

EVALUATION OF AUTOMATICALLY EXTRACTED LANDMARKS FOR FUTURE DRIVER ASSISTANCE SYSTEMS

Claus Brenner and Sabine Hofmann

Institute of Cartography and Geoinformatics
Leibniz Universität Hannover
Appelstrasse 9a, 30167 Hannover, Germany
{claus.brenner, sabine.hofmann}@ikg.uni-hannover.de

KEY WORDS: LIDAR, Mobile Laser scanning, Extraction, Matching, Mapping, Accuracy, Navigation

ABSTRACT:

In the future, vehicles will gather more and more spatial information about their environment, using on-board sensors such as cameras and laser scanners. Using this data, e.g. for localization, requires highly accurate maps with a higher level of detail than provided by today's maps. Producing those maps can only be realized economically if the information is obtained fully automatically. It is our goal to investigate the creation of intermediate level maps containing geo-referenced landmarks, which are suitable for the specific purpose of localization.

To evaluate this approach, we acquired a dense laser scan of a 22 km scene, using a mobile mapping system. From this scan, we automatically extracted pole-like structures, such as street and traffic lights, which form our pole database. To assess the accuracy, ground truth was obtained for a selected inner-city junction by a terrestrial survey. In order to evaluate the usefulness of this database for localization purposes, we obtained a second scan, using a robotic vehicle equipped with an automotive-grade laser scanner. We extracted poles from this scan as well and employed a local pole matching algorithm to improve the vehicle's position.

1 INTRODUCTION

1.1 Driver Assistance Systems

In recent years, research and development in the area of car driver information and safety systems has been quite active. Beside the goal to provide information to the driver, active safety systems should completely prevent accidents, instead of just reducing the consequences. Therefore, vehicles are using more and more on-board sensors such as cameras, laser scanners or radar to gather spatial information about their environment. There are numerous examples for driver assistance systems, like adaptive cruise control, collision warning/ avoidance, curve speed warning/ control, and lane departure warning. Compared to car navigation, these so-called advanced driver assistance systems (ADAS) operate at a very detailed scale, requiring large scale maps, and thus, even more detailed field surveys. The NextMAP project has investigated the future map database requirements and the possible data capturing methods (Pandazis, 2002). The key point is to align the evolution of in-vehicle sensors with advances in mapping accuracy and map contents in such a way that sensible ADAS applications can be realized (at minimum cost). For example, additional mapping of speed limit, stop, priority, right of way, give way and pedestrian crossing signs, traffic lights, as well as the capture of the number of lanes, the road and lane width, and of side obstacles such as houses, walls, trees, etc. has been proposed. However, the NextMAP project has not dealt with the commercial aspects of producing such maps with very high detail and accuracy. Some investigations with regard to this have been made in the 'Enhanced Digital Mapping' project (EDMap, 2004). ADAS applications have been classified, according to their requirements, into 'WhatRoad', 'WhichLane', and 'WhereInLane', the latter requiring a mapping accuracy of ± 0.2 m. It was concluded that contemporary mapping techniques would be too expensive to provide such maps for the 'WhichLane' and 'WhereInLane' case.

A key observation is that these findings assume a 'traditional' map production, i.e. the map is a highly abstract representation

in terms of a vector description of the geometry, with additional attributes. However, depending on the application, this is not always necessary. For example, the accurate positioning of vehicles using relative measurements to existing map features does not necessarily require a vector map in the usual sense. Rather, any map representation is acceptable as long as it serves the purpose of accurate positioning. This means that map production to derive a highly abstract representation (which usually involves manual interaction) is not required, but rather a more low-level representation is suitable as well (which can be derived fully automatically and thus, inexpensively).

