High Definition Mapping for Inland Waterways: Techniques, Challenges and Prospects

Lukas Hösch[1], Alonso Llorente[1], Xiangdong An[1], Juan Pedro Llerena[2], Daniel Medina[1]

***Abstract*— Inland waterway transport (IWT) is an efficient way of mass good transportation with low energy consumption and reasonable ecological impact. As for other transport systems, there is a need for reliable advanced driver-assistance functions and increased autonomy. In this regard, High Definition (HD) Maps play a role by enhancing vessels' localisation and perception capabilities. Obtaining HD Maps is an important step towards the ongoing digitalisation of water-based transport and forms the basis for assistant systems to captains of inland vessels or even higher automated ships. This work discusses the workflow for HD map generation by using vessels as information platforms (i.e., by having precise localisation and environment-detecting perception capabilities). An architecture for HD map generation from vessels provided with geodetic equipment for precise localisation and environment-detecting perception is presented. An overview on standard techniques for HD mapping within the automobile and robotics domain are discussed, as well as the particular challenges and needs present for water-based applications. Finally, the conceptual use and appropriate data exchange of HD Maps are drafted.**

I. Introduction

With the volume of freight traffic increasing over the past few years, road- and rail-borne transportation tend to reach their capacity limits. Inland waterway transport (IWT) constitutes an appealing alternative to land-based transportation. Being a mode of carriage with a long history, IWT offers the advantages of high safety, low energy consumption and cost-efficiency. With the currently used large inland vessels and comparably long transportation time, IWT is most suitable for the transportation of bulky and heavy freight. In addition to the efficiency of mass transportation, the reduced effect on habitat damage of inland waterways emphasises the environmental safety of this mode of transport.

With the goal of increasing the operation volume for IWT, the safety on the related infrastructure needs to be ensured. Indeed, the catastrophic effects of an important waterway being blocked were illustrated by the container ship *Ever Given*, causing severe traffic obstruction of the *Suez Canal* in March 2021. Even though, the economic effects of a blocked inland waterway may be less harmful in comparison, safe traffic circulation on inland waterways is an important goal. One contribution to this goal is to ensure accurate and well-updated chart data.

Within the context of the German-national research project *AutonomSOW* [1], we treat the vessel as *information platform* to provide services for predictable and networked transport processes on inland waterways. Thus, a vessel becomes an information platform when being able to reliably and precisely estimate its pose and perceive its surroundings. To that end, robust and multi-modal sensor fusion is required to obtain: *a)* geo-referenced positioning and attitude via multi-antenna Global Navigation Satellite Systems (GNSS) and inertial measurements; *b)* environment perception and relative distance measuring with sensors such as cameras, LiDARs or RADARs. Following the data collections by the information platform(s), the process of estimating the spatial HD mapping can be carried out. The derived HD maps can assist other vessels in safe navigation of the inland waterway. In fact, the current chart data is available on online platforms. However, this data is limited to static representation and can only be updated after the execution of a dedicated (and costly for the geodetic national representatives) survey campaign. Thus, the proposed information platform processing provides up-to-date and low-cost HD maps that can enhance the safety operation for IWT. The overall workflow is illustrated in Fig. 1: first, the information platform collects geo-referenced measurements from its surrounding; then, the data is cleaned and processed into a 3D HD maps; finally, the estimated maps are sent to the users and the new information is incorporated in the chart.

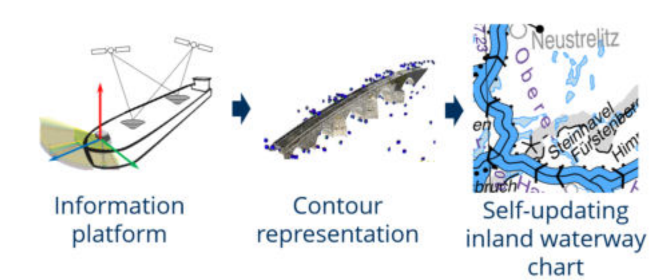

Fig. 1. Architectire of a selft-updating inland waterway chart. Images are supported by [2] and [3].

The rest of the paper is organised as follows. Section II provides an overview of recent activities and applications regarding SLAM (Simultaneous Localisation and Mapping) solutions for the automotive domain and discusses their applicability to IWT. In section III approaches for geo-referenced navigation and visual perception are presented and their results are evaluated. Furthermore, the generation of the BerlinIWT dataset is introduced. A final architecture of an IWT-related mapping approach is debated in section IV. Finally, section V gives an outlook on future activities.

