IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 14 August 2023, accepted 25 August 2023, date of publication 5 September 2023, date of current version 12 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3312186

METHODS

Synthetic Scan Formation for Underwater Mapping With Low-Cost Mechanical Scanning Sonars (MSS)

TIM HANSEN, (Member, IEEE), AND ANDREAS BIRK, (Member, IEEE)
Robotics Group, Constructor University, 28759 Bremen, Germany

Corresponding author: Tim Hansen (thansen@constructor.university)

This work was supported in part by the Deutsche Forschung Gemeinschaft (DFG) in the Project Unconstrained Synthetic Aperture Sonar (U-SAS) and the Project 3D Mapping of the Memorial Submarine Bunker Valentin by Air-, Ground-, and Underwater Robots (Valentin-3D) funded by the German Federal Ministry of Education and Research (BMBF).

ABSTRACT Mechanical Scanning Sonars (MSS) are popular devices for Unmanned Underwater Vehicles (UUV), i.e., Remotely Operated Vehicles (ROV) and Autonomous Underwater Vehicles (AUV), as they function under low visibility conditions and over extended ranges. They are comparatively low-cost and easy to integrate. But they require motion-compensation due to the low updates rates caused by the mechanical scanning and the slow speed of sound. We present here a new form of scan formation for MSS where the data from single beams is embedded into a pose-graph. The rendering of scans is not as usual based on only core navigation sensors, but it can improve in the spirit of a synthetic aperture. To this end, online Simultaneous Localization and Mapping (SLAM) is used to form scans from the single beams. These can be optimized and improved scans in turn lead to improved registration results in subsequent steps. This Synthetic Scan Formation (SSF) leads to better mapping results than state-of-the-art SLAM with MSS. The method is validated with several real-world experiments. First, different trajectories with precise ground-truth in a pool with a gantry set-up are used. Second, results from field trials in a WW-II submarine bunker are presented. It is shown that there are clear quantitative and qualitative improvements, and that SSF can be used in real-time for mapping during a mission.

INDEX TERMS Underwater mapping, marine robotics, sonar, synthetic aperture (SA), simultaneous localization and mapping (SLAM), registration, digital humanities, cultural heritage.

I. INTRODUCTION

Sonar [1] is an essential tool for underwater perception as it works in low visibility conditions and over extended ranges. But it also has its challenges due to its inherent limitations like the slow speed of sound, limited focus, or high noise levels [2], especially in comparison to light or RF-signals. The family of sonar sensors consists of multiple groups of device types that significantly differ in their functioning principles, complexity, and intended application scenarios [3].

The associate editor coordinating the review of this manuscript and approving it for publication was Xuebo Zhang.

One of the most simple types of sonar is the mechanical scanning sonar (MSS), also known as scanning sonar (SS) or mechanical scanning imaging sonar (MSIS). An MSS consists of a transducer to form a single beam and to measure the amplitudes of the returns along the time axis, i.e., the intensity and time-of-flight of each return. As the name suggests, the transducer and hence the beam is mechanically rotated, which allows to form a polar image by stepping through a sequence of angles. MSS have the advantage that they are low cost and they are easy to integrate due to their small size and low power-consumption. The significant disadvantage is the time per scan, namely in the range of multiple (tens of) seconds depending on factors like the range

and the field of view. The BlueRobotics Ping360 [4], e.g., needs 35 sec for a 360^o scan with 50 m range.

The polar images, respectively the ranges or the geometric features extracted from them, can be combined by registration and additional Simultaneous Localization and Mapping (SLAM) into 2D maps. But when an MSS is mounted on an Unmanned Underwater Vehicle (UUV), there is the need for motion compensation to avoid distorted scans. The state of the art is to generate the motion-compensated scans in fixed time-windows, e.g., whenever the sensor has completed a full rotation between minimum and maximum angles, using the vehicle's navigation [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18].

Note that underwater localization is very challenging as there is, among others, no access to Global Navigation Satellite Systems (GNSS) due to the strong attenuation of RF-signals in water, and that much less reliable alternatives than for land or aerial robots must be used [19], [20], [21]. Motion compensation based on navigation data is hence error-prone.

Here, we propose to generate a node in a pose-graph for each pose where a single scan-line along a beam is generated. The single scan-lines do not contain enough information to be used for registration; among others, there is hardly any overlap between them. Therefore, scan formation, i.e., the combination of multiple scan-lines into one image is still needed. But this can at least partially be based on SLAM in our approach, which leads to better localization estimates, and hence better input to the front-end, than just navigation alone. This Synthetic Scan Formation (SSF) hence produces scans that are not rigid, but which improve in their quality over time. The idea is in principle also applicable to other types of sonar like sidescan sonar. The source code of our method is freely available at https://github.com/constructor-robotics/scan_formation_underwater_mapping.

The rest of this article is structured as follows. Synthetic scan formation is formally introduced in Sec.II. A discussion including a presentation of its relation to Synthetic Aperture Sonar (SAS) is provided in Sec.III. Information on the methods used in the implementation is given in Sec.IV; this includes a presentation of the vehicle and its sensors that are used in all experiments and field trials. In Sec.V, the settings for the experiments are described. Results from a first set of real-world experiments are presented in Sec.VI, where a gantry set-up in a pool is used for validation with precise ground-truth data. In Sec.VII, results from two field trials are presented, which are conducted in different parts of a WW-II bunker for the production of submarines; they demonstrate the usefulness of synthetic scan formation in a concrete use-case, namely the digitization of cultural heritage. Sec.VIII concludes the article.

II. ADAPTIVE FORMATION OF SYNTHETIC SCANS

To formalize Synthetic Scan Formation (SSF), we use the following notations. A scan-line sl_t is a vector $I[i_{tof}]$ of the amplitudes of the returns of a single ping, i.e., the intensities measured by a single beam at time t. The index i_{tof} corresponds to the time-of-flight of each return, i.e., given the velocity of sound in water, it determines the distance d to the point where the back-scatter originated.

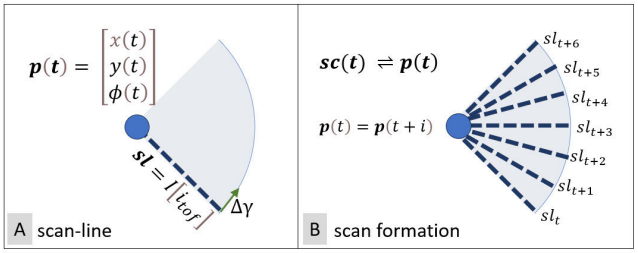


FIGURE 1. A Mechanical Scanning Sonar (MSS) has a beam that provides a 1D image or scan-line sl_t of backscatter intensities ordered by time of flight (tof), i.e., range [A]. The MSS mechanically rotates the beam, which leads to a scan $sc(t)$ in form of a polar image if the sensor pose $p(t+i)$, here $0 \leq i < 7$, is fixed during that time [B].

The sonar may have a minimum sensing distance d_{min}. The maximum range d_{max}, which depends among others on the power of the ping, is usually a parameter that can be selected. A MSS mechanically rotates the beam, i.e., the scan-line $sl(\gamma)$ from a minimum angle γ_{min} to a maximum angle γ_{max} with a step-width $\Delta\gamma$. The three parameters γ_{min}, γ_{max}, $\Delta\gamma$ can typically be set by the user.

A scan $sc(t)$ is simply a collection of k consecutive scan-lines sl_{t+i} recorded at time $t+i$ with $0 \leq i < k$. Given a sonar at a fixed location, a scan $sc(t)$ forms a polar image $I[\gamma][d]$ of the sonar's surroundings.

