2023 IEEE International Conference on Robotics and Automation (ICRA 2023)
May 29 - June 2, 2023. London, UK

Using Registration with Fourier-SOFT in 2D (FS2D) for Robust Scan Matching of Sonar Range Data

Tim Hansen and Andreas Birk

***Abstract*— In this paper, we introduce Fourier-SOFT 2D (FS2D) as a new robust registration method. FS2D operates in the frequency domain where it exploits the well-known decoupling of rotation and translation. The challenging part of determining the rotation parameter is solved here based on a projection of the Fourier magnitude on a sphere and the SO(3) Fourier Transform (SOFT). The underlying use case is underwater mapping with sonar, i.e., with very noisy and partially overlapping environment data under non-trivial localization and navigation challenges. Fourier-SOFT 2D is compared with openly available registration methods on two real-world datasets and a simulated dataset. Results show the robustness of FS2D, i.e., its capabilities to handle large amounts of noise and occlusions of consecutive scans. The implementation in C++ is openly available.**

I. Introduction

The mapping of underwater structures is particularly difficult. First of all, localization is challenging due to the lack of access to Global Navigation Satellite Systems (GNSS) and the limitations of odometry [1]. Second, the perception of the environment, which can, respectively must aid underwater localization, is challenging, too. Vision for example is limited to specific application scenarios with decent visibility conditions [2], [3]. Sonar in contrast works under low visibility conditions, but it also suffers from multiple limitations, e.g., the slow speed of sound, limited spatial resolution, and high amounts of noise [4]–[6].

Here, we deal with the registration, i.e., the spatial alignment, of 2D sonar scans from multi-beam echo sounders (MBES) or from scanning sonars with motion-compensation [4]. More precisely, 2D range sonar scans consist of a polar representation of amplitude information over time-of-flight of the ping returns, i.e., the range information is only implicitly encoded in the scan that is affected by various noise effects. The scans are hence somewhat a chimera of images and 2D range-clouds. Successful 2D registration recovers the change in poses $(\Delta x, \Delta y, \Delta\alpha)$ between the scans, which can then be used for navigation and mapping including also the option of Simultaneous Localization and Mapping (SLAM). Typical application scenarios are the mapping of natural or man-made environments like cliffs or harbors, whereas our interest in this task stems from a project on the digitization of cultural heritage (Fig.1).

Fig. 1: The WWII submarine-bunker Valentin (top) is mapped by air-, ground-, and underwater-robots as part of a project on the digitization of cultural heritage [7]. Especially the flooded parts of the bunker (center) - including maze-like basements - are of interest to historians as there are differences to the original construction plans, e.g., unexpected walls or structures (bottom).

2D registration is in general a well-investigated topic since several decades [8]–[10]. But it is quite challenging to use standard registration methods for sonar scans and images due to the noise and structural errors in the data [11]–[23]. We introduce here with *Fourier-SOFT in 2D (FS2D)* a new 2D registration method, which is particularly well suited for noisy data and hence sonar scans. The implementation is openly available[1].

The method is described in some detail in the next Section II. Section III uses an openly available 2D sonar scan dataset and a dataset recorded in the Bunker Valentin (Fig. 1) to compare FS2D with multiple standard registration methods. In addition, a simulated dataset with precisely known

Authors are with the School of Computer Science and Engineering at Constructor University, Bremen, Germany. <thansen, abirk>@constructor.university

The presented research was partially support 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) financed by the German Federal Ministry of Education and Research (BMBF).

[1] https://github.com/constructor-robotics/fourier-soft-2d.git

ground truth is used to benchmark consecutive scan matching. Section IV concludes the paper.

II. Fourier-SOFT in 2D (FS2D)

Fourier-SOFT in 2D (FS2D) is a spectral method, i.e., it operates in the frequency-domain. According methods have recently become more popular again due to their higher robustness compared to feature-based methods; see, e.g., [24] for a more detailed review and discussion.

FS2D exploits the well-known fact that translation and rotation - plus scale if of interest - are decoupled in the frequency domain [25], [26]. Suppose function $f(\mathbf{x})$ is rotated by function $g(\alpha,\beta,\gamma)$. The Fourier magnitude $|\tilde{F}(\mathbf{k})|$ of function $\tilde{f}(\mathbf{x}) = g(\alpha,\beta,\gamma)f(\mathbf{x})$ is

$$|\tilde{F}(\mathbf{k})| = |F(g(\alpha,\beta,\gamma)\mathbf{k})|, \tag{1}$$

where the capital letter denotes the function in the frequency domain.