[1]Authors are with the Institute of Communications and Navigation of the German Aerospace Center (DLR), name.surname@dlr.de

[2]Juan Pedro Llerena Caña is with the Applied Artificial Intelligence Group (GIAA), Carlos III University of Madrid, llerena@inf.uc3m.es

II. Related work

The interest on HD maps has substantially increased along with higher autonomy levels for vehicular applications. Thus, the first references to digital, 3D or HD maps can be found in the context of mobile mapping systems (MMSs) for automobile cases [4]–[6]. From a research perspective, one finds contributions on HD mapping in the fields of photogrammetry, vehicular technology and robotics. This section provides an overview on the relevant works for HD mapping in relation to MMSs in automotive scenarios and Simultaneous Localisation and Mapping (SLAM) in robotics. Finally, the characteristic challenges for HD mapping at IWT are analysed.

A. HD maps for road-based vehicles

The current HD mapping research is focused in a top-down approach, where roads are represented as imaginary centre lines, whom new attributes are added depending on the complexity of the scenario. These attributes could be traffic lights, road lines and the shape of the road in case only static objects are considered. As the requirements increase, new attributes are added in the map.

In the perspective of projects such as OpenStreetMap (OSM) [7], which exemplifies the concept of Volunteered Geographic Information (VGI), individuals actively contribute their local knowledge and data to collectively construct a comprehensive geo-database. The data sourced from OSM has been employed in various applications, such as vehicle localisation using 3D laser scanning [8].

Another example can be found in [9] which introduces the open-source framework Lanelet2. In addition to the common purposes in autonomous driving, localisation, and motion planning, this work also enables potential applications of maps to achieve highly automated driving (HAD). The Lanelet2 map was employed in [10] for the evaluation of semantic localisation.

B. Methods for Simultaneous Localisation and Mapping

The SLAM problem addresses the task of mapping an unknown environment while concurrently establishing the target's pose within it through the use of perception systems. Most recent activities distinguish Full SLAM from Online SLAM [11].

- *Full SLAM* involves estimating all the robot states and map computations by considering a set of all observations and controls taken from the whole path. This post-processing procedure comes with drawbacks regarding the computational effort.
- *Online SLAM* aims to estimate the robot's path and the map at each time step in real-time. Computational efficiency is increased by avoiding to consider the entire duration.

Further variables in the definition of the SLAM problem are [11]:

- *Map approach*. In volumetric SLAM, the map is sampled at a high resolution to achieve a realistic reconstruction of the environment. In feature-based SLAM, the algorithm tends to be more efficient, since only certain objects represented in the map.
- *Environment considerations*. Most activities assume static environments [12], [13], [14], where dynamic motions are considered as noise or need to be excluded [15]. In contrast, [16] and [17] offer dynamic SLAM solutions.
- *Exploration strategies* A passive SLAM does not involve path planning as an external entity has control over the robot. In active SLAM as described by [18] the robot actively explores its environment.

The approach of GMapping [19] is one of the most well-known SLAM algorithms in the community. It uses Rao-Blackwellized particle filters to construct a 2D grid map. More recent examples for passive online SLAM are mentioned in the following. LeGO-LOAM [20] is a lightweight and ground-optimised LiDAR odometry and mapping method estimating a 3D volumetric map. RTAB-Map [21] is an open-source library that combines LiDAR and visual information for large-scale and long-term online mapping operations suitable for the generation of 3D volumetric maps as well as 2D maps. LIO-SAM [22] integrates real-time LiDAR and IMU data to construct a 3D volumetric map in complex environments.

C. Challenges for IWT-related SLAM

The outdoor mapping approach pursued within the *AutonomSOW* project differs from most indoor SLAM algorithms developed for robotic applications. First, IWT is an outdoor activity and GNSS allow for global positioning such that the mapping problem get simplified. Also, inland vessels travel far distances in the same direction. In comparison to road vehicles or robots, the time until known sections are passed again is significantly longer. This may pose challenges regarding long-term memory and recognition of known features, as well as loop-closure techniques implemented in many well-known SLAM solutions. Another challenge in IWT-related environments is the standardisation of visual features. For instance, in automobile urban scenarios the scan of the surrounding is classified into edges and planar surfaces before being mapped to the appropriate key-frame. Thus, adjacent buildings and other planar structures can easily be recognised. Additionally, lane markings are well-defined and consistent across different locations. On inland waterways in contrast, the overall environment appearance is very irregular. Most of the time, the shore is covered with trees and smaller vegetation making the extraction of edges and planar surfaces complicated. Complex structures like wide bridges and waterway locks change the general appearance of the scene. The comparably low number of waterway locks makes it challenging to ensure generally valid mapping in these environments. Factors like varying water levels and large vessels moored in vicinity to the sensor platform heavily affect the detected ranges and complicate the recognition of known features.