Fig.1 illustrates this with a very simple example with $\gamma_{min} = -45^o$, $\gamma_{max} = +45^o$, and $\Delta\gamma = 15^o$, i.e., the sensor has 7 scan-lines sl_t in its field of view through which it is mechanically stepping at the time-steps t to $t+6$. The symbol $\rightleftharpoons$ is used here to denote an association between two entities. For example, $sc(t) \rightleftharpoons p(t)$ symbolizes the pose $p(t)$ associated with the scan $sc(t)$.

Motion-compensation has to be used to avoid distortions on a moving platform (Fig.2). To form a scan, the 2D pose $p_s(t) = [x_s(t), y_s(t), \phi_s(t)]^T$ of the sensor at each time-step t needs to be taken into account, which depends in a fixed and

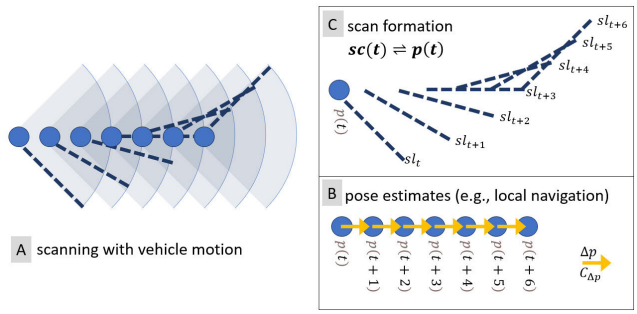


FIGURE 2. On a moving platform [A], the scan-lines sl_{t+i} can not just be combined into a polar image. But given pose-estimates $p(t+i)$ [B] for each sl_{t+i}, e.g., from navigation (Alg.1), they can be projected (Alg.2) to form a scan $sc(t)$ [C].

Algorithm 1 local Navigation

input : DVL, INS, ...(core navigation sensors)
const. : size k of the time window
output: pose-estimates $p(t), p(t+1), \ldots, p(t+k)$ with related uncertainties $C(t), C(t+1), \ldots, C(t+k)$

any navigation filter; here EKF

Algorithm 2 scan Formation

input : scan-lines $sl(t), sl(t+1), \ldots, sl(t+k)$ and their pose-estimates $p(t), p(t+1), \ldots, p(t+k)$
const. : size k of the time window; maximum index i_{max} of a scan-line; range difference Δd between two consecutive indexes i and $i+1$ in a scan-line
output: scan $sc(t)$ and its pose-estimate $p(t)$

$[x_o, y_o, \phi_o]^T = p(t)$
for $j \in \{0, \ldots, k\}$ **do**
  $[x, y, \phi]^T = p(t+j)$
  $I'_{1D}[.] = sl(t+j)$
  for $i \in \{0, \ldots, i_{max}\}$ **do**
    $x_{pixel} = x_o + \cos(\phi_o + \phi) \cdot (i \cdot \Delta d)$
    $y_{pixel} = y_o + \sin(\phi_o + \phi) \cdot (i \cdot \Delta d)$
    $I_{2D}[x_{pixel}][y_{pixel}] = I'_{1D}[i]$
$sc(t) = I_{2D}$
return $sc(t) \rightleftharpoons p(t)$

known way on the 2D vehicle pose $p(t) = [x(t), y(t), \phi(t)]^T$. For the sake of simplicity, we just refer to the vehicle pose $p(t)$ in the following.

Algorithm 3 scan Registration

input : two scans $sc_1()$ and $sc_2()$
const. : -
output: spatial transformation Δp between the scans, i.e., $p_2(t) = p_1(t) \oplus \Delta p$, with its uncertainty $C_{\Delta p}()$

any 2D registration algorithm; here FMS

As discussed in the introduction, the state-of-the-art is to use the pose-estimates $p(t+i)$ from the core navigation with

Algorithm 4 SLAM Back-End

input : pose-graph G
const. : -
output: optimized pose-graph G'

any graph SLAM optimizer; here iSAM

dead-reckoning (Alg.1) to form a scan $sc(t)$ (Alg.2). More precisely, a local trajectory estimation from navigation is used, i.e., a short sequence of estimated rigid transformations $\Delta p(t+i) = [\Delta x(t+i), \Delta y(t+i), \Delta\phi(t+i)]^T$ with $p(t+i+1) = p(t+i) \oplus \Delta p(t+i)$.

This is a reasonable approach as local navigation, i.e., motion estimation in rather short time-periods, tends to be reasonably accurate. The navigation filter also provides initial uncertainty estimates $C(t+i)$, which can be used to generate an uncertainty for the scan. So, the result is a rigid scan $sc(t)$ with an associated pose $p(t)$: $sc(t) \rightleftharpoons p(t)$.

For graph-based Simultaneous Localization and Mapping (SLAM) according to the state of the art, each scan $sc(t)$ and its pose $p(t)$ is associated with a vertex v_i, and an edge (v_i, v_{i+1}) representing the motion estimate between the (relative) poses is added (Fig.3, [A]). This motion estimate forms a constraint with an uncertainty that can be taken from the navigation filter. Registration of scan pairs (Alg.3) is essential to generate additional constraints, i.e., edges, including loop-closures. Scans are therefore formed every h time-steps with $h < k$ to ensure that there is some overlap between sequential scans. The pose-graph can then be optimized with any back-end (Alg.4). The different steps are summarized in Fig.4.

For SSF, we enter each scan-line sl_t into a node v_i together with its initial pose estimate $p(t)$ and the estimated motions between them as an edge (v_i, v_{i+1}) (Fig.3, [B]). At first glance, there is no benefit in this as the individual scan-lines do not contain enough information for registration and a scan formation still needs to be used. The major advantage is that the pose associated with each sl_i can get optimized by online SLAM, i.e., the quality of scans formed from them improves.

This is illustrated in more detail in Fig.5. The SSF proceeds in rounds n. The important aspect is that the scan $sc^{n_1}(t)$ in round n_1 can differ in appearance from the "same" scan $sc^{n_2}(t)$ in round n_2 as the pose-estimates of the scan-lines that form it can change.

SSF is initialized for $n = 0$ from local navigation (Fig.5, [A]), i.e., the initial pose- and uncertainty-estimates $p_i^{(0)}$, $C_i^{(0)}$ are provided from Alg.1 just like in the state of the art. Based on this, a scan $sc_i^{(0)}$ can be formed with Alg.2 for the first time. Overlapping scan pairs $sc_i^{(0)}$ and $sc_{i+h}^{(0)}$ can be registered (Fig.5, [B]), and the resulting graph is optimized (Fig.5, [C]).

But unlike the state of the art, SSF does not stop here. The optimized pose-estimates $p_i^{(1)}$ can be used for a new formation of scans (Fig.5, [A]), which are more focused due to improved localizations. This in turn leads to improved registration results (Fig.5, [B]), i.e., further improvements of the (relative) pose-estimates and their related uncertainties $p_i^{(1)}$, $C_i^{(1)}$, which again can be used for optimization (Fig.5, [C]).