As usual, the difficult part for a registration method is to determine the rotation component. Once it is known, the two degrees of freedom (dof) $\Delta x, \Delta y$ of the translation can be easily recovered by correlation as shown already in the early days of computer vision [27].

For the 2D rotation, i.e., the 1-dof $\Delta\alpha$, we propose here an approach based on the SO(3) Fourier transform (SOFT) [28]. This may at first glance look awkward as SOFT is - as the name suggests - linked to SO(3) transformations; but it leads to very robust results as shown later on. SOFT provides an efficient correlation on the $\mathbb{S}^2$ sphere using spherical harmonics, where the sphere can be represented by these harmonics, similar to the Fourier transform. This is also used in Fourier-Mellin-SOFT (FMS) [29], a very robust 3D registration method providing 7-dof, i.e., translation, rotation, and scale in 3D. Naively, one may assume that FMS could just be applied on 2D scans with Zero-padding in the 3rd dimension, but this would be much too sparse data and it would lead to computational overhead.

To be applicable, FS2D differs from FMS in the way the spectral data is projected.

More precisely, given a regular $N \times N$ array with intensity values $f(\mathbf{x})$ with $\mathbf{x} = [x, y]^T$. The corresponding 2D Discrete Fourier Transform (DFT) is denoted with $F(\mathbf{k})$ and the frequency coordinates $\mathbf{k} = [u, v]^T$. As sonar scans are not square but so-to-say fan-shaped in polar coordinates, Zero-padding is used to enter them in $f(\mathbf{x})$. Roughly speaking, each scan is rendered as a square image in the spatial domain and then converted to the frequency domain.

The core idea is to project the spectral magnitude $|F(\mathbf{k})|$ of the frequency domain representation of each scan $f(\mathbf{x})$ onto a sphere $\omega(\theta,\phi) \in \mathbb{S}^2$. For an equalization of the amplitudes of the spectral structures, the data on the sphere is processed with Contrast Limited Adaptive Histogram Equalization (CLAHE) [30].

The resampling process is done as follows:

$$f_{\mathbb{S}^2}(\theta,\phi) = \sum_{j=0}^{2B-1}\sum_{k=0}^{2B-1} |F(\mathbf{k})| \tag{2}$$

$$\mathbf{k} = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} r\sin\theta\cos\phi + N/2 \\ r\sin\theta\sin\phi + N/2 \end{bmatrix} \tag{3}$$

$$\phi = \frac{\pi\phi_k}{B}; \theta = \frac{\pi\theta_j}{2B} \tag{4}$$

$$\phi_k = 0, ..., 2B-1; \theta_j = 0, ..., 2B-1 \tag{5}$$

$$r = N/2 - 1 \tag{6}$$

The bandwidth B is $N/2$ and the resulting 2D function $f_{\mathbb{S}^2}(\theta,\phi)$ describes the unwrapped $\mathbb{S}^2$ sphere.