III. Generation of HD maps

The inputs for the map generation in the presented application can be divided in two main parts: precise navigation and visual perception. Fig. 2 provides an overview of the sensor inputs and the further processing. The *Position, Navigation and Timing Unit* or PNT Unit provides geo-referenced pose information and time synchronization. The *Perception Unit* receives spatial mapping data from one or multiple LiDAR sensors as well as camera data for semantic scene understanding. The resulting point cloud, segmentation solutions and estimated vehicle's pose are finally fed to the *Map Processing* step.

In our information platform, the Perception Unit uses *ROS (Robotic Operating System)* as middleware and the processing is done in Python and C++. The PNT Unit is a DLR-developed platform that uses a C++ *RT (real-time) Framework* for processing the inertial and GNSS data. The rest of the section details the processing performed by the three modules afore-described.

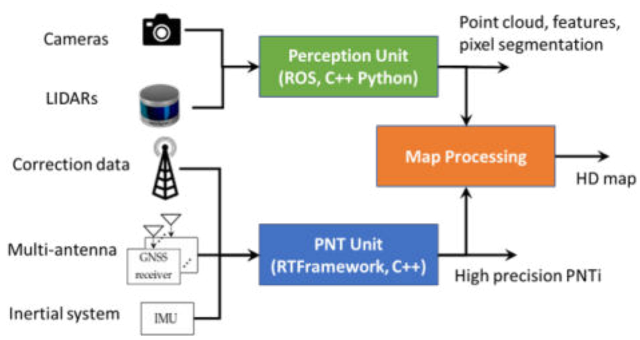

Fig. 2. Overview of the sensor inputs for HD map generation within our information platform.

A. Geo-referenced navigation

Geo-referenced navigation is the determination of a vehicle's dynamics (position, speed and orientation) over time with respect to a global geographic frame. GNSS and inertial measurement unit (IMU) constitute the primary sensors used for geo-referencing, such that GNSS allows for the *global* localization of the vehicle, while IMU tracks subtle and sudden motion and bridges short GNSS outages. By exploiting GNSS carrier phase observations and correction data, advanced techniques such as Real Time Kinematic (RTK) and Precise Point Positioning (PPP) provide decimetre-to-centimetre level positioning accuracy [23]–[25]. A hybridisation denoted as PPP-RTK combines the conventional PPP filtering with regional corrections and integer ambiguity resolution to provide quasi-instantaneous cm-accurate positioning even in remote locations (i.e., when base stations for RTK processing are unavailable) [26], [27]. In other words, PPP-RTK is based on precise satellite orbit, clock, signal biases and optional atmospheric products from a GNSS network of stations, represented as SSR (State State Representation) information and then broadcast to users.

The dependencies on GNSS goes beyond positioning and includes the provision of time and attitude information. Thus, a GNSS receiver is used as control unit for time synchronization (using, for instance, a local network time protocol (NTP) for the connected sensors and subsystems), triggering sensors and time-stamping the collected data. Also, absolute and drift-less orientation information can be obtained from vehicles equipped with a multi-antenna setup. The attitude accuracy is proportional to the distance between antennas, which becomes a limiting factor for miniaturized platforms but it is convenient for vessels' applications. The joint position and attitude (JPA) [28] or the array PPP [29] models consists on exploiting the inertial and GNSS observations across multiple antennas for an enhanced navigation performance.

One disadvantage of GNSS-based navigation is that it is vulnerable under a harsh environment and cannot obtain continuous positioning solutions, e.g. passing through a bridge in Fig. 3. In the illustrated example, our research information platform achieves sub-degree and cm-level orientation and positioning accuracy, respectively. The estimation is based on a multi-antenna PPP-RTK solution with inertial integration.

