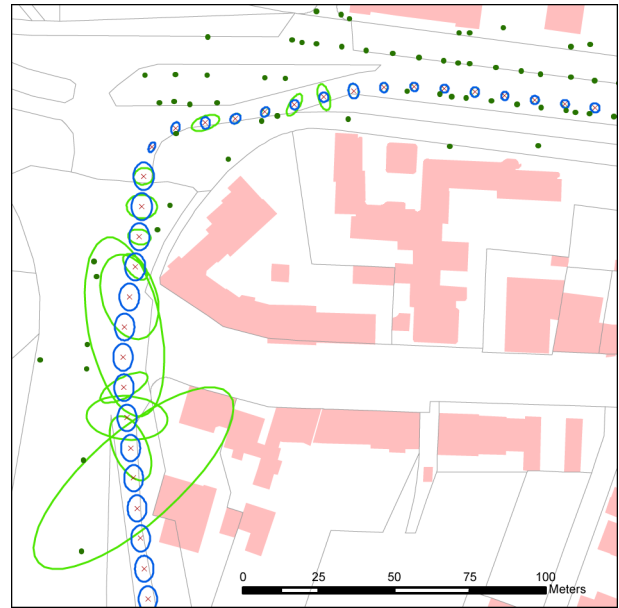

Figure 5. Positioning accuracy with filtered positions drawn as error ellipses for opening angle 100°, (blue: precise IMU, green: without IMU). All ellipses scaled by a factor of 20. (Upper right: green ellipses coincide with blue ellipses.)
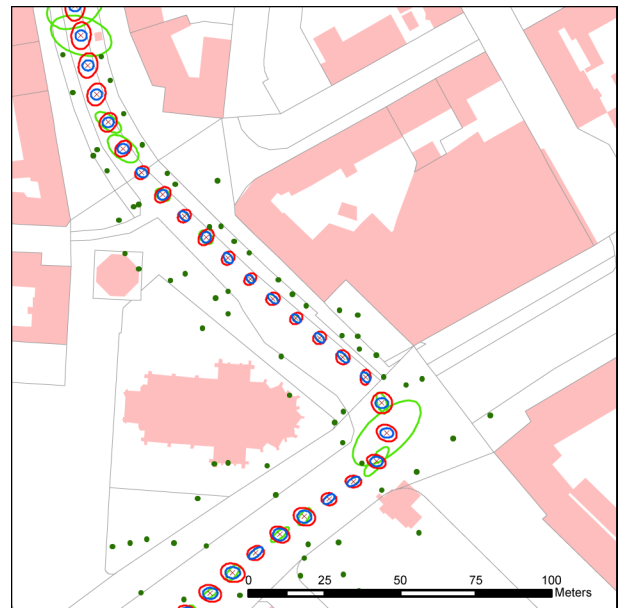

Figure 6. Positioning accuracy drawn as error ellipses for opening angle 100°, (green: without IMU, blue: precise IMU, red: automotive IMU). All ellipses scaled by a factor of 20. (Middle: green ellipses coincide with blue ellipses.)

Another important question is the maximum allowed distance between objects for positioning. Or to put it another way, how far can we drive without using any landmark objects for localization but still reach an acceptable positioning accuracy? Results are shown in Table 1, for a camera system with opening angle 100° and 0.1 m positioning accuracy at the starting point.

max. pos. error	Precise IMU	Automotive IMU
0.2 m	650 m	30 m
0.3 m	990 m	80 m

Table 1. Maximum distance without landmark update with a maximum positioning error of 0.2 m and 0.3 m.