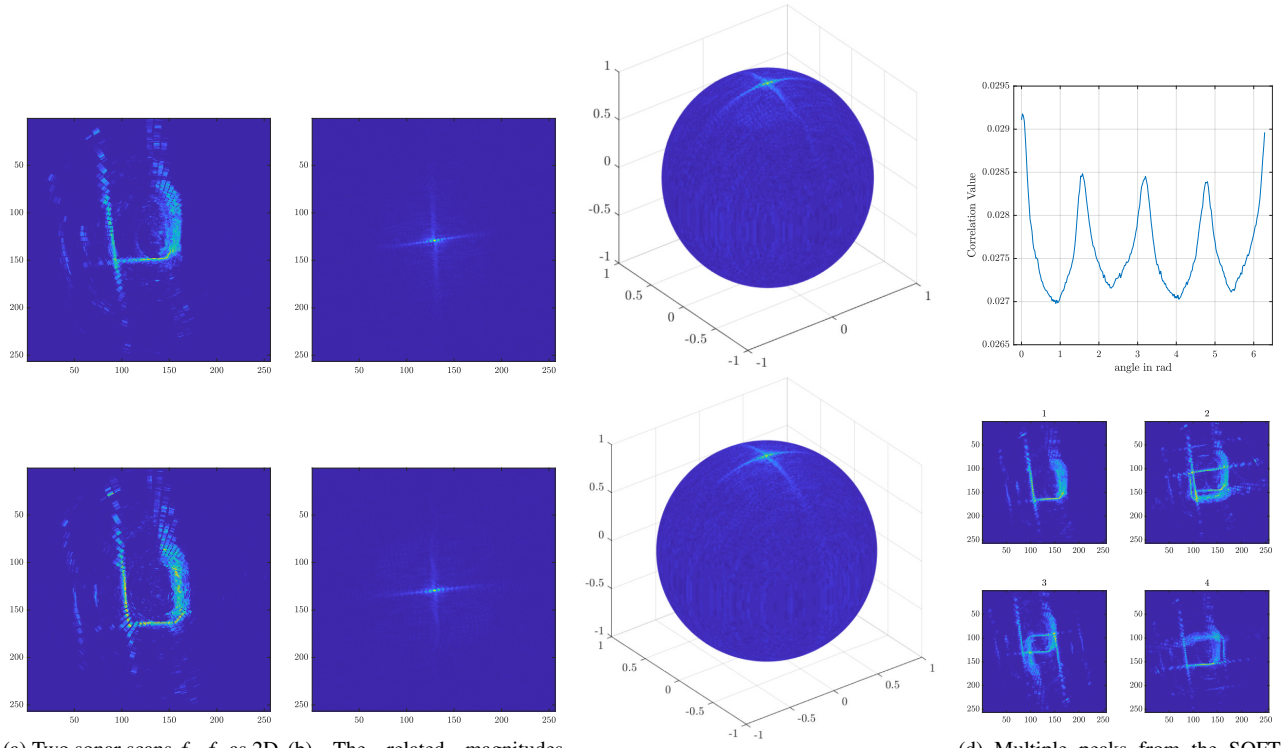
Given two 2D scans that are accordingly represented, SOFT can then be directly applied to determine $\Delta\alpha$. More precisely, the correlation between the two spheres leads to a pulse along the 1st dimension of the rotation $g(\alpha,\beta,\gamma) \in SO(3)$, which can be determined using a peak detection [31].

Like in other spectral registration methods, the magnitude and shape of the peak can be used as a basis for estimating the uncertainty of the registration result [32], respectively the existence of multiple peaks can be used to handle mutually exclusive ambiguities in the sensor-data or the environment in robust SLAM [33], [34]. But these aspects are out of the scope of this paper. Furthermore, the results presented later on indicate that with only registration reasonable navigation and mapping can be achieved, which is, e.g., of interest for real-time processing with limited computational resources on Autonomous Underwater Vehicles (AUV) during a mission.

The FS2D process is illustrated in Fig. 2 with a simple example. As usual with correlation methods, the highest peak in the result space can be used for global registration, i.e., when no other information about the relative motion between scans is given. When an initial guess is known, the initial guess can be used to find the closest peak rotation $\Delta\alpha$ and use an initial guess translation to find the closest 2D peak, in the correlation, which represents the translation Δx and Δy. It is important to highlight, that FS2D requires no parameter tuning. Only the size of N has to be chosen, which represents a trade-off between accuracy and computation speed.

III. Experiments and Results

For an evaluation of the robustness of FS2D, we use several popular registration methods for which open implementations are available for comparison. Concretely, these methods are Generalized ICP [35], Super4PCS [36], NDT D2D [37] and NDT P2D [38]. Each method is implemented in C++. Thresholding is used when the input is required to be a point cloud. Two variants of our FS2D method are used. A global variant, which takes the best matching solution, i.e. highest peak, of the correlation processes, and a local variant, which utilizes the initial guess transformation. Additionally, we use two grid-sizes N to process the input data, which represent a fast method ($N = 64$) and a more accurate method ($N = 256$). For all experiments, the computation is performed on an i9-10850K CPU.



(a) Two sonar scans f_1, f_2 as 2D arrays (b) The related magnitudes $|F_1|, |F_2|$ of the scans (c) Resampled Magnitudes on $\mathbb{S}^2$ (d) Multiple peaks from the SOFT and the related scan alignments below

Fig. 2: A simple illustrative registration example with two scans and minor motion between them. There are multiple peaks in the end for the angle determination where the highest one, here close to Zero, respectively 2π, indicates the correct solution. If an initial guess is given, it can be used, e.g., the best peak from each of the corresponding translation correlations can be used to select the best fitting solution.