This concept can be applied in an offline as well as in an online fashion. Offline, i.e., when processing recorded data after a mission, rounds can simply be iterated until the motion estimates from registration do not (substantially) change anymore. Online, i.e., during a mission, a new round can, e.g., be triggered when sufficiently many new

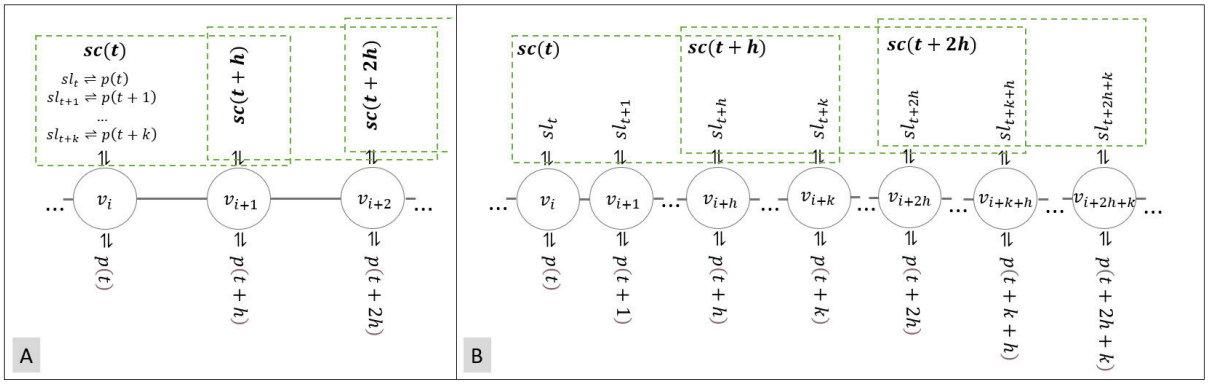


FIGURE 3. The state of the art is to use complete, rigid scans $sc(t)$ and their estimated pose $p(t)$ [A]. For synthetic scan formation, a single node v_i is linked to an individual scan-line sl_t and its estimated pose $p(t)$ [B], which can change over time using optimization, which allows in turn to form new, more focused scans (Fig.5).

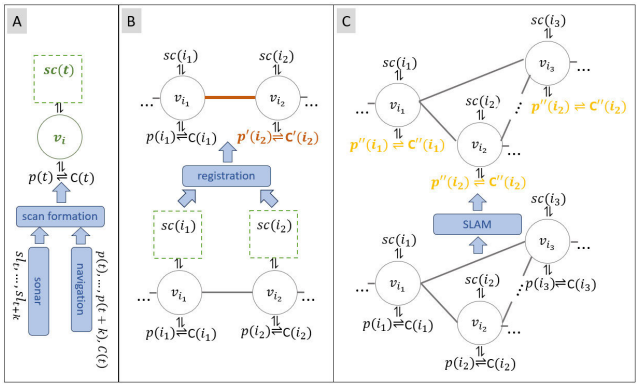


FIGURE 4. The state of the art uses rigid scans from motion compensation with local navigation [A], which get registered to generate additional constraints [B] and then optimized in a SLAM back-end [C].

scan-lines have been gathered to form an initial new scan with motion-compensation from navigation.

III. MOTIVATION AND DISCUSSION

The trade-off in SSF is an increased computational effort. Before a scan can be used for registration, it needs to be newly rendered with scan-formation when one of the pose-estimates of its scan-lines has been optimized. But this is also the strong point of SSF. It operates so to say on a level below conventional SLAM by optimizing the data that serves as input for the approaches according to the state of the art [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18]. As demonstrated in the experiments and field trials, this leads to improved results and the increase in time is minor; the synthetic scan formation is even usable for real-time mapping during a mission.

At first glance, SSF bears some relation to non-rigid registration, respectively non-rigid SLAM. This has, up to our knowledge, not been investigated for MSS before. More importantly, our method is very light-weight. It does not require regularization nor assumptions about the smoothness or distribution of the underlying data like popular non-rigid methods [22], [23], [24], [25].

The inspiration for SSF is derived from Synthetic Aperture Sonar (SAS), which is a widely researched principle for generating sonar data with an improved spatial resolution. The core idea of SAS is to obtain a sequence of sonar measurements by moving the sensor, which are combined to achieve enhanced representations of the sampled environment [26], [27], [28], [29]. This principle is derived from Synthetic Aperture Radar (SAR), which has been a standard remote sensing technique since multiple decades [30]. But the core concept has also been extended to completely different sensors and vehicles like thermal vision on aerial drones [31], [32], [33].

SAS is a very active field of research, see, e.g., [34], [35], [36], [37], [38], [39], [40], [41], [42], and [43]. But while the inspiration comes from SAS, there are also some substantial differences to it in SSF. First and foremost, SAS methods operate on the signal processing level of the raw data from the transducer(s). SSF operates so to say one level above, i.e., it uses the data from scan-lines as 1D images from the already processed transducer data. The synthetic aperture, respectively the improved focus of a scan in SSF is derived from the optimization of the estimated poses of the scan-lines. Second, SAS tends to require constraints on the sensor trajectory, typically in form of straight lines, to facilitate the underlying signal processing [26], [27], [28], [29]. This is not the case for SSF where the vehicle, or more precisely the sonar beam, can follow any arbitrary trajectory while the scan is formed. Third, SSF does not require access to the low signals to/from the transducer as it operates one level above it. This is a very important practical aspect as this data is usually not available in commercial off-the-shelf (COTS) sonars where this access is proprietary to the device manufacturer.

IV. IMPLEMENTATION: METHODS, VEHICLE, AND SENSORS

Synthetic scan formation as introduced in Sec.II is agnostic to the registration method and the SLAM back-end used in it.

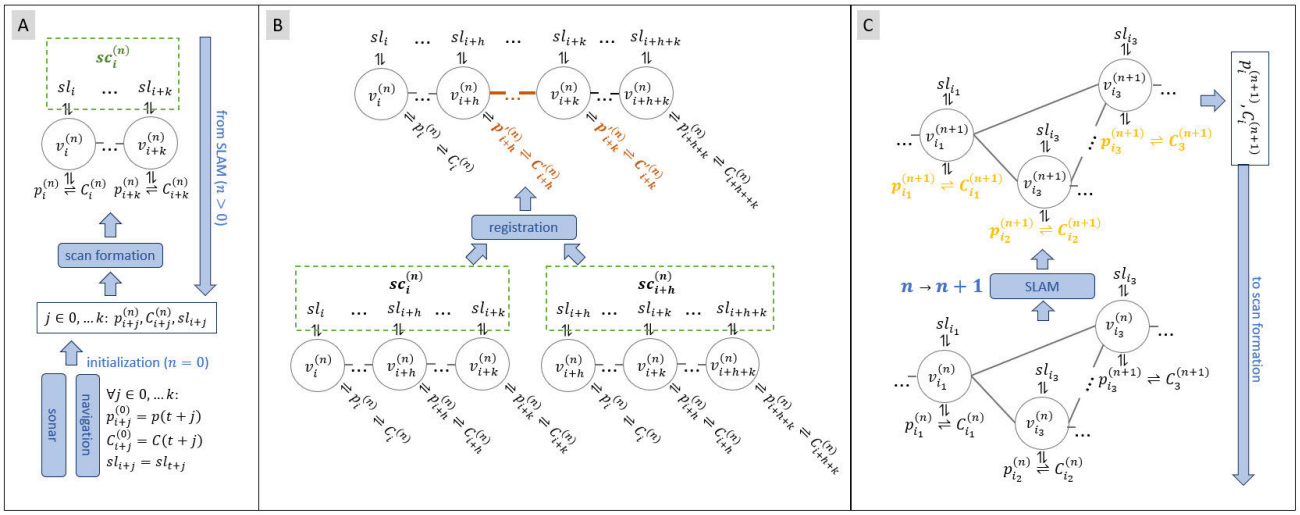


FIGURE 5. Synthetic scan formation proceeds in rounds n. Scan formation [A] is initialized with motion compensation from navigation ($n = 0$), but it uses an optimized pose estimates $p^{(n)}$ for each scan-line $sl^{(n)}$ when a scan is formed again ($n > 0$). Within a round, each registration of two scans sc_i^n and sc_{i+j}^n introduces $j + k$ new constraints [B]. After optimization with SLAM [C], there are new pose-estimates for the scan-lines, which can be used for scan formation in the next round ($n \rightarrow n + 1$).