4.3 Positioning Accuracy Maps

Based on the results of the simulation new maps of position accuracies were created (Figure 7, Figure 8, and Figure 9). These maps indicate the possible accuracy on each point along the roads. As shown here, the accuracy was only calculated along the driven path. From these maps, e.g. areas can be selected, where the positioning accuracy is lower (orange and red areas) and higher (green and yellow areas) than when using GPS. It is obvious, that in urban environments, especially at intersections, a lot of pole-like objects are present, which leads to a reliable and accurate positioning; on highways there are less of these objects, so often a positioning is impossible (Figure **7**, blue areas).
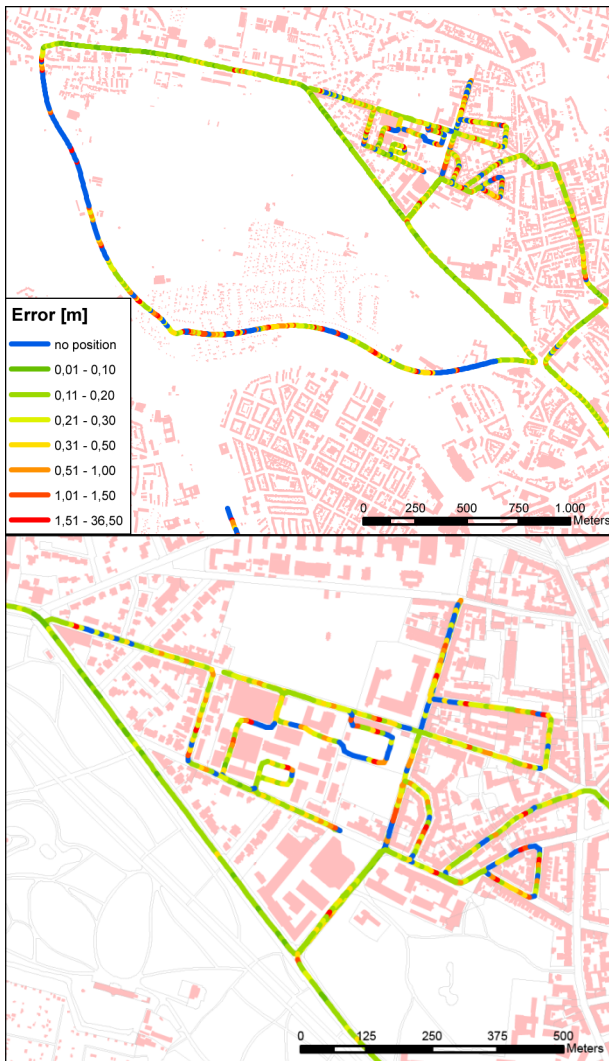

Figure 7. Map of accuracies without filter for a camera system with 100° opening angle.

Using the Kalman filter led to far better results. In Figure 8 the positioning accuracy in general is better than 0.3 m. With an automotive IMU the results were also improved (Figure 9). Even better results can be achieved when using all visible landmarks in the filter process. This is especially the case in densely-built up areas (compare Figure 7 and Figure 9, lower picture).

In general, however, GPS and landmark based navigation are somewhat complementary, as in open areas, where GPS works well, there are less possibly obtruding landmarks and vice versa.
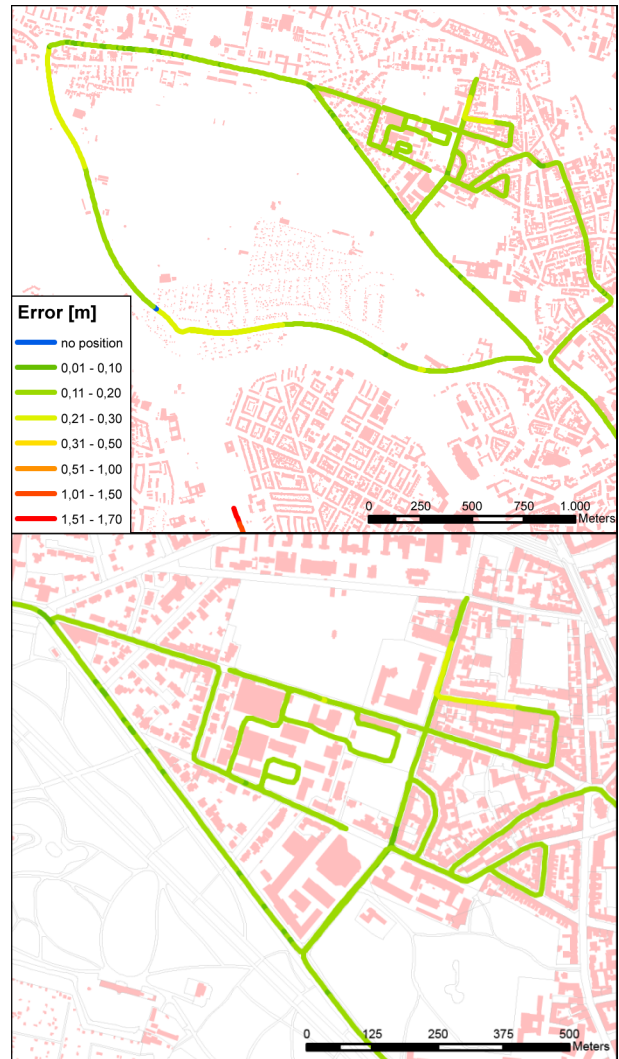