The comparison is done with three different datasets. Each dataset is a recorded list of sonar measurements, with a corresponding angle of the mechanical rotating sonar. The sonar data is then corrected by a motion compensation, measured by a Doppler Velocity Log (DVL) to create 360-degree underwater scans. The motion compensation can not be expected to be perfect. The registration is hence more challenging than using a Multi-Beam Echosounder (MBES), for which our method can of course be applied, too. The first one is the Fluvia Nautic dataset [39], sometimes also denoted as St.Pere dataset with 214 scans. In addition, we recorded a dataset inside the Bunker Valentin, which is located in Bremen, Germany (Fig. 1). The dataset with 121 scans covers a large rectangular basin of 150 m length and 10 m width (Fig.6). The Valentin dataset was recorded with a BlueROV2 [40], a Ping360 scanning imaging sonar [41] and the Water Linked A50 DVL [42]. The last dataset was created in a Gazebo simulation, where a complex labyrinth is created and a simulated BlueROV2 is collecting the sonar data. In that dataset, absolute ground truth is known and can be used for quantitative evaluation of registering consecutive scans with the different registration methods.

The focus of the analysis of the performance of the different methods is two folded. First, the rotation between two scans is the most challenging part of the registration, and it requires therefore a dedicated analysis. Second, the robustness of the registration methods, i.e., the influence of noise, is here of particular interest.

We therefore created a number of different experiments using simulated and real-world data for different aspects of the comparison. First, simple small motions of the same real world scans are used to create a baseline of the different methods. Second, the scans are step wise rotated to each other to study the effect of bigger rotational motions. Third, larger motions, added noise, and changing viewpoints with occlusions are used to test the robustness of the different methods. In the end, the registration of consecutive scans from the Gazebo simulation as well as from the Bunker Valentin dataset is used to create maps with each method.

A. *Random Motion in Real World Data without Noise*

The real world datasets, namely Fluvia Nautic and Bunker Valentin, are used to establish a baseline performance of all methods under ideal conditions, namely when each scan is to be registered with a copy of itself that is randomly rotated and translated by a small amount. The rotation angle is randomly chosen from within $\pm 5°$, while the translation in $x-y$ direction is randomly between ± 2 m. The randomly chosen rotation or translation follows a uniform distribution. The initial guess, when needed, is the identity matrix. For each scan, 50 different rotations plus translations are generated. After the generation process, the different registration methods are benchmarked in the rotation and translation error using the angle distance and the L^2 norm. Additionally, the different computation times are recorded.

TABLE I: Results when registering each sonar scan from two real-world datasets with a slightly moved copy of itself. The initial guess is the initial pose of each scan.

Dataset mean ± std		GICP	Super4PCS	NDT		FS2D		global FS2D	
				D2D 2D	P2D	N = 64	N = 256	N = 64	N = 256
Small Motion Bunker Valentin	Error Angle in degree	1.9 ± 1.5	3.3 ± 21	6.5 ± 29.6	1.1 ± 3.4	1.6 ± 0.8	**0.37 ± 0.22**	1.6 ± 0.8	0.37 ± 0.22
	Error L2 Norm in m	1.2 ± 0.65	4 ± 1500	0.49 ± 1.6	0.43 ± 0.96	0.4 ± 0.21	**0.1 ± 0.04**	0.4 ± 0.21	0.1 ± 0.04
	Computation Time in ms	**9.7 ± 6.1**	1251 ± 3135	20 ± 14	317 ± 417	12 ± 1.4	1428 ± 58	13 ± 1.6	1451 ± 58
Small Motion Fluvia Nautic	Error Angle in°	2 ± 1.3	2.3 ± 17	**0.3 ± 3.5**	0.9 ± 1.2	1.6 ± 0.9	0.6 ± 0.5	1.7 ± 1.9	0.6 ± 0.5
	Error L2 Norm in m	1.4 ± 0.59	32 ± 1354	**0.08 ± 0.7**	0.25 ± 0.7	0.9 ± 0.5	0.3 ± 0.3	0.9 ± 0.6	0.3 ± 0.3
	Computation Time in ms	21 ± 15	22760 ± 62260	46 ± 26	980 ± 980	**11 ± 1.9**	1327 ± 100	12 ± 1.3	1335 ± 61