And it is in principle also applicable to other types of sonar like sidescan sonar. In this section, the specific methods are described that are used to implement it for the experiments and results presented later on.

The pose-graph is embedded into a factor graph [44] optimized with iSAM2 [45] from the GTsam library [46]. Loop-closures are simply proximity based. In the front-end, Fourier-SOFT in 2D (FS2D) is used for registration. It is a spectral registration method that is well suited for low-quality, noisy sonar scans [47]. For all experiments and fields trials in this article, a grid-size of $N = 128$ is used in FS2D.

The registration needs to be accompanied by an uncertainty estimation. FS2D operates in the frequency domain where the different signal processing methods used in there generate in theory Dirac pulses for each degree of freedom (dof). In reality with imperfect data, a peak detection is needed to determine the registration results. Here, the 0-th dimensional persistent homology for 2D images [48] is used to this end. In the spirit of [49], a Principal Components Analysis (PCA) in a circle around each peak is used to determine a co-variance matrix.

The core navigation is based on the standard state-of-the-art approach for underwater dead-reckoning [19], [20], [21]. An Extended Kalman Filter (EKF) estimates the current pose of the robot $\boldsymbol{x}_{ekf} = [x, y, z, \psi, \theta, \phi]^\top$ using the input from a Doppler Velocity Log (DVL) and an Inertial Measurement Unit (IMU).

As mentioned before, MSS are low-cost sonars, which allow a wide variety of use-cases. The same holds for the rest of the hardware used in the experiments and field trials presented later on.

The vehicle is a BlueROV2 from BlueRobotics in the heavy configuration with a payload skid [50] (Fig.7b). The MSS is a Ping360 sonar from the same company [4], which is even within this class of devices one of the most affordable, if not the least costly device. As DVL, a Waterlinked A50 is used [51]. The IMU is a Xsens MTi-300 [52].

In the standard remote operation mode, the vehicle is connected with a tether to a control station in form of a laptop or embedded PC, to which the data is transmitted for processing. Our group has also developed multiple compute bottles based on different Intel NUCs for PC-like computation power on the vehicle itself, which can be used for on-board processing or autonomy if needed.

V. SETTINGS FOR THE EXPERIMENTS

Throughout all experiments and trials, we use the same settings for the three methods that are used for a comparison, namely mapping with dead-reckoning, state-of-the-art SLAM, and synthetic scan formation.

For the Ping360 sonar, the full 360^o field of view is used, i.e., $\gamma_{min} = 0^o$ and $\gamma_{max} = 360^o$. As described below, the three different test environments differ in size. Therefore, different settings for the maximum range d_{max} and for the stepping angle $\Delta\gamma$ are used. Both parameters not only influence the content of a scan, but also the time it takes to acquire the raw data. Therefore, they differ depending on the environment they are used in. The according values are shown in Tab.1.

The state-of-the-art approach to SLAM with MSS data can be formulated in our framework in a straight-forward way without any performance disadvantages. As described in Sec.II, each scan is formed once from k scan-lines using motion-compensation with navigation after the acquisition.

The core aspect of SSF is that synthetic scans can be re-generated from already recorded scan-lines if there is a

TABLE 1. Sonar parameters for the three test environments, i.e., (A) pool, (B) bunker basement, (C) large submarine basin. The number k of scan-lines per scan follows from $\Delta\gamma$; it is provided for the sake of simplicity.

	(A)	(B)	(C)
max range d_{max} (m)	3	20	30
step-width $\Delta\gamma$ (deg)	0.9	0.9	4.5
#scan-lines k per scan	400	400	80

TABLE 2. Accuracy of the core navigation with dead reckoning (dr), the state-of-the-art SLAM (sota), and Synthetic Scan Formation (SSF) for the different trajectories in the pool. The errors are computed via the Euclidean distances of the estimated locations to the accurate ground for each pose.

trajectory	error (mean ±std) as L2-norm (m) dr	sota	SSF
circles	0.2899 ± 0.1986	0.0300 ± 0.0162	0.0304 ± 0.0175
S-curves	0.4030 ± 0.2782	0.0404 ± 0.0211	0.0357 ± 0.0176
rectangles	0.3836 ± 0.2155	0.1449 ± 0.1002	0.0959 ± 0.0698

better pose-estimation for one or multiple scan-lines. It hence benefits from loop-closures, which tend to improve the poses and hence the quality of the synthetic scans.

In addition to the so to say global, proximity-based loop-closures, we attempt here also local loop-closures every $1/4$ of a scan, i.e., every 90^o. Note that this is a rather simple strategy to facilitate SSF as introduced in Sec.II. It nonetheless already leads to improved maps and small computational overhead as presented in detail in the following sections.

VI. POOL EXPERIMENTS WITH GROUND TRUTH

Ground truth data is a major challenge for underwater robotics in general, especially when it comes to localization as, e.g., GNSS are not accessible. Here, we use real-world experiments in a pool with a gantry mechanism where our vehicle's sensors are mounted on. The gantry serves as the propulsion for our ''vehicle'' with its sensors, i.e., the MSS, DVL, and IMU. The Computerized Numerical Control (CNC) of the gantry provides accurate motions and highly precise localization, thus allowing a quantitative analysis. The down-side is that the tank is only 2 m wide and 4 m long.

We use three different motion patterns in the experiments, namely a circle, a combination of S-curves, and a rectangular pattern. They form increased levels of difficulty, especially the rectangular patter with its abrupt changes at corners. The results for core navigation with dead reckoning, state-of-the-art SLAM, and SLAM with our scan formation are shown in Fig.6 and Tab.2.

The maps generated with core navigation, i.e., an EKF with DVL and IMU data, are always of low quality (Fig. 6a, 6b & 6g). This is also reflected in the errors, i.e., the Euclidean distances of the estimated locations to the ground truth (Tab.2).

SLAM with registration of the sonar data always substantially improves the results. In the case of the very simple trajectory of a circle, both the state-of-the-art SLAM and SSF perform similarly (Fig. 6b & 6c) with high accuracy (Tab.2). When the trajectories get a bit more complex, i.e., the S-curves and the rectangles, a higher accuracy of SSF can be clearly measured (Tab.2). For the maps, respectively the estimated trajectories, only small performance gains in minor details are visible due to the small size of the tank. This is very different for much larger environments in the fields trials presented in the next section.

TABLE 3. The computation times for processing one scan with dead-reckoning (dr), state-of-the-art SLAM (sota), and Synthetic Scan Formation (SSF), and the times for the raw sonar-data acquisition for a scan in (A) the pool, (B) the bunker basement, and (C) the large submarine basin.

	method (1 step)			sonar (1 scan)		
	dr	sota	SSF	(A)	(B)	(C)
time (sec)	0.18	0.4	0.6	20	32	10

VII. FIELD TRIALS

The field trials presented in this section are related to the use-case of the digitization of cultural heritage, for which robotic mapping technologies are meanwhile widely used [53], [54], though there are still many challenges when used underwater [55]. Concretely, we are interested in the digitization of the World-War-II submarine bunker Valentin (Fig.7a) with air-, ground-, and underwater robots [56] including 3D models from photogrammetry and Lidar data [57].