Also, accuracy considerations nowadays usually assume that the position of a vehicle relative to the map is obtained using an absolute measurement for both, typically using the global positioning system (GPS) for the vehicle and geo-referenced maps (possibly obtained by GPS measurements as well). However, when the vehicle uses relative measurements to known objects in the environment (such as poles and crash barriers), a very high absolute map accuracy is not required. Thus, one is relieved from unreasonable high requirements regarding absolute map accuracy and absolute vehicle positioning, nevertheless making relative centimeter level positioning feasible.

1.2 Mapping Based on Mobile Laser Scanning

Mobile mapping systems using laser scanners combined with GPS and inertial sensors are well suited for the production of large scale maps, since they reach a relative accuracy of down to a few centimeters. With a rate of more than 100,000 points per second, they capture the environment in great detail (Kukko et al., 2007).

The problem of building a representation of the environment using laser scanners has also been investigated in robotics (although, of course, laser scanners were not the only sensors considered). Iconic representations, such as occupancy grids, have been used as well as symbolic representations consisting of line maps or landmark based maps (Burgard and Hebert, 2008). One of the major problems to solve is the simultaneous localization

and mapping (SLAM), which incrementally builds a map while a robot drives (and measures) in unknown terrain (Thrun et al., 2005). Often, 2D laser scans (parallel to the ground) are used for this, but it has also been extended to the 3D case (Borrmann et al., 2008). However, it is probable that it will still take some time until such full 3D scanners are available in vehicles. What can be expected in the near future, though, are scanners which scan in a few planes such as the close-to-production IBEO Lux scanner (IBEO, 2009), which scans in four planes simultaneously.

Apparently, it would be unreasonable to provide dense georeferenced point clouds for the entire road network, as this would imply huge amounts of data to be stored and transmitted. On the other hand, as pointed out in the previous section, obtaining a highly abstracted representation usually requires manual editing, which would be too expensive. Following the work in robotics, a landmark based map would be a suitable approach. For example, to provide a map useful for accurate positioning, landmarks should fulfill the following requirements: (1) they should be unique in a certain vicinity, either by themselves or in (local) groups, (2) their position / orientation should be stable over time, (3) few of them should be needed to determine the required transformation with a certain minimum accuracy, (4) they should be reliably detectable, given the available sensors and time constraints.

In this paper we describe the use of pole-like objects as landmarks. We assess the accuracy of the extracted poles using a terrestrial survey and investigate the applicability of pole matching for the problem of vehicle localization.

2 DATA ACQUISITION

In the following sections, we describe the acquisition of the two data sets used. For our experiments, we obtained two different data sets of the same area in Hannover, Germany. The first data set is a very dense point cloud, obtained by the Streetmapper mobile mapping system and is described in section 2.1. The second data set was acquired by a vehicle, equipped with an automotive-grade scanner and is described in section 2.2.

2.1 Reference Data

A dense laser scan of a number of roads in Hannover, Germany, was acquired using the Streetmapper mobile mapping system. This system was jointly developed by 3D laser mapping Ltd., UK, and IGI mbH, Germany (Kremer and Hunter, 2007). The scan was acquired with a configuration of four scanners. Two scanners Riegl LMS-Q120 were pointing up and down at an angle of 20° , one was pointing to the right at an angle of 45° . Another Riegl LMS-Q140 was pointing to the left at an angle of 45° . The LMS-Q120 has a maximum range of 150 m and a ranging accuracy of 25 mm. All scanners were operated simultaneously at the maximum scanning angle of 80° and scanning rate of about 10,000 points/s. Positioning was accomplished using IGI's TERRAcontrol GNSS/IMU system which consists of a NovAtel GNSS receiver, IGI's IMU-IIId fiber optic gyro IMU operating at 256 Hz, an odometer, and a control computer which records all data on a PC card for later post processing to recover the trajectory. The scanned area contains streets in densely built-up regions as well as highway like roads. The total length of the scanning trajectory is 21.7 kilometers, 70.7 million points were captured. On average, each road meter is covered by more than 3,200 points.