In Tab. I the results are shown. The Bunker Valentin results have the lowest error in the FS2D methods. The increased dimension of $N = 256$ creates a more accurate result compared to $N = 64$. In contrast to the accuracy, the computation time increases by a factor of roughly 30 when increasing the grid size, while the dimension only increases by a factor of 4. This is to be expected as the complexity of SOFT as the most computation intensive part is $O(B^4)$ with $B = N/2$. The fastest method is the GICP. The computation time of the GICP is directly dependent on the number of points in the point cloud. Hence, it can be seen in the second dataset, where generally more points are in each scan, that the computation time increases.

Large outliers are present in the Super4PCS and NDT D2D 2D methods. The high standard deviation is the result of these outliers, which shows less robustness and high errors on a number of registrations.

Only the Fluvia Nautic dataset leads to different results in the NDT method cases. The mean is better than in the FS2D cases, but the standard deviation is still higher compared to the FS2D. Thus, the number of outliers is reduced with this dataset. Additionally, the computation time of the Super4PCS method compared to the first dataset is greatly increased.

In general, it can be seen that the computation time varies a lot in the iterative methods, especially compared to the spectral FS2D methods that have a fixed computation time.

B. Step-wise rotation of each Scan

The goal of this experiment is to test the robustness against rotations. Therefore, no translation component is used here in the motion between the scans. The rotation angle is stepwise increased by 5 degrees from an initial rotation of 5 degrees up to 120 degrees. Each method uses the initial guess of 0 degrees, i.e., the orientation of the original scan that is to be registered with its rotated counterpart. An exception are of course the global methods since they do not need an initial guess. This is done for both the Fluvia Nautic and the Bunker Valentin datasets.

In Fig. 3 and Fig. 4 the resulting mean error of the registration results of the rotated scans is shown. The standard deviation is indicated by the vertical line to the corresponding method. The dark-red line indicates the error of the initial guess without any registration.

It can be seen that the results differ based on the dataset. While the Bunker Valentin dataset includes mainly square structures, the Fluvia Nautic dataset also includes other

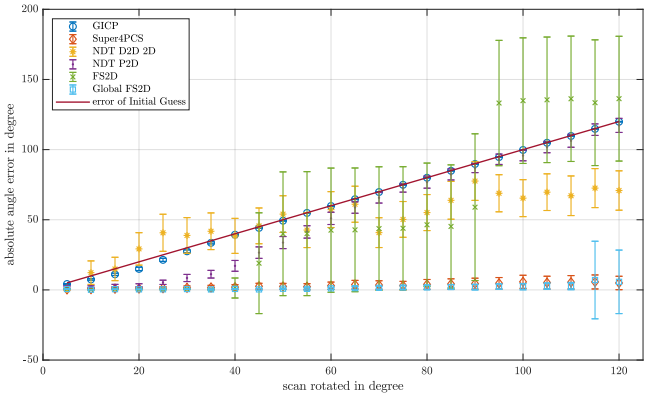


Fig. 3: Bunker Valentin dataset rotated by constant steps of 5 degree. The initial guess is always 0 degree, while the scans are rotated against each other.

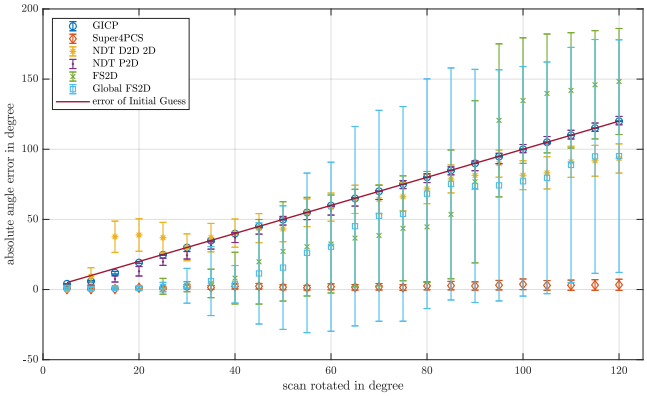


Fig. 4: Fluvia Nautic dataset rotated by constant steps of 5 degree. The initial guess is always 0 degree, while the scans are rotated against each other.