The bunker, which is now a memorial, was built from 1943 to 1945 with massive use of forced laborers with the intention to produce Type-XXI submarines there. While being the largest armament project of the German navy, the bunker could not be completed before the end of World War II. Up to 8,000 forced laborers worked on the bunker construction site every day, and many of them lost their lives [56].

Here, we present results from underwater mapping of two parts of the bunker. The first is the flooded basement of the bunker (Fig.7b&7c), for which open questions with respect to its use exist [58]. The second is a large basin from where the finished submarines could leave the bunker into the Weser river to move on to the sea (Fig.7d). Maps generated with the different methods are shown in Fig.'s 8 & 10.

Like in the pool-experiments, the core-navigation has significant drifts and it is not use-able for mapping in both cases (Fig.'s 8a & 10a), which demonstrates the need for registration of the sonar scans. More importantly, synthetic scan formation leads to clearly visible, qualitative improvements compared to the state-of-the-art SLAM.

For the basement, SSF (Fig.8c) captures more details than state-of-the-art SLAM (Fig.8b), e.g., an internal wall structure (Fig.9). It also produces proper representations of two connections to other parts of the bunker at both ends of the explored environment (Fig.9). Both aspects are very

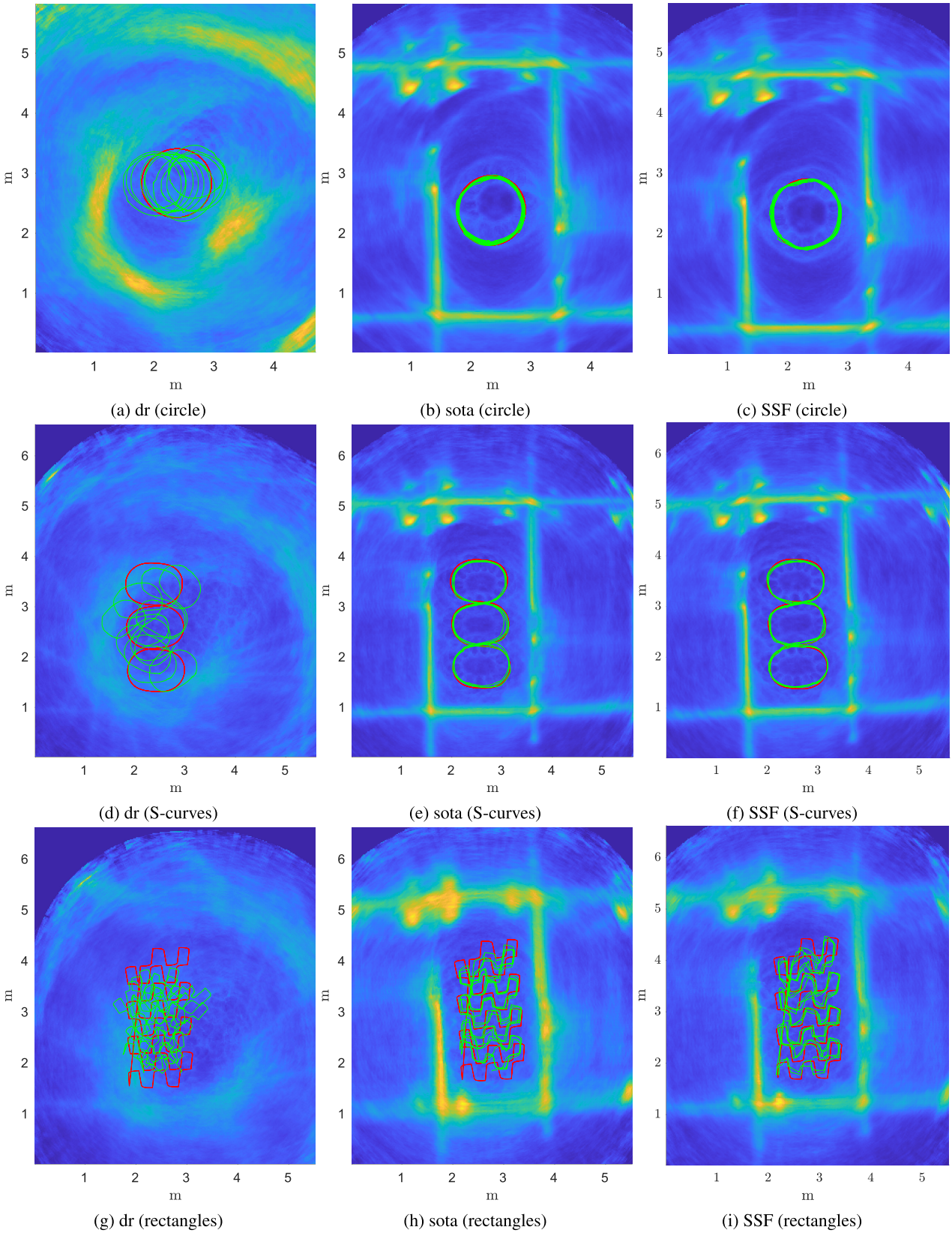
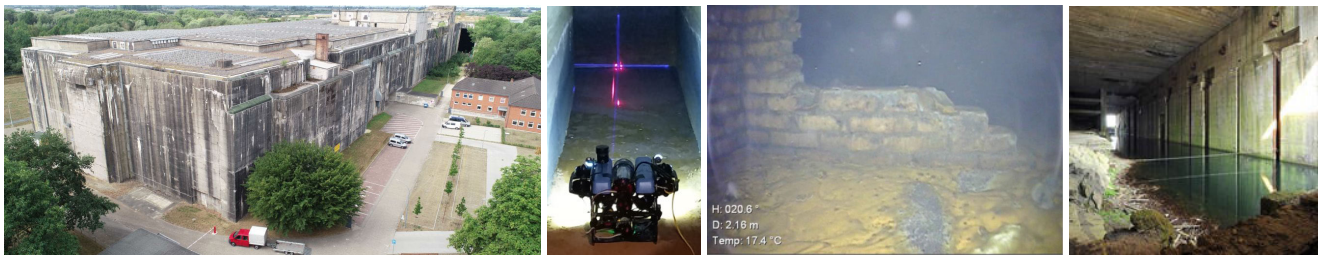
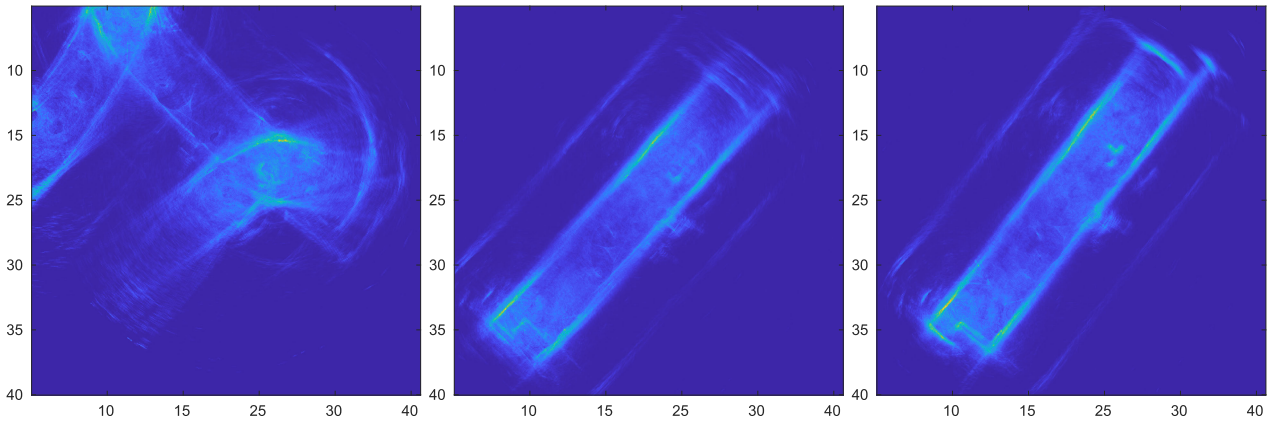


FIGURE 6. Maps generated with core navigation with dead reckoning (dr), state-of-the-art SLAM (sota), and Synthetic Scan Formation (SSF). The ground truth is shown in red for each of the three trajectories in form of a circle, a combination of multiple S-curves, and an exploration pattern consisting of rectangles. The estimated trajectories are shown in green for each method.