2.2 Vehicle Data

The second scan was obtained by RTS Hanna, a vehicle which was developed by the Real Time Systems Group (RTS) at the

Institute for Systems Engineering at the Leibniz Universität Hannover, Germany (Figure 1). As main sensor they use two pairs of custom-built 3D laser range scanners for environmental detection, the 'RTS-ScanDriveDuo'. Each pair consists of two 2D SICK LMS 291-S05 laser scanners, which are mounted on the ScanDrive, a rotary unit. In vertical direction, the scanners cover an area of 90° , when reading remission values. Based on the center point of a single scanner, 20° of the upper half-space and 70° of the lower half-space are recorded. The angular resolution is 1° in vertical direction. The horizontal scan range of each scanner pair is 260° , with an angular resolution of approximately 2° . Taking the driving direction as 0° , the left scanner covers the area from -210° to $+50^\circ$, the right scanner covers the area from -50° to $+210^\circ$. With a rotational speed of $150^\circ/\text{s}$, it takes 2.4 s for a full rotation. Using a pair of scanners where each scanner only covers one half-space, the scanning time is reduced to 1.2 s. New scan lines are provided every 13 ms. Therefore, each 3D scan consists of 180 scan lines. The scanners have a maximum range of 30 m and a ranging accuracy of 60 mm.

For positioning, RTS-Hanna uses a Trimble AgGPS 114 without differential corrections. A speed sensor at the gear output and an angular encoder at the steering provide the odometer data.

For our experiments, all data, including GPS positions and headings, odometer data and scanned point clouds, were stored on a PC card for post processing. A single point cloud contains post processed data of one scanner rotation calculated relative to a fixed center point. Each scan consists of approximately 6,000 object points.

To compare results, this second scan was acquired along the trajectory of the first scan, excluding highway like roads because of speed limitations.



Figure 1: The RTS Hanna robotic vehicle. The point clouds were acquired using the four laser scanners mounted on the two ScanDrives on top of the vehicle. Image courtesy of RTS.

3 FEATURE EXTRACTION

We decided to use poles, such as poles of traffic signs, traffic lights, and trees, as landmarks. Poles are usually abundant in street corridors, especially in inner city scenes. It has been shown earlier by (Weiss et al., 2005), that a combination of GPS, odometry, and a (four-plane) laser scanner can be used to estimate the position of a vehicle, given a previously acquired map of poles.

The reported accuracies were in the range of a few centimeters.

In order to form matches, poles have to be extracted in both the reference scan and the vehicle scan, as described in the following two subsections.

3.1 Extraction of Poles from Streetmapper Data

The automatic extraction of simple shapes (like cylinders) from laser scanning data has already been done in other contexts, e.g. the alignment of multiple terrestrial scans in industrial environments (Rabbani et al., 2007). However, as can be seen from Figure 2, left, single poles are not hit by very many scan rays. Thus, methods which rely on the extraction of the surface, of surface normal vectors, or even of curvature, are not applicable.

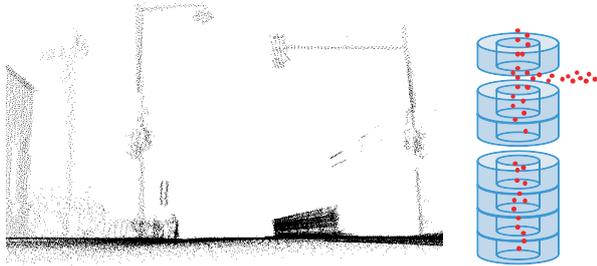


Figure 2: Left: Part of the Streetmapper scene containing a sign post, street (and traffic) light, and traffic light. Right: Illustration of the pole extraction algorithm using a stack of hollow cylinders.