geometries. Therefore, the results differ depending on the registration method. It can be seen, that GICP, NDT D2D 2D, and NDT P2D can keep the error low in the beginning, but with higher rotations, the error also increases. While NDT P2D can longer keep some correct estimates than the other two methods, it also converges to the dark-red line of the initial guess error. Our FS2D can hold the correct angle longer, namely until an angle of about 40 degree is exceeded; afterwards it also gets unstable. More precisely, while some registrations still lead to the correct angle, multiple registrations fail and the method loses its reliability. Using the global registration of our FS2D, the correct angle can be found much better, i.e., not using the increasingly

erroneous initial guess is here an advantage for FS2D.

In the Bunker Valentin dataset, there exist four outliers in the last 4 rotation angles determined by FS2D, which increase the standard deviation and the mean. In contrast, in the Fluvia Nautic dataset, the global variation makes the FS2D perform better. The best performing registration in this noise-free scenario is the Super4PCS, which can correctly register the datasets without a big error with increasing rotation.

The different results are among others also a result of multiple differences in the datasets. The sensors for the motion compensation in the Fluvia Nautic dataset, which was published in 2008, are not as up-to-date than the ones used for the recent recording of the Valentin dataset. Also, the marina covered in the Fluvia Nautic dataset is more complex than the simple rectangular basin of the bunker.

C. Robustness Analysis

The robustness of a registration method is critical for its usage in real world environments. Therefore, we investigate in this section the influences of (A) noise, (B) larger motions, and (C) the effects of changes in the viewpoint like occlusions.

First (A), the experiment of Sec. III-A is repeated with added noise. The noise consists of salt and pepper noise, where measurements of the sonar are randomly corrupted in a uniform manner, plus Gaussian white noise on every sonar measurement. The first simulates the noise in the signal processing of sonar while the second component relates to the typical changes in intensity depending on, e.g., viewing angles.

Second (B), larger motions of $\pm 15°$ and ± 10 m between the scans are used in a variation of the experiment of Sec. III-A.

Third (C), an additional dataset is used to test for the effects of changes in the viewpoints, especially occlusions. So, each of the two to be registered scans do not perfectly overlap; each one typically contains areas not contained in the other. This dataset is generated with a realistic sonar simulation in Gazebo using an environment that resembles layouts found in the basement of the bunker. To generate the data, a simulated BlueROV with a Ping360 sonar and a DVL is used to explore the environment. The advantage of the simulation is that it guarantees accurate ground truth data for the poses of the consecutive scans. Also, this dataset includes realistic noise and realistic motions, i.e., changes in angle and translation, between the scans.

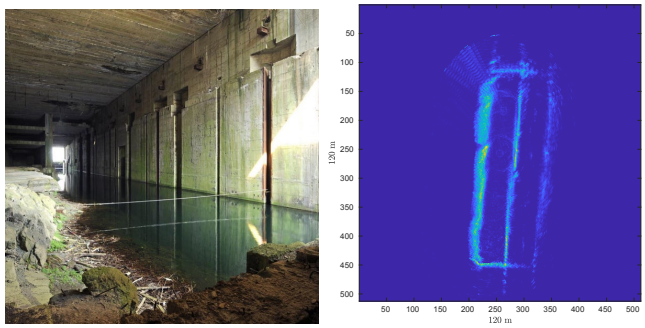
In Tab. II the errors of the methods in the different robustness experiments are shown. First and foremost, it can be noticed that FS2D has the best performance with respect to robustness among all methods. This comes at the cost of higher computation times - which are nonetheless still fully sufficient for real-time performance given the time it takes to record a sonar scan.

As a result that is at first glance very surprising, it can be noticed that errors may slightly decrease with noise added to the scans. The main reason for that is that the likelihood of outliers that are completely wrong decreases, i.e., there is a tendency to end up in local minima closer to the initial guess if the registration fails under the presence of noise.

As can be expected, larger transformations lead to a big error in the non-global methods that require a good initial guess. Especially the global FS2D method has a much smaller error compared to the version with an initial guess. If the initial guess from, e.g., very noisy odometry, is quite error-prone, it can even be counter-productive and lead to worse results.