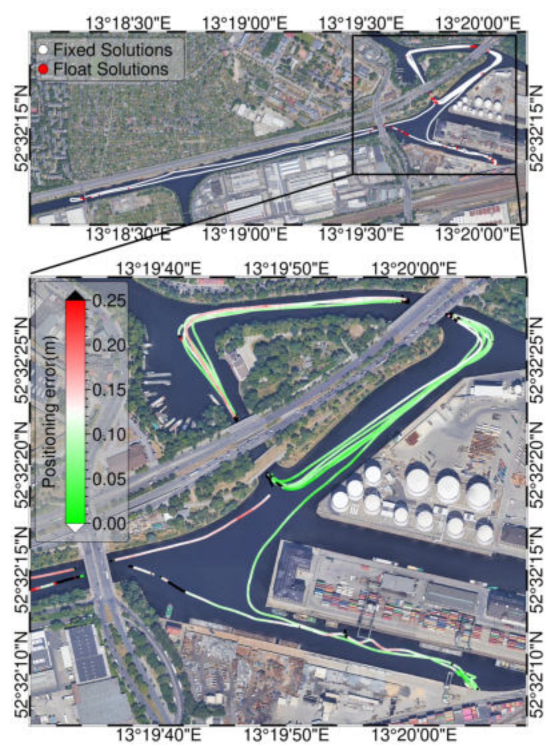

Fig. 3. Positioning performance for the PNT Unit on-board DLR's vessel AURORA. Data collection performed in Berlin, November 2022.

B. Visual perception

In the general usage of visual perception systems, the following information levels can be distinguished from the perspective of a user [30]:

- Physical description: pose, speed and shape of objects
- Semantic description: categories of objects

- Intention prediction: likelihood of the object's behaviour.

While the level of intention prediction is mostly related to autonomous driving, for the application of inland waterway spatial mapping, both physical and semantic description are required: semantic description is needed to distinguish relevant infrastructure from other objects and classify it appropriately. Providing the precise shape of objects, physical description in required for the surveying process.

The physical description of an object is comparably easy using point clouds derived from a LiDAR sensor. For instance, Fig. 4 shows a bird's eye view (BEV) for the data collected by one of the platform's LiDARs, that has been acquired on the urban waterways of Berlin in November 2022. The height of the points is encoded in the colourscale. The sensor's position is indicated by the red triangle in position $0, 0$, pointing in positive x-direction. Points are represented following the right-hand-rule. In this scenario, the furthermost point is located at a distance of 120 m from the sensor, which is surprising, as it exceeds the maximum range provided from the manufacturer. In direct vicinity to the sensor's position, a number of closely located points can be seen, which refer to reflections from the vessel's superstructures and sensor mountings. The water surface does not provide any reflections. However, river banks and shore structures can be distinguished on either side of the waterway, making the determination of navigable area rather straight-forward. Indeed, similar conditions prevail on most inland waterways. Still, waterway crossings, wide rivers or lakes lead to the effect that at least parts of the shore area are not detected correctly, which can lead to problems when distinguishing between navigable from (possibly) occluded areas. In this plot, parts of the river banks are marked with negative heights as the frame centre is located at the position of the sensor, which was mounted onto the sensor platform. Consequently, parts of the river bank are located below the sensor position. Behind the river banks, a number of higher points (located in the order of 10 to 30 m) can be seen. These reflections refer to trees, bushes or buildings. In the right part of the plot, the structures of a bridge are visible approximately 60 m ahead of the sensor position. Beside the edge of the bridge perpendicular to the waterway, two more lines of points are remarkable, continuing from the edge of the bridge into the x-direction. The associated echos are reflected from under the bridge.

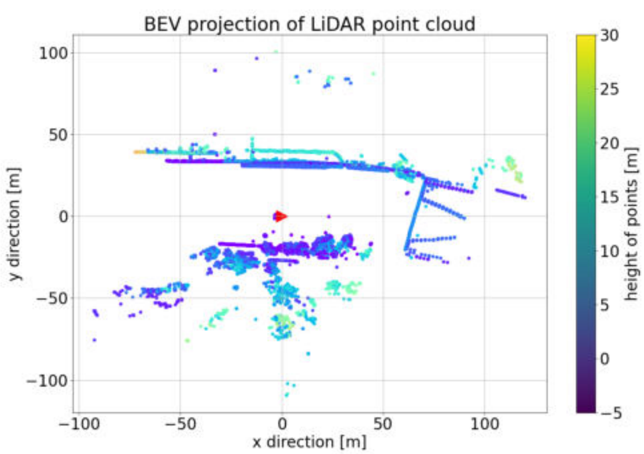

Fig. 4. BEV projection of a Velodyne point cloud. Vessel position is (0,0), pointing in the x-direction (red triangle). Due to the sensor mounting, some points are located below the z-plane.