Instead, we use a simple geometric model for pole extraction, assuming that poles are upright, there is a kernel region of radius r_1 , where laser scan points are required to be present, and a hollow cylinder, between r_1 and r_2 ($r_1 < r_2$) where no points are allowed to be present (Figure 2, right). The structure is analyzed in stacks of hollow cylinders. A pole is confirmed when a certain minimum number of stacked cylinders is found. After a pole is identified, the points in the kernel are used for a least squares estimation of the pole center. The method also extracts some tree trunks of diameter smaller than r_1 , which we do not attempt to discard.

In the entire 22 km scene, a total of 2,658 poles were found fully automatically, which is one pole every 8 meters on average. In terms of data reduction, this is one extracted pole per 27,000 original scan points. Although the current implementation is not optimized, processing time is uncritical and yields several poles per second on a standard PC. Note that this method owes its reliability to the availability of full 3D data, i.e., the processing of the 3D stack of cylinders. We classified the extracted poles manually and found 46% trees, 21% street lights, 5% traffic lights, 5% tram poles, 8% sign posts and other poles and 9% of non identifiable pole-like objects. The group of non identifiable objects contains all pole-like objects which are located along roads, but could not be identified definitely by manual inspection, e.g. since too few laser beams had hit the pole. The error rate (false positives) was about 6% on average, with higher rates of up to more than 30% in densely built-up areas. False positives are mainly due to occlusions which generate narrow bands of scan columns, similar to poles.

3.2 Extraction of Poles from RTS Hanna Data

Hanna scan data consists of scan frames, where one frame is a horizontal sweep of any of the four scanners. One orientation is given for each single frame. As the vehicle moves during the



Figure 3: Extracted poles (blue dots) in a larger area with buildings from the cadastral map (green). The thin dotted line is the trajectory of the Streetmapper van.

scan, scanpoints are motion compensated in such a way that the single, given orientation holds for all points of the frame.

Unfortunately, applying the same algorithm as in the Streetmapper case to extract poles does not work well. The major problem is the low point density. From the 3D point cloud of a single scan, it is almost impossible to tell if a column of stacked points is due to an actual pole or just a consequence of the low density (Figure 4). On the other hand, if one tries to combine several scans along the trajectory in order to obtain a higher density, the positioning errors will lead to several images of the very same pole, as can be seen in Figure 5. The extracted poles may be more than one meter apart, even if they were derived from successive scans.

Therefore, we extracted the poles from single scans, using the raw data rather than the 3D point cloud. Since the points are recorded in succession, column by column, a neighborhood can be defined on the row and column indices. This means that we perform our analysis on the depth image, indexed by the horizontal and vertical scan angles. One has to keep in mind, though, that this image is not an exact polar representation due to the forward movement of the vehicle during the scan. We extract poles by searching for vertical columns which have height jumps to both their left and right neighboring columns. The resulting point subsets are checked for a minimal height. Then, a principal component analysis is performed and the selected column subset is accepted based on the eigenvalues, equivalent to fitting a line and checking the residual point distances.

Figures 4 and 5 show poles which were extracted using this method. In total, 1,248 poles were detected in 1,384 scans of the first trajectory and 1,794 poles in 3,237 scans for the second trajectory. Due to the limited range and resolution of the scanners, there are around 50% (first trajectory) to 60% (second trajectory) of scans where no poles were found and around 30% of scans where only one pole was found. In both trajectories, the maximum number of detected poles in any scan was 6. Figure 6 shows the distribution of the number of scans in which n ($0 \leq n \leq 6$) poles were found.

Figure 7 compares the reference poles, extracted using the Streetmapper van, with the poles obtained from the RTS Hanna scans. The latter were extracted in each scan separately, as described above. They were then transformed to the global coordinate system using the transformation provided for each scan

4 DATA ANALYSIS

4.1 Verification of Reference Data

Before using the reference data provided by the Streetmapper system, we have to make a statement on the quality of this data. Therefore, an analysis of the accuracy of position of the extracted landmarks is required. Ground truth was obtained along a selected inner-city junction by a terrestrial survey, using a total station to measure every single pole manually. The accuracy of position is < 10 cm for every measured pole.