Finally, the Gazebo based experiment provides a quantitative analysis that is very close to realistic use-cases with occlusions, noise, and navigation-data from an Inertial Measurement Unit (IMU) and DVL. Once there is only partial overlap in the scans combined with noise, most registration methods are severely challenged. Here, the noisy but reasonable initial guess from navigation can boost the performance of FS2D. Regarding the computation time, a significant gain can be achieved when using FS2D with a grid-size of $N = 64$ instead of $N = 256$ at minor losses in robustness. Nonetheless, the time to record a scan with some well 20 sec is the dominant factor in mapping with sonar as mentioned before. Hence, the higher robustness of a larger grid-size is preferable even if it takes about 1.3 sec to register the scans with $N = 256$. It is of course also conceivable to use an intermediate grid with $N = 128$ depending on the application.

D. Mapping with consecutive scans



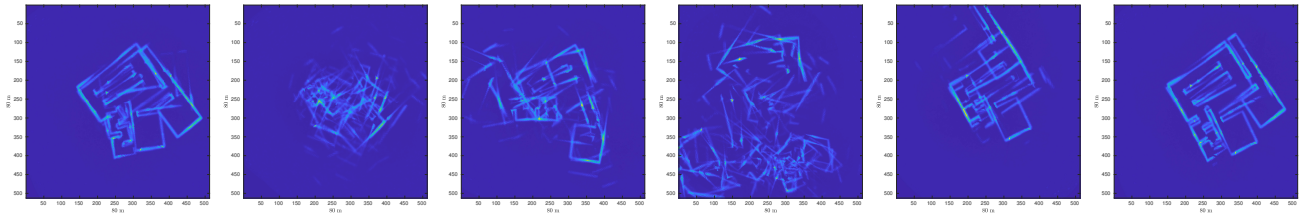
(a) Basin for submarine deployment (b) FS2D-map of the basin

Fig. 6: The Valentin dataset with 121 scans of the basin for the deployment of the submarines (left) is well suited for a simple qualitative assessment of a mapping result. The basin is a simple rectangular shape of 150 m × 10 m, which is well captured by a map using FS2D registration of the consecutive scans (see also Fig.7 for a comparison with other methods).

Finally, we use the different methods to generate maps by registering consecutive scans from the Gazebo and the Valentin dataset. A grid-size of $N = 256$ is used for FS2D. In Fig. 5 the resulting maps of the underwater environment simulated in Gazebo are shown. Mapping with FS2D captures the real layout the best. With GICP, the core layout is also recognizable, but there are larger structural errors. In the other cases, including the global version of FS2D, outliers in the registrations lead to severely wrong maps. The maps created with the Bunker Valentin dataset are shown in Fig. 7. In this case, all methods show roughly the outline of the

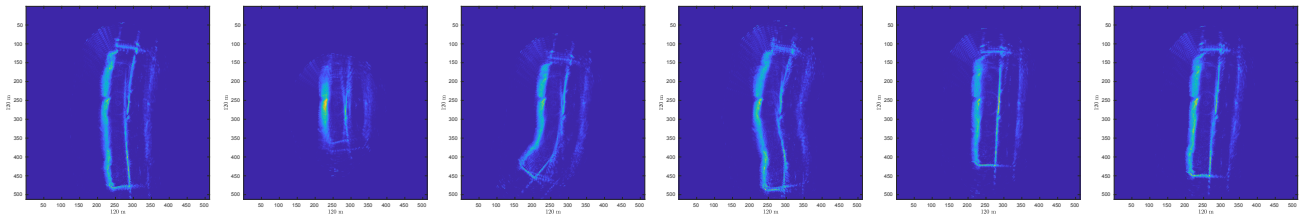
TABLE II: Results of the experiments to test robustness with (A) noise, (B) larger motion, and (C) consecutive scans with occlusions and realistic noise and motions of a BlueROV with a Ping360 sonar.