As can be seen from the point cloud in Fig 4, essential scene understanding is indeed possible also from LiDAR point clouds, but requires prior knowledge about the scene. Full semantic scene understanding from point clouds only is therefore unfeasible.

Camera-based RGB images offer rich semantic information and can therefore be applied easily for the level of semantic description. Stereo-cameras could also be used for distance estimation, but exhibit lower maximum range than LiDAR sensors and decreased accuracy performance. It is therefore common practice to combine camera and LiDAR sensors in order to leverage the flaws of each other [30]. Semantic segmentation describes the task of assigning every single pixel in an image to a pre-defined class. During the process, pixels, that belong to the same semantic class are grouped together and assigned to the same class [31]. In the literature, this task is also referred to as pixel-wise classification [32]. The model used in our segmentation problem is an extension of the Deeplab algorithm, the DeeplabV3, suggested by [33], that has been designed specifically for semantic segmentation.

Semantic segmentation is a typical machine-learning (ML) task. For the performance of a NN, it is crucial to carry out the training process with a dataset of sufficient size and variety suitable for the application. In contrast to automotive applications, the availability of training datasets for water-based transportation is extremely limited, especially for semantic segmentation. Various datasets collected for inland shipping applications aim to identify navigable areas and therefore focus on water segmentation only [34], [35]. The MaSTr1325 dataset [36] has been collected in the coastal waters around Koper, Slowenia to support the development of small unmanned surface vehicles (USV). While the coastal, marine environment differs significantly from most inland waterway scenes, the variety of classes provided in MaSTr1325 is an advantage to our application. Still, as the categories are limited to "water", "sky", "obstacle", "void", and "uncertain", the dataset does not offer enough class labels for infrastructure-related semantic segmentation.

For the first application of bridge surveying and mapping, we are therefore currently developing our own dataset. The goal of the BerlinIWT dataset is detailed semantic segmentation of IWT-related infrastructure. For data acquisition, a first measurement campaign has been carried out on the urban waterways of Berlin during two days in June 2022. From the 11 h of data, the most relevant sections have been chosen. Sections are considered as relevant based on a high number of bridges that have been passed, while ensuring the largest possible variety of the captured scene

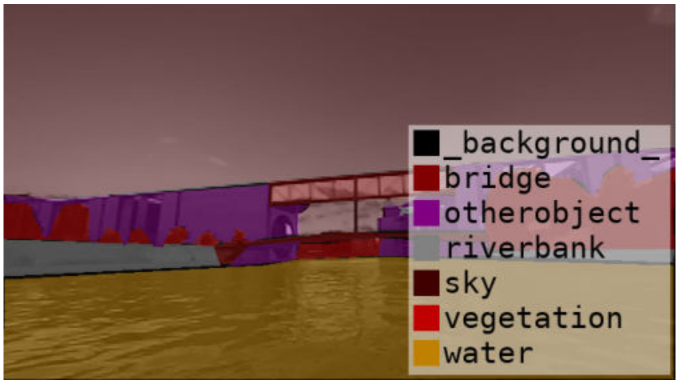

Fig. 5. Labelled example from the BerlinIWT dataset under development. Not all of the defined classes were used in this picture.

(i.e. weather conditions, width of the waterway, distance to the bridge). Additionally, also images recorded inside waterway locks are selected for annotation. In total, 200 images have been chosen. For this first attempt, a small team of four in-house annotators worked on the data for two weeks. Depending on the complexity of the image, the annotation of a single took between 10 and 20 min. Revising the annotated data with the means of quality control left 171 images that can be used for training. In order to avoid constraints on the possible use-cases of the dataset, a number of IWT-infrastructure related classes are defined, such as "bridge", "river bank", "water mark" (navigational marks and aids) and "waterway lock". Additionally, the classes "water", "sky", "object" and "vegetation" are defined. However, for the application in discussion, only the classes "water", "sky", "object" and "bridge" are used. Unlabelled pixels are classified as background.
With its extremely small size, the BerlinIWT dataset alone is not feasible for training a NN. It is therefore combined with the mentioned MaSTr1325 dataset [36] for coastal navigation. Further details on the exact training conditions are beyond the scope of this work. Nevertheless, we would like to emphasise the decent results achieved after a computation time of 20 h: The mean intersection over union (IoU) assessed on the test set amounts to 93.12 %. The IoU is computed from the true positives divided by the sum of true and false positives and false negatives. The mean IoU is the value of the IoU across all classes [37]. As the generation of the BerlinIWT dataset is an ongoing activity, an increase of the available training data is expected. Furthermore, the result can be improved by further hyperparameter optimisation and longer training on a more powerful machine. Still, the achieved results suffice for essential semantic scene understanding required for the described mapping purpose. Therefore, very next activities will focus on extrinsic calibration of LiDAR and camera in order to incorporate the segmentation solution into LiDAR point clouds.