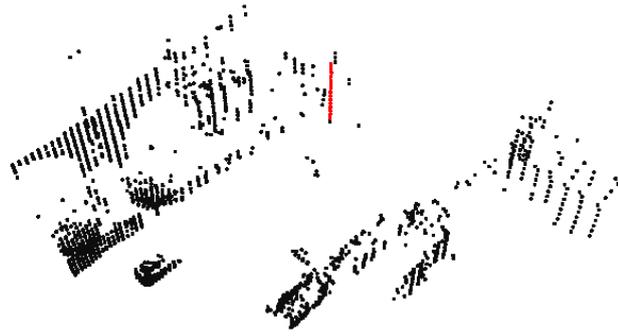


Figure 4: Single pole (red dots) extracted from RTS Hanna data (black dots). Ground points were removed for better visibility.

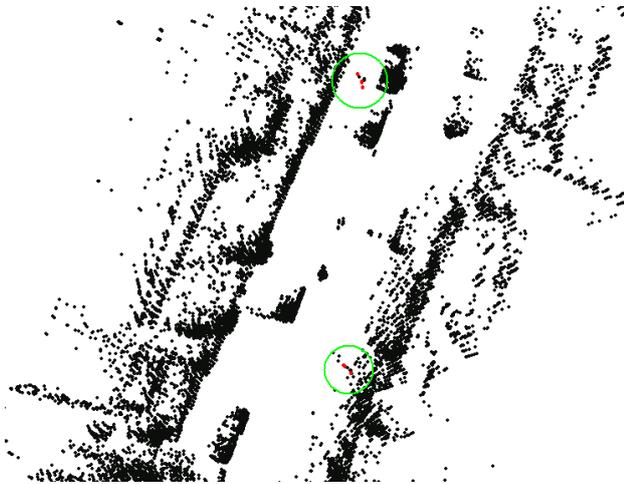


Figure 5: Poles (red dots inside green circles) extracted from RTS Hanna data (black dots). Ground points were removed for better visibility. Extracted poles appear at several locations due to the positioning error of the vehicle.

frame, which is based on GPS positioning and odometry. In Figure 7, left, the situation is especially obvious at the parking lot in the upper right, where a systematic shift is present (corresponding poles are marked by circles). Again, it can be seen that single poles appear multiple times in the RTS scans due to the low positioning accuracy. Figure 7, right, shows another scene at an intersection, where the reference poles are to the left and right of the street, whereas the poles extracted from RTS Hanna often appear inside the buildings or in the middle of the road. See Figure 12 for final results on these two scenes.

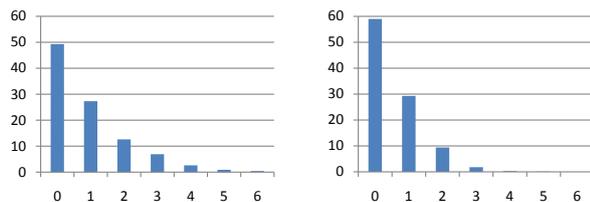


Figure 6: Histograms of the percentage of scans with n extracted poles, for the first (left) and second (right) trajectory of the RTS vehicle.

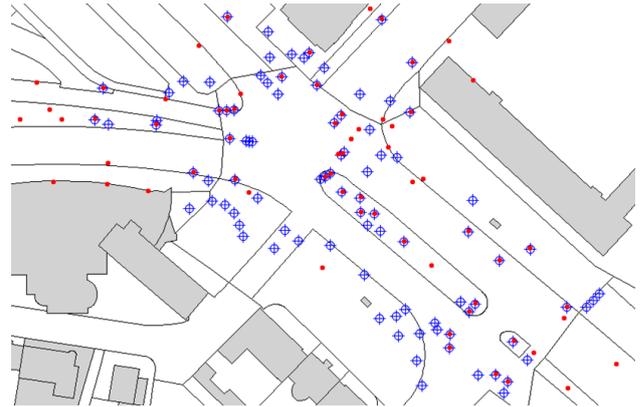


Figure 8: Accuracy of position: ground truth obtained by a terrestrial survey (blue crosshairs) compared to automatically extracted poles (red dots) of the reference map.