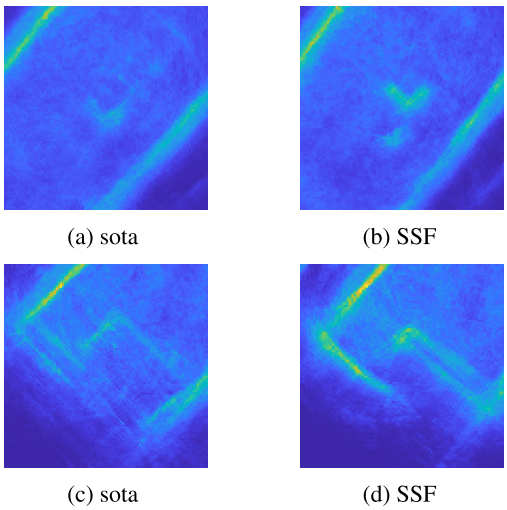
(a) (b) (c) (d)

FIGURE 7. (a) An aerial view on the WW-II bunker Valentin where the field trials are conducted as part of work on the digitization of cultural heritage. (b) The entrance to the flooded basement. (c) Part of a wall in the flooded basement. (d) A large basin that was intended to deploy the assembled submarines into the Weser river from where they could head to the sea.



(a) dr (b) sota (c) SSF

FIGURE 8. Three maps of the flooded basement (Fig.7b&7c), which are generated by (a) core navigation with dead reckoning (dr), (b) SLAM with state-of-the-art scan formation (sota), and (c) can Formation (SSF). Our method leads to straighter walls, clearly visible exit structures at both ends of the main room, and it captures smaller parts like the L-shaped, broken wall shown in Fig.7c.



(a) sota (b) SSF

(c) sota (d) SSF

FIGURE 9. Zoomed details from the maps of the basement with sota-SLAM and our method (Fig.8. Top: An L-shaped wall (Fig.7c) and a pillar-structure below it. Bottom: An exit leading into a very wide corridor. Note that the shorter concrete wall is seen from two sides, i.e., our method captures its thickness.

useful for the exploration of the basement as the visibility there is very poor and there are debris as well as partially destroyed constructions. Our method also leads to straighter and more pronounced representations of the main walls (Fig.8), which also indicates some benefits for the overall map on a global scale.

The improvement for the overall map is even more pronounced in the case of the large basin (Fig.7d). SSF compensates the drift in the core navigation very well (Fig.10c) and it leads to much straighter walls than the state-of-the-art SLAM (Fig.10b). Note that the straight, but broad area in the left lower part of the maps originates from a part of the basin that got destroyed and that hence consists of a kind of beach of rubble (Fig.7d).

Tab.3 shows the average computation times for processing one sonar scan with dead-reckoning (dr), state-of-the-art SLAM (sota), and Synthetic Scan formation (SSF) including loop-closures and re-rendering. Furthermore, the times it takes to record the data for one scan with the sonar are given, which depend on the maximum range setting of the sonar together with the sampling rate of the angle (Tab.1). The three time-values for the sonar correspond to the experiments in (A) the pool, (B) the basement, and (C) the large basin.

Our method adds about 50% computation overhead compared to the state-of-the-art SLAM while leading to

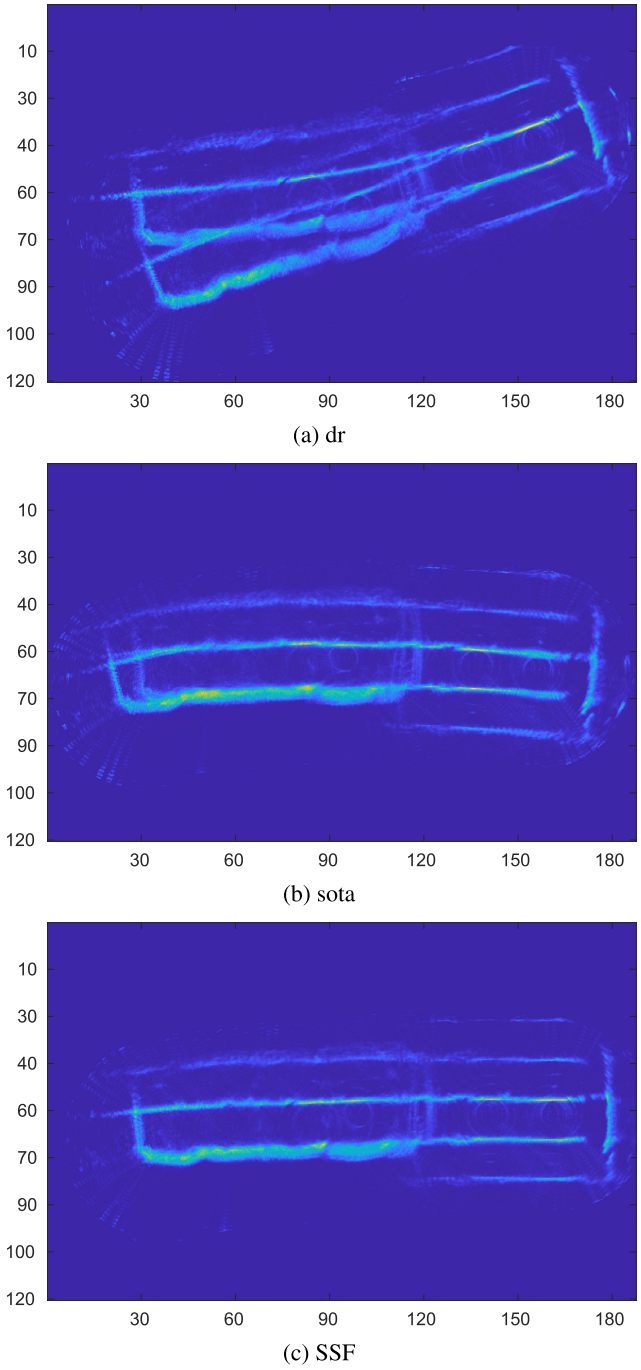


FIGURE 10. Three maps of the large, rectangular basin in the bunker (Fig.7d), which are generated again by (a) dead reckoning, (b) sota SLAM, and (c) Synthetic Scan Formation.

improved results as discussed above. More importantly, the processing of a scan is significantly shorter than its data-acquisition with the sonar. Hence, SSF is very well suited for real-time processing.

VIII. CONCLUSION

Synthetic Scan Formation (SSF) was introduced. It operates on a level above conventional Synthetic Aperture Sonar (SAS) and below state of the art Simultaneous Localization and Mapping (SLAM) with sonar, i.e., it does not require raw data from the transducer(s) for signal processing, but it improves the focus of scans that serve as input to the registration and the optimization backend.