C. Map estimation

With all the required inputs (segmented spatial mapping data as well as precise position and attitude information) in place, the conditions for map generation are fulfilled. Indeed, the acquired data from an information platform is likely to be less precise than comparable map information collected by professional survey teams. This drawback is expected to be leveraged by the quantity of the measurements: the goal is not to equip a single inland vessel with the necessary sensors, but rather to use a sufficient amount (or if possible even all) inland vessels as sensor platforms. On the one hand, a large amount of sensor platforms ensures that data can be acquired simultaneously in different places and consequently is likely to be up to date. However, the expectation that different sensor platforms pass through the same sections of inland waterways multiple times is more important at this point. This effect can be exploited to interpolate the range measurements numerically and overcome the drawback of decreased accuracy. Furthermore, various surveys of the same sections are a robust method to filter noisy measurements and undesired artefacts, which may even allow to use static SLAM algorithms (as described in section II-B) in conditions with large dynamic influence. As real-time capabilities are not required for the application Full SLAM algorithms could be used. For the sake of computational efficiency, also Online SLAM would be applicable. Mapping entire parts of the IWT-related infrastructure, the most feasible map approach appears to be a 3D volumetric map. Considering inland vessels as information platforms, passive SLAM is the exploration strategy of choice.

IV. HD MAPS FOR IWT APPLICATIONS

For traditional survey methods, a qualified team and high-precision equipment is required. Using inland vessels as sensor platforms is a comparably cost-efficient solution for the mapping problem. Furthermore, it is easy to ensure undisturbed traffic circulation and flawless survey operation at the same time.

The sensors presented in this work would be required already to ensure the safe passage of autonomous vessels. One of the motivations for the project of AutonomousSOW [1] is therefore to profit from synergy effects using the sensors required by autonomous vessels for navigation and perception tasks for generating high-definition maps. These high-definition maps are a further requirement for the safe passage of possible autonomous vessels. Particularly, driverless systems require more detailed information than lately is provided to captains of inland vessels.

Currently, horizontal and vertical bridge clearances are available online from the EuRIS portal [38] and can be downloaded in a JSON format. Bridges are assigned with their name and an additional identifier. Their position is encoded in the kilometre of the inland waterway where the bridge is located. Additionally, facility data is provided as well as the information whether the bridge is permanently installed or can be operated for the passage of larger ships.

Straight-forward intermediate steps could focus on updating the data available on EuRIS as the data format is already defined and available on an appropriate platform. For providing whole bridge contours, the data format of the existing platform could be adapted in an appropriate way. Still, for testing and validation purposes, an additional platform might be necessary with a data format adapted as closely as possible to the JSON representation used by EuRIS. For this purpose, the most feasible representation of bridge contours needs to be decided. It is desirable not to broadcast dense (and therefore memory-intensive) 3D LiDAR point clouds. Instead non-linear 3D polynomials could be used to approximate the bridge contours and transfer contour information at comparable precision and lower memory consumption. Another possibility of pattern representation are convex polygons. Wrapping the bridge contours with these geometrical structures without precision losses is expected to be feasible. Still, broadcasting 3D polygons might come with the drawback of increased memory consumption when compared to the polynomial representation. 2D polygons in comparison are considered to be unfeasible for the representation of 3D structures. As intermediate steps, assistant systems could use up-to-date bridge contours to notify the captain in case bridges along the planned passage appear to be dangerous to navigation. Within the scope of this development step, the appropriate representation, availability of the platform and especially performance of mapping algorithms under real-world conditions could be tested and assessed extensively. This is strongly necessary, as various error sources need to be considered when the LiDAR-based range information is registered into a global map. Even though, the GNSS navigation method in place works at a high-precision standard, the position estimation can generally be deteriorated by multipath effects, interference or intentional attacks. As the mapping task can only be performed using the estimated position, accuracy losses on this parameter also directly influence the position accuracy of the detected objects. Robust and reliable global positioning is extremely important for the use-case described in this work as uncertainties on the global position directly introduce uncertainties of the generated map. Finally, LiDAR-based range determination comes with the additional, potential error source of possible interference with other laser sources or sunlight.