As ground truth and extracted reference poles are both given in DHDN 1990, 3rd meridian strip, calculating the residuals is very easy. We found 28 poles with a maximum distance of 32 cm from their corresponding pole. Using larger search radii only leads to mismatches between poles. Calculating the residuals using these points, we achieved a maximum distance between corresponding poles of 22.9 cm and a minimum distance of 1.4 cm. The root mean square error was 12.1 cm. Figure 9 shows the error distribution.

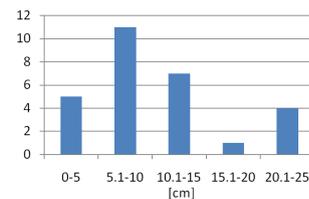


Figure 9: Histogram of pole measurement errors.

Comparing the extracted poles to the ground truth also revealed the problem of up-to-dateness. As shown in Figure 10, some of the extracted poles, marked with red crosses, have a large distance to ground truth poles, marked with black circles (the outer circle has a distance of 50 cm to the center of the pole). In this case there is no gross error in the feature extraction algorithm, but the junction was rebuilt between scanning the area and measuring poles manually.

5 EXPERIMENTS

The goal of our experiment is to find out if the position of the vehicle can be determined using a matching of poles. To this end, we implemented a local matching algorithm, as described in the following (see also Figure 11).



Figure 7: Comparison of reference poles (green squares) and poles extracted using the RTS Hanna vehicle (red dots).

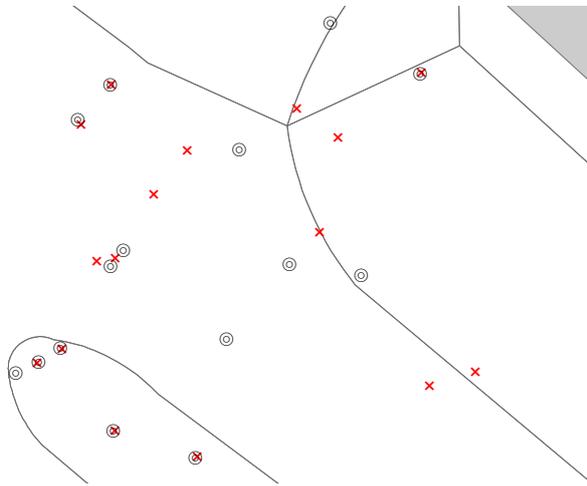


Figure 10: Detailed view of the junction: large distances between extracted poles (red crosses) and ground truth (black circles, radius: inner circle 25 cm, outer circle 50 cm) are not measurement errors but due to a rebuild of the junction.

The vehicle maintains a current transformation, T , and a current set of active poles, P , which is a set of 2D points. As the vehicle moves forward, new scans become available and some of them lead to newly detected poles. Those poles are then added to the active set. At the same time, old poles above a given horizon distance ε_h are removed from the active set. The reason for this is that the positions in the active set can only be assumed to be approximately correct within a certain distance, as their transformation is based on the given scan frame orientation, which uses GPS and odometry.

Whenever the active set is updated, a match to the global database of poles, Q , is attempted. For each point in the active set, all neighbors in the database within a distance smaller than a search distance ε_s are retrieved. For any translation vector between a pole in the active set and a corresponding element among its neighbors, the active set is transformed and the number of ‘hits’ of all poles in the active set to the database are counted. A hit is assumed when the distance is below a ‘matching’ threshold ε_m . This is similar to random sampling consensus (RANSAC), however, due to the low number of elements we can afford check-