Figure 8. Map of accuracies using a precise IMU for filtering for a camera system with 100° opening angle. The accuracy along the highway (upper picture, left) depends on the direction of travel (from south to north).
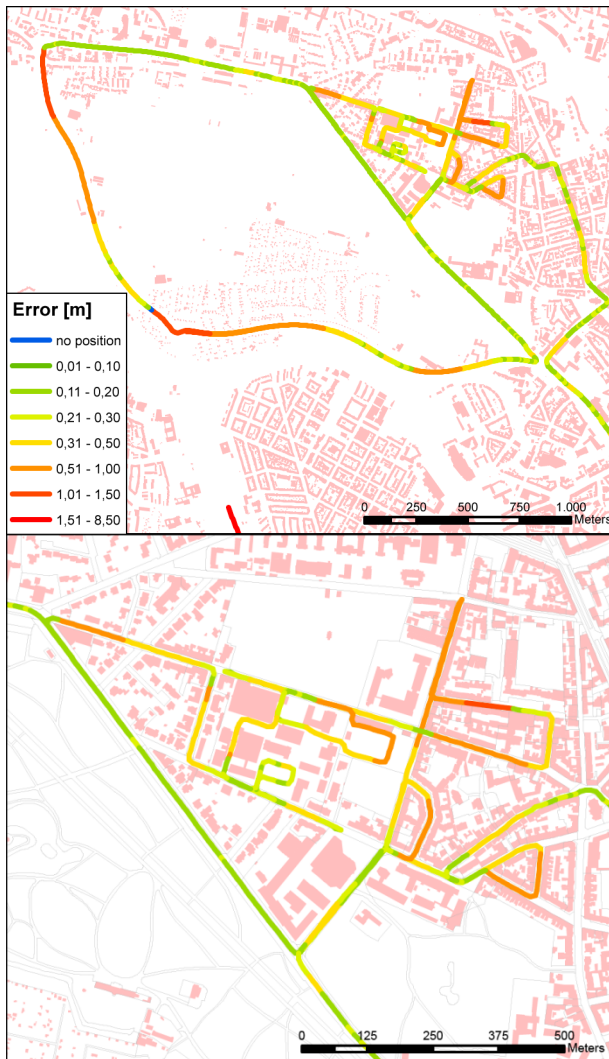
Figure 9. Map of accuracies using an automotive IMU for filtering for a camera system with 100° opening angle. The accuracy along the highway (upper picture, left) depends on the direction of travel (from south to north).

5. OUTLOOK

The results of the presented method show a high potential for localization based on landmarks. Especially at inner-city intersections, where lateral and longitudinal positioning accuracy is very important, the simulation gives promising results.

As already mentioned, positioning was not successful in all cases. In some areas especially when not using a filter the accuracy was low due to too few or unfavorably distributed objects. As shown in the second part of the simulation, using a standard Kalman filter helps to improve not only the positioning accuracy but also bridges the gap between landmarks where no positioning would be possible. Furthermore, using additional features, such as planes, which we can find in urban areas, will help to improve reliability and accuracy. We can expect, that the combination of GPS, an automotive IMU and landmarks leads to a reliable and accurate localization.

The current simulation reconstructed the positions based on stereo-reconstruction. Another simulation will be undertaken using only mono-images. Furthermore, we plan to verify the simulation with a data acquisition in a mobile mapping system.

6. REFERENCES

Brenner, C., 2009: Extraction of Features from Mobile Laser Scanning Data for Future Driver Assistance Systems, *Advances in GIScience: Proceedings of 12th AGILE Conference on GIScience, Lecture Notes in Geoinformation and Cartography*, Springer, Berlin, pp. 25-42.

Brenner, C., Hofmann, S., 2010: Evaluation of Automatically Extracted Landmarks for Future Driver Assistance Systems, *Proceedings of the Joint International Conference on Theory, Data Handling and Modelling in GeoSpatial Information Science*, Hongkong, Vol. 38, No. 2, pp. 361-366.

Brenner, C., 2010: Vehicle localization using landmarks obtained by a LIDAR mobile mapping system, *Proceedings of the ISPRS Commission III Symposium on Photogrammetric Computer Vision and Image Analysis*, Paris, pp. 139-144.

Burgard, W., Hebert, M., 2008: *Springer Handbook of Robotics.* Springer, chapter World Modeling, pp. 853-869.

Kremer, J., Hunter, G., 2007: Performance of the streetmapper mobile lidar mapping system in real world projects. *Photogrammetric Week 2007*, Wichmann, pp. 215-225.

Weiss, T., Kaempchen, N., Dietmayer, K., 2005: Precise ego localization in urban areas using laserscanner and high accuracy feature maps. *Proc. 2005 IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, pp. 284-289.