V. Outlook

With this work, we present our concept of a self-updating inland waterway chart in order to pave the way towards the ongoing digitalisation of water-based transportation. The process for estimating HD maps for IWT is based on treating the surveying or participating vessels as information platforms. Such platforms shall be equipped with PNT and perception units, with GNSS, LiDARs and cameras being the most prominent sensors to be used. Our current HW and SW prototype has been presented, including the estimation techniques used by our PNT and Perception units. With regards to geo-referencing, our multi-antenna PPP-RTK solution allows for nearly instantaneous cm- and sub-degree accuracy for position and attitude estimates, respectively. Our Perception Unit provides remarkable results in semantic segmentation. The camera-based segmentation has been achieved with comparably low-effort and can be integrated in the follow-up perception subsystems. Still, the activities regarding the construction of the BerlinIWT dataset will be enlarged for more robust segmentation. A number of SLAM algorithms have been reviewed with respect to the application of IWT-related mapping. Meaningful combinations and adaptions will allow to exploit the advantages of the most suitable algorithms while leveraging their drawbacks. Finally, the setup of an appropriate data interface has been discussed while considering the most efficient way of transferring spatial mapping information.

Future activities will focus on the expansion of the BerlinIWT dataset, extrinsic calibration of the visual sensors and the implementation of a suitable multi-sensor, multi-modal SLAM algorithm and the setup of the data interface.

Acknowledgment

This work has been partially funded by the German Federal Ministry for Digital and Transport (BMDV) project AutonomSOW II and DLR Programmdirektion Verkehr with the project FuturePorts.

References

[1] A. GmbH. (2022) Projektsteckbrief autonomsow ii. [Online]. Available: https://www.autonomsow.de/autonomsow2.html

[2] B. Riveiro, P. Arias, J. Armesto, J. Caamaño, and M. Solla, "From geometry to diagnosis: Experiences of geomatics in structural engineering," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*, vol. 39, pp. 291–296, 2012.

[3] D. W. und Schifffahrtsverwaltung des Bundes. (2022) Bundeswasserstraßenkarten. [Online]. Available: https://www.gdws.wsv.bund.de/DE/service/karten/01_karten/karten-node.html

[4] C. C. Goad, "The ohio state university highway mapping project: the positioning component," in *Proceedings of the 47th Annual Meeting of The Institute of Navigation (1991)*, 1991, pp. 117–120.

[5] K. Novak, "The Ohio State University Highway Mapping System: The Stereo Vision System Component," in *Proceedings of the 47th Annual Meeting of The Institute of Navigation (1991)*, 1991, pp. 121–124.

[6] A. Geiger, J. Ziegler, and C. Stiller, "StereoScan: Dense 3d Reconstruction in Real-time," in *2011 IEEE intelligent vehicles symposium (IV)*. Ieee, 2011, pp. 963–968.

[7] P. Neis, D. Zielstra, and A. Zipf, "The street network evolution of crowdsourced maps: Openstreetmap in germany 2007–2011," *Future Internet*, vol. 4, no. 1, pp. 1–21, 2011.

[8] P. Ruchti, B. Steder, M. Ruhnke, and W. Burgard, "Localization on openstreetmap data using a 3d laser scanner," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 5260–5265.

[9] F. Poggenhans, J.-H. Pauls, J. Janosovits, S. Orf, M. Naumann, F. Kuhnt, and M. Mayr, "Lanelet2: A high-definition map framework for the future of automated driving," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 1672–1679.

[10] J.-H. Pauls, K. Petek, F. Poggenhans, and C. Stiller, "Monocular localization in hd maps by combining semantic segmentation and distance transform," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 4595–4601.

[11] C. Stachniss, J. J. Leonard, and S. Thrun, *Simultaneous Localization and Mapping*. Cham: Springer International Publishing, 2016, pp. 1153–1176. [Online]. Available: https://doi.org/10.1007/978-3-319-32552-1_46

[12] A. Eliazar and R. Parr, "Dp-slam: Fast, robust simultaneous localization and mapping without predetermined landmarks," in *IJCAI*, vol. 3. Acapulco, Mexico, 2003, pp. 1135–1142.

[13] M. Montemerlo, S. Thrun, D. Koller, B. Wegbreit *et al.*, "Fastslam 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges," in *IJCAI*, vol. 3, no. 2003, 2003, pp. 1151–1156.