-
- 1: Q is the reference point set
 - 2: Initialize the transformation T to identity
 - 3: Initialize the current active set of points P to be empty
 - 4: **while** new points p become available **do**
 - 5: $P \leftarrow P \cup \{p\}$
 - 6: Remove all points p' from P with $\text{dist}(p', p) > \varepsilon_h$
 - 7: For each point $p \in P$ compute
 - $N_\varepsilon(p) \leftarrow \{n \mid \text{dist}(n, p) < \varepsilon_s\}$
 - 8: **for** all pairs p, q with $p \in P, q \in N_\varepsilon(p)$ **do**
 - 9: Compute transformation T' mapping p to q
 - 9: Map all $p \in P$ using $p' = T'(p)$ and count all p' for which $\exists n \in N_\varepsilon(p)$ with $\text{dist}(n, p') < \varepsilon_m$
 - 10: Let T' be the transformation with the highest count
 - 11: Find closest point pairs using T' and determine $T \leftarrow \text{best L.S. estimation based on those pairs}$
-

Figure 11: Algorithm used for local point pattern matching.

ing all possible translation vectors instead of having to pick them randomly. The transformation T' with the largest consensus set is estimated using a least squares estimation of a similarity transform. It is then taken as the new local transformation T . We used $\varepsilon_h=100$ m for the horizon, $\varepsilon_s=15$ m as search radius, and $\varepsilon_m=2$ m as match radius in our experiments.

The results for the two scenes of Figure 7 can be seen in Figure 12, which shows the positions of the vehicle according to the given orientation per scan frame (red dots) and the corrected position after applying the transformation T (green squares). For example, in the left image, the trajectory starts at the parking lot, where both positions are identical in the beginning (only the red dots are visible). After a few meters, the poles are matched and the green trajectory is corrected to its true position. The corrected trajectory is also clearly better in the left and lower part, where the original trajectory obviously suffers from GPS shadowing by the trees. Similarly, in the right image, the original coordinates are far off the correct lane not only after the U-turn, but also when driving straight on (presumably also due to shadowing). As there is no ground truth for the real trajectory, e.g. obtained with a GPS/IMU system with substantially higher accuracy, our analysis does not include an evaluation for the entire trajectory. Clearly, there are also many parts in the trajectory where our lo-