The core idea is to use pose-estimates for each single beam or scan-line, i.e., the array of amplitude measurements from a single ping ordered by time-of-flight and hence distance, to form the synthetic scans. Instead of just employing the core navigation data to combine multiple scan-lines into each scan once, SSF uses online SLAM to update the pose-estimates of the scan-lines and new, improved scans can be (re-)rendered. Improved scans lead to improved registration results in the subsequent SLAM processing and they hence lead to an overall improved map quality.

Synthetic Scan Formation is validated in several real-world experiments. Quantitative results are presented from experiments in a pool with precise ground-truth through motions with a gantry set-up. Using different trajectories, it is shown that our methods increases the accuracy. Furthermore, qualitative results from field trials within a project for the digitization of cultural heritage are presented. Two different areas from the memorial bunker Valentin, a WW-II bunker for the production of submarines, are explored and mapped. SSF leads to visible improvements while it is still use-able for real-time processing during the mission.

ACKNOWLEDGMENT

The authors would like to thank the Institute of Mechanics and Ocean Engineering, Technical University of Hamburg, for giving them access to their research pool with the gantry.

REFERENCES

[1] A. W. Cox, *Sonar and Underwater Sound*. Toronto, ON, Canada: Univ of Toronto Press, 1974.

[2] J. P. Marage and Y. Mori, *Sonar and Underwater Acoustics*. Hoboken, NJ, USA: Wiley, 2013.

[3] R. P. Hodges, *Underwater Acoustics: Analysis, Design and Performance of Sonar*. Hoboken, NJ, USA: Wiley, 2010.

[4] BlueRobotics. (2023). *Ping360 Scanning Imaging Sonar*. [Online]. Available: https://bluerobotics.com/store/sensors-sonars-cameras/sonar/ping360-sonar-r1-rp/

[5] T. Hansen, F. Buda, and A. Birk, "Underwater exploration with sonar of the flooded basement of a WW-II submarine bunker in the context of digitization of cultural heritage," in *Proc. MTS/IEEE OCEANS*, Jun. 2023, pp. 1–5.

[6] Z. Xu, H. Qiu, M. Dong, H. Wang, and C. Wang, "Underwater simultaneous localization and mapping based on 2D-SLAM framework," in *Proc. IEEE Int. Conf. Unmanned Syst. (ICUS)*, Oct. 2022, pp. 184–189.

[7] L. Chen, A. Yang, H. Hu, and W. Naeem, "RBPF-MSIS: Toward Rao–Blackwellized particle filter SLAM for autonomous underwater vehicle with slow mechanical scanning imaging sonar," *IEEE Syst. J.*, vol. 14, no. 3, pp. 3301–3312, Sep. 2020.

[8] M. Jiang, S. Song, J. M. Herrmann, J.-H. Li, Y. Li, Z. Hu, Z. Li, J. Liu, S. Li, and X. Feng, "Underwater loop-closure detection for mechanical scanning imaging sonar by filtering the similarity matrix with probability hypothesis density filter," *IEEE Access*, vol. 7, pp. 166614–166628, 2019.

[9] M. Jiang, S. Song, F. Tang, Y. Li, J. Liu, and X. Feng, "Scan registration for underwater mechanical scanning imaging sonar using symmetrical Kullback–Leibler divergence," *Proc. SPIE*, vol. 28, no. 1, Feb. 2019, Art. no. 013026.

[10] A. Mallios, P. Ridao, D. Ribas, and E. Hernández, "Scan matching SLAM in underwater environments," *Auto. Robots*, vol. 36, no. 3, pp. 181–198, Mar. 2014, doi: 10.1007/s10514-013-9345-0.

[11] A. Burguera, Y. González, and G. Oliver, "Underwater SLAM with robocentric trajectory using a mechanically scanned imaging sonar," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2011, pp. 3577–3582.

[12] A. Mallios, P. Ridao, M. Carreras, and E. Hernández, "Navigating and mapping with the SPARUS AUV in a natural and unstructured underwater environment," in *Proc. OCEANS MTS/IEEE KONA*, Sep. 2011, pp. 1–7.

[13] H. Bülow, M. Pfingsthorn, and A. Birk, "Using robust spectral registration for scan matching of sonar range data," in *Proc. 7th Symp. Intell. Auton. Vehicles (IAV)*, 2010, pp. 611–616.

[14] A. Burguera, G. Oliver, and Y. Gonzàlez, "Scan-based SLAM with trajectory correction in underwater environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2010, pp. 2546–2551.

[15] A. Mallios, P. Ridao, E. Hernandez, D. Ribas, F. Maurelli, and Y. Petillot, "Pose-based SLAM with probabilistic scan matching algorithm using a mechanical scanned imaging sonar," in *Proc. OCEANS EUROPE*, May 2009, pp. 1–6, doi: 10.1109/OCEANSE.2009.5278219.

[16] E. Hernández, P. Ridao, D. Ribas, and A. Mallios, "Probabilistic sonar scan matching for an AUV," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2009, pp. 255–260, doi: 10.1109/IROS.2009.5354656.

[17] D. Ribas, P. Ridao, J. D. Tardós, and J. Neira, "Underwater SLAM in man-made structured environments," *J. Field Robot.*, vol. 25, nos. 11–12, pp. 898–921, Nov. 2008.

[18] D. Ribas, P. Ridao, J. Domingo Tardos, and J. Neira, "Underwater SLAM in a marina environment," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, P. Ridao, Ed., Oct. 2007, pp. 1455–1460.

[19] F. Maurelli, S. Krupiński, X. Xiang, and Y. Petillot, "AUV localisation: A review of passive and active techniques," *Int. J. Intell. Robot. Appl.*, vol. 6, no. 2, pp. 246–269, Jun. 2022, doi: 10.1007/s41315-021-00215-x.

[20] L. Paull, M. Seto, S. Saeedi, and J. J. Leonard, *Navigation for Underwater Vehicles*. Berlin, Germany: Springer, 2018, pp. 1–15, doi: 10.1007/978-3-642-41610-1_15-1.

[21] L. Paull, S. Saeedi, M. Seto, and H. Li, "AUV navigation and localization: A review," *IEEE J. Ocean. Eng.*, vol. 39, no. 1, pp. 131–149, Jan. 2014.

[22] S. Monji-Azad, J. Hesser, and N. Löw, "A review of non-rigid transformations and learning-based 3D point cloud registration methods," *ISPRS J. Photogramm. Remote Sens.*, vol. 196, pp. 58–72, Feb. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271622003380

[23] B. Jian and B. C. Vemuri, "Robust point set registration using Gaussian mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1633–1645, Aug. 2011.

[24] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–2275, Dec. 2010.

[25] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Comput. Vis. Image Understand.*, vol. 89, nos. 2–3, pp. 114–141, Feb. 2003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1077314203000092

[26] P. Vouras, K. V. Mishra, A. Artusio-Glimpse, S. Pinilla, A. Xenaki, D. W. Griffith, and K. Egiazarian, "An overview of advances in signal processing techniques for classical and quantum wideband synthetic apertures," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 2, pp. 317–369, Mar. 2023.

[27] M. Amani, F. Mohseni, N. F. Layegh, M. E. Nazari, F. Fatolazadeh, A. Salehi, S. A. Ahmadi, H. Ebrahimy, A. Ghorbanian, S. Jin, S. Mahdavi, and A. Moghimi, "Remote sensing systems for ocean: A review (Part 2: Active systems)," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 1421–1453, 2022.

[28] R. E. Hansen, *Introduction to Synthetic Aperture Sonar*. Rijeka, Croatia: InTech, 2011.