[14] G. Grisetti, R. Kümmerle, C. Stachniss, and W. Burgard, "A tutorial on graph-based slam," *IEEE Intelligent Transportation Systems Magazine*, vol. 2, no. 4, pp. 31–43, 2010.

[15] S. Li and D. Lee, "Rgb-d slam in dynamic environments using static point weighting," *IEEE Robotics and Automation Letters*, vol. 2, no. 4, pp. 2263–2270, 2017.

[16] Y. Fan, Q. Zhang, Y. Tang, S. Liu, and H. Han, "Blitz-slam: A semantic slam in dynamic environments," *Pattern Recognition*, vol. 121, p. 108225, 2022.

[17] M. Henein, J. Zhang, R. Mahony, and V. Ila, "Dynamic slam: The need for speed," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 2123–2129.

[18] J. A. Placed, J. Strader, H. Carrillo, N. Atanasov, V. Indelman, L. Carlone, and J. A. Castellanos, "A survey on active simultaneous localization and mapping: State of the art and new frontiers," *IEEE Transactions on Robotics*, pp. 1–20, 2023.

[19] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with rao-blackwellized particle filters," *IEEE transactions on Robotics*, vol. 23, no. 1, pp. 34–46, 2007.

[20] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 4758–4765.

[21] M. Labbé and F. Michaud, "Rtab-map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation," *Journal of Field Robotics*, vol. 36, no. 2, pp. 416–446, 2019.

[22] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2020, pp. 5135–5142.

[23] J. Zumberge, M. Heflin, D. Jefferson, M. Watkins, and F. Webb, "Precise Point Positioning for the efficient and robust analysis of GPS data from large networks," *Journal of geophysical research: solid earth*, vol. 102, no. B3, pp. 5005–5017, 1997.

[24] R. B. Langley, "RTK GPS," *GPS World*, vol. 9, no. 9, pp. 70–76, 1998.

[25] P. J. Teunissen and O. Montenbruck, *Springer handbook of global navigation satellite systems*. Springer, 2017, vol. 10.

[26] P. Teunissen and A. Khodabandeh, "Review and principles of PPP-RTK methods," *Journal of Geodesy*, vol. 89, no. 3, pp. 217–240, 2015.

[27] X. An, R. Ziebold, and C. Lass, "From RTK to PPP-RTK: towards real-time kinematic precise point positioning to support autonomous driving of inland waterway vessels," *GPS Solutions*, vol. 27, no. 2, p. 86, 2023.

[28] D. Medina, J. Vilà-Valls, A. Hesselbarth, R. Ziebold, and J. García, "On the recursive joint position and attitude determination in multi-antenna GNSS platforms," *Remote Sensing*, vol. 12, no. 12, p. 1955, 2020.

[29] P. J. Teunissen, "A-PPP: Array-aided precise point positioning with global navigation satellite systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 2870–2881, 2012.

[30] Y. Li and J. Ibanez-Guzman, "Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems," *IEEE Signal Processing Magazine*, vol. 37, no. 4, pp. 50–61, 2020.

[31] X. Liu, Z. Deng, and Y. Yang, "Recent progress in semantic image segmentation," *Artificial Intelligence Review*, vol. 52, pp. 1089–1106, 2019.

[32] J. Taipalmaa, N. Passalis, and J. Raitoharju, "Different color spaces in deep learning-based water segmentation for autonomous marine operations," in *2020 IEEE international conference on image processing (ICIP)*. IEEE, 2020, pp. 3169–3173.

[33] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.

[34] J. Taipalmaa, N. Passalis, H. Zhang, M. Gabbouj, and J. Raitoharju, "High-resolution water segmentation for autonomous unmanned surface vehicles: A novel dataset and evaluation," in *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE, 2019, pp. 1–6.

[35] Y. Cheng, M. Jiang, J. Zhu, and Y. Liu, "Are we ready for unmanned surface vehicles in inland waterways? the usvinland multisensor dataset and benchmark," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3964–3970, 2021.

[36] B. Bovcon, J. Muhovič, J. Perš, and M. Kristan, "The mastr1325 dataset for training deep usv obstacle detection models," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 3431–3438.

[37] M. A. Rahman and Y. Wang, "Optimizing intersection-over-union in deep neural networks for image segmentation," in *International symposium on visual computing*. Springer, 2016, pp. 234–244.

[38] A. NV. (2023) Euris. [Online]. Available: https://www.eurisportal.eu/?KL=en