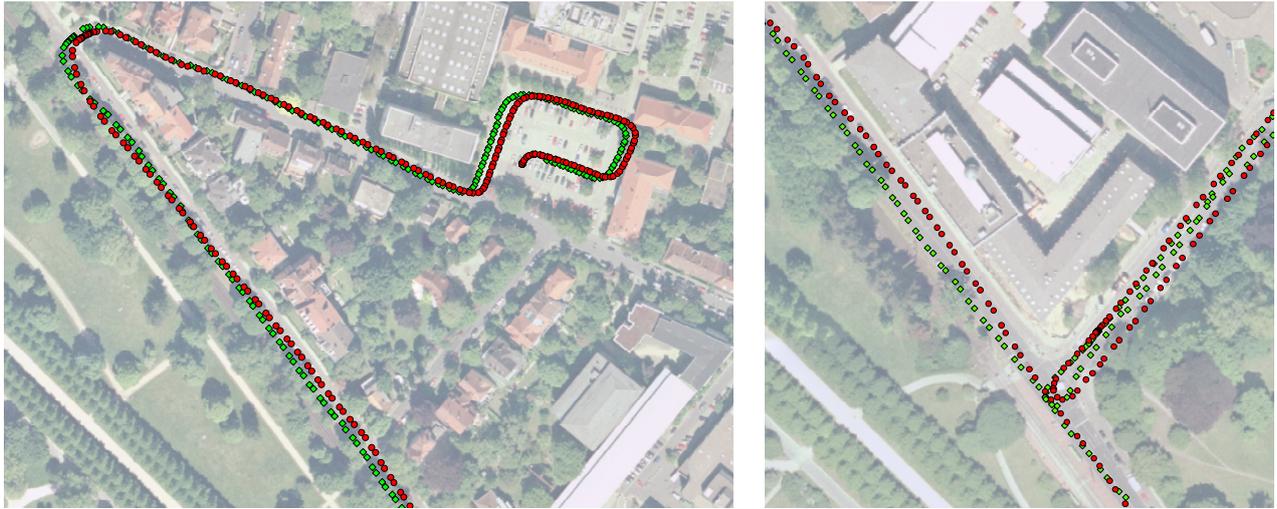


Figure 12: Comparison of original vehicle trajectories (red dots), as obtained from the RTS Hanna GPS/odometry positioning system and corrected trajectories (green squares), using pole matching.

cal matching algorithm failed so far. Sometimes, too few poles lead to failure. On the other hand, along the alleys, there are often too many poles and our local matching algorithm finds wrong correspondences, leading to a wrong transformation.

6 DISCUSSION AND OUTLOOK

In this paper, we presented an approach to use automatically extracted landmarks for the positioning of vehicles. We focused on pole-like objects and extracted them both in a high-resolution, high accuracy scan (the reference database) as well as in low-resolution, low accuracy scans of a robotic vehicle, using two different algorithms. We assessed the accuracy of the extracted poles and made first experiments regarding the improvement of the trajectory using a local pole matching.

As we mentioned, pole matching did not succeed in all cases. Too few as well as too many poles may cause problems, and erroneous poles in the vehicle scans and in the database lead to wrong feature associations and thus, transformations. We think that a key to a successful matching is the short-term stability of the vehicle trajectory. If a high precision of the poles can be obtained locally, many false matches can be ruled out by tight distance bounds. However, even then, there is sometimes the problem of too few poles in the database and it is obvious that a key element for a reliable localization will be to include additional landmarks.

For the future, we plan to include such additional features, e.g. local planes. Also, the local matching algorithm should be extended and all measurements should be included in a Kalman filter solution to estimate the vehicle position.

ACKNOWLEDGEMENTS

We thank Matthias Hentschel from the Real Time Systems Group at the Institute for Systems Engineering of the Leibniz Universität Hannover for providing the RTS Hanna scans.

REFERENCES

Borrmann, D., Elseberg, J., Lingemann, K., Nüchter, A. and Hertzberg, J., 2008. Globally consistent 3D mapping with scan matching. *Journal of Robotics and Autonomous Systems (JRAS)* 56(2), pp. 130–142.

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

EDMap, 2004. Enhanced digital mapping project final report. Technical report, United States Department of Transportation, Federal Highway Administration and National Highway Traffic and Safety Administration, <http://ntl.bts.gov/lib/jpodocs/repts.te/14161.htm>, (accessed 30 Nov. 2009).

IBEO, 2009. Ibeo Lux laser scanner. <http://www.ibeo-as.de> (accessed 30 Nov. 2009).

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

Kukko, A., Andrei, C.-O., Salminen, V.-M., Kaartinen, H., Chen, Y., Rönnholm, P., Hyypä, H., Hyypä, J., Chen, R., Haggren, H., Kosonen, I. and Čapek, K., 2007. Road environment mapping system of the finnish geodetic institute – FGI roamer. In: *IAPRS (ed.), Proc. Laser Scanning 2007 and SilviLaser 2007*, Vol. 36 Part 3/W 52, pp. 241–247.

Pandazis, J.-C., 2002. NextMAP: Investigating the future of digital map databases. In: *e-Safety Congress, Lyon*, http://www.ertico.com/en/activities02/completed_projects/websites/nextmap_website.htm (accessed 30 Nov. 2009).

Rabbani, T., Dijkman, S., van den Heuvel, F. and Vosselman, G., 2007. An integrated approach for modeling and global registration of point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 61(6), pp. 355–370.

Thrun, S., Burgard, W. and Fox, D., 2005. *Probabilistic Robotics*. The MIT Press, Cambridge, Mass.

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