[29] M. P. Hayes and P. T. Gough, "Synthetic aperture sonar: A review of current status," *IEEE J. Ocean. Eng.*, vol. 34, no. 3, pp. 207–224, Jul. 2009.

[30] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 1, pp. 6–43, Mar. 2013.

[31] D. C. Schedl, I. Kurmi, and O. Bimber, "An autonomous drone for search and rescue in forests using airborne optical sectioning," *Sci. Robot.*, vol. 6, no. 55, Jun. 2021, Art. no. eabg1188. [Online]. Available: https://www.science.org/doi/abs/10.1126/scirobotics.abg1188

[32] A. Birk, "Seeing through the forest and the trees with drones," *Sci. Robot.*, vol. 6, no. 55, 2021, Art. no. eabj3947. [Online]. Available: https://robotics.sciencemag.org/content/6/55/eabj3947

[33] I. Kurmi, D. Schedl, and O. Bimber, "Airborne optical sectioning," *J. Imag.*, vol. 4, no. 8, p. 102, Aug. 2018. [Online]. Available: https://www.mdpi.com/2313-433X/4/8/102

[34] M. Ma, J. Tang, H. Wu, P. Zhang, and M. Ning, "CZT algorithm for the Doppler scale signal model of multireceiver SAS based on shear theorem," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5201412.

[35] Z. Tian, H. Zhong, J. Tang, and J. Zhang, "Azimuth-invariant motion compensation and imaging chirp scaling algorithm for multiple-receiver synthetic aperture sonar," *IEEE Access*, vol. 10, pp. 114060–114076, 2022.

[36] D. J. Pate, D. A. Cook, and B. N. O'Donnell, "Estimation of synthetic aperture resolution by measuring point scatterer responses," *IEEE J. Ocean. Eng.*, vol. 47, no. 2, pp. 457–471, Apr. 2022.

[37] S.-M. Steele and A. P. Lyons, "Development and experimental validation of endfire synthetic aperture sonar for sediment acoustics studies," *IEEE J. Ocean. Eng.*, vol. 47, no. 2, pp. 472–482, Apr. 2022.

[38] B. Thomas and A. Hunter, "Coherence-induced bias reduction in synthetic aperture sonar along-track micronavigation," *IEEE J. Ocean. Eng.*, vol. 47, no. 1, pp. 162–178, Jan. 2022.

[39] M. Ma, J. Tang, H. Zhong, and H. Wu, "Multireceiver synthetic aperture sonar chirp scaling algorithm considering intrapulse Doppler shift," *IEEE J. Ocean. Eng.*, vol. 47, no. 2, pp. 433–444, Apr. 2022.

[40] M. Wu and S. Yan, "Motion compensation for OFDM inverse synthetic aperture sonar imaging based on compressed sensing," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.

[41] X. Zhang, H. Wu, H. Sun, and W. Ying, "Multireceiver SAS imagery based on monostatic conversion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 10835–10853, 2021.

[42] S. A. V. Synnes, R. E. Hansen, and T. O. Sæbø, "Spatial coherence of speckle for repeat-pass synthetic aperture sonar micronavigation," *IEEE J. Ocean. Eng.*, vol. 46, no. 4, pp. 1330–1345, Oct. 2021.

[43] B. Thomas, A. Hunter, and S. Dugelay, "Phase wrap error correction by random sample consensus with application to synthetic aperture sonar micronavigation," *IEEE J. Ocean. Eng.*, vol. 46, no. 1, pp. 221–235, Jan. 2021.

[44] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA, USA: MIT Press, 2009.

[45] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping with fluid relinearization and incremental variable reordering," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2011, pp. 3281–3288.

[46] D. Frank and K. Michael, *Factor Graphs for Robot Perception*. Hanover, MD, USA: Now Publishers, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/8187520

[47] T. Hansen and A. Birk, "Using registration with Fourier-SOFT in 2D (FS2D) for robust scan matching of sonar range data," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2023, pp. 3080–3087.

[48] H. Edelsbrunner and J. Harer, *Computational Topology: An Introduction*. Providence, RI, USA: American Mathematical Society, 2010.

[49] M. Pfingsthorn, A. Birk, S. Schwertfeger, H. Bülow, and K. Pathak, "Maximum likelihood mapping with spectral image registration," in *Proc. IEEE Int. Conf. Robot. Autom.* Piscataway, NJ, USA: IEEE Press, May 2010, pp. 4282–4287.

[50] BlueRobotics. (2023). *BlueROV2*. [Online]. Available: https://bluerobotics.com/store/rov/bluerov2/

[51] WaterLinked. (2023). *A50 DVL*. [Online]. Available: https://waterlinked.com/dvl-a50

[52] Xsense. (2023). *MTi-300 AHRS*. [Online]. Available: https://www.movella.com/products/sensor-modules/xsens-mti-300-ahrs

[53] V. Poulopoulos and M. Wallace, "Digital technologies and the role of data in cultural heritage: The past, the present, and the future," *Big Data Cogn. Comput.*, vol. 6, no. 3, p. 73, Jul. 2022. [Online]. Available: https://www.mdpi.com/2504-2289/6/3/73

[54] F. Galeazzi, "Towards the definition of best 3D practices in archaeology: Assessing 3D documentation techniques for intra-site data recording," *J. Cultural Heritage*, vol. 17, pp. 159–169, Jan. 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S129620741500120X

[55] F. Menna, P. Agrafiotis, and A. Georgopoulos, "State of the art and applications in archaeological underwater 3D recording and mapping," *J. Cultural Heritage*, vol. 33, pp. 231–248, Sep. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1296207417308555
[56] A. Birk, F. Buda, H. Bülow, A. G. Chavez, C. A. Müller, and J. Timpe, *Digitizing a Gigantic Nazi Construction: 3D-Mapping of Bunker Valentin in Bremen*. Berlin, Germany: De Gruyter, 2022, pp. 133–168, doi: 10.1515/9783110714692-006.
[57] H. Bülow, C. A. Müller, A. G. Chavez, F. Buda, and A. Birk, "A divide and conquer method for 3D registration of inhomogeneous, partially overlapping scans with Fourier Mellin SOFT (FMS)," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 8594–8601.
[58] A. Birk, F. Buda, and T. Hansen, "Dating wall constructions from brick sizes in the flooded basement of a WW-II submarine bunker for the digitization of cultural heritage," in *Proc. MTS/IEEE OCEANS*, Jun. 2023, pp. 1–6.

TIM HANSEN (Member, IEEE) received the B.S. and M.S. degrees in mechatronics engineering from the Hamburg University of Technology, Germany, in 2020. He is currently pursuing the Ph.D. degree with Constructor University, Bremen, Germany. Before the M.S. degree, the focus lied in micro underwater robotics, with the focus on state estimation. His current research interest includes underwater robotic systems, that are able to observe the environment, while still being low-cost.

ANDREAS BIRK (Member, IEEE) received the B.Sc. (Diploma) and M.Sc. degrees in computer science and the Ph.D. (Dr. rer.nat) degree from Universität des Saarlandes, Saarbrücken, Germany, in 1989, 1993, and 1995, respectively. Since 2011, he has been a Full Professor in electrical engineering and computer science with Constructor University, Bremen, Germany, where he started as an Associate Professor, in 2001. From 1996 to 2001, he was a Postdoctoral Researcher with the AI-Laboratory, Vrije Universiteit Brussel (VUB), Brussels, Belgium, where he was among others as a Research Fellow of the Flemish Society of Applied Research (IWT), from 1998 to 2001.