Dataset mean ± std		GICP	Super4PCS	NDT		FS2D		global FS2D	
				D2D 2D	P2D	64	256	64	256
(A) Noise Bunker Valentin	Error Angle in°	2.4 ± 1.5	3.2 ± 9.7	4 ± 21	1.4 ± 3.4	1.7 ± 0.93	**0.4 ± 0.3**	1.7 ± 2.4	0.44 ± 0.3
	Error L2 Norm in m	1.4 ± 0.6	0.6 ± 0.6	0.4 ± 1	0.4 ± 0.8	0.5 ± 0.3	**0.1 ± 0.05**	0.5 ± 0.3	0.1 ± 0.05
	Computation Time in ms	**7.8 ± 3.7**	3090 ± 4002	20 ± 11	335 ± 378	11 ± 1.5	1337 ± 24	12 ± 1.4	1344 ± 25
(A) Noise Fluvia Nautic	Error Angle in°	2.1 ± 1.3	2 ± 5.7	0.7 ± 5.7	1.2 ± 1.4	1.7 ± 0.95	**0.6 ± 0.6**	1.7 ± 2.1	0.6 ± 0.6
	Error L2 Norm in m	1.4 ± 0.6	2 ± 2.7	0.32 ± 1.2	0.33 ± 0.77	0.97 ± 0.55	**0.29 ± 0.28**	0.97 ± 0.6	0.29 ± 0.28
	Computation Time in ms	**9.5 ± 4.7**	1713 ± 1043	31 ± 9.9	328 ± 335	12 ± 1.1	1360 ± 66	12 ± 1.1	1368 ± 66
(B) Larger Motion Bunker Valentin	Error Angle in°	7.9 ± 5.3	1.6 ± 12	49 ± 66	13 ± 26	1.5 ± 0.95	0.4 ± 0.26	1.5 ± 0.95	**0.4 ± 0.26**
	Error L2 Norm in m	7.5 ± 3	1.7 ± 73	6.3 ± 6.9	6.9 ± 5.9	2.7 ± 4.6	3 ± 5.2	0.6 ± 0.3	**0.13 ± 0.07**
	Computation Time in ms	**9.7 ± 9.4**	1522 ± 3486	28 ± 18	494 ± 507	12 ± 1.4	1427 ± 58	13 ± 1.5	1450 ± 58
(B) Larger Motion Fluvia Nautic	Error Angle in°	7.5 ± 4.5	1.3 ± 11	17 ± 40	5.9 ± 11	1.8 ± 1.7	**0.9 ± 1.4**	3.8 ± 17	1.1 ± 5.8
	Error L2 Norm in m	7.6 ± 2.9	9.3 ± 797	4.7 ± 9.4	4.6 ± 5.7	1.7 ± 2.3	1.3 ± 3.4	1.8 ± 2.7	**0.6 ± 1.3**
	Computation Time in ms	26 ± 26	19070 ± 43260	66 ± 42	1213 ± 1093	**11 ± 1.5**	1327 ± 64	12 ± 1.2	1335 ± 62
(C) Consecutive Scans (Gazebo)	Error Angle in°	2.2 ± 2.6	41 ± 59	8.6 ± 22	31 ± 30	2 ± 1.6	**1 ± 0.76**	45 ± 64	26 ± 57
	Error L2 Norm in m	0.56 ± 0.4	5.3 ± 5.3	2.1 ± 3.7	4.8 ± 4.2	1.8 ± 2.3	**0.53 ± 0.45**	5.3 ± 6.4	2.9 ± 5.2
	Computation Time in ms	**7.3 ± 3.5**	58430 ± 154500	6.5 ± 3.5	284 ± 204	11 ± 1.1	1347 ± 24	13 ± 1.2	1391 ± 27



(a) GICP (b) Super4PCS (c) NDT D2D (d) NDT P2D (e) FS2D Global (f) FS2D

Fig. 5: The maps generated by the different methods with consecutive registration of the scans from the Gazebo dataset with a simulated BlueROV, a Ping360 sonar, and DVL based navigation for initial guesses. The FS2D registration leads to a result that is closest to the actual layout of the environment.



(a) GICP (b) Super4PCS (c) NDT D2D (d) NDT P2D (e) FS2D Global (f) FS2D

Fig. 7: The maps generated by the different methods with consecutive registration of 30 scans from the real world Valentin dataset. The FS2D registration leads again to a result that is closest to actual environment in form of a rectangular basin of size 150 m × 10 m (Fig6).

basin with more or less distortions. The best mapping result is again achieved with FS2D.

IV. Conclusion

In this paper, the Fourier-SOFT 2D (FS2D) registration method was introduced. FS2D operates in the frequency domain. It can hence use the well-known decoupling of rotation and translation. As usual, the difficulty is to determine the underlying rotation of the to be registered data. This is solved in FS2D in a novel way based on a projection of the Fourier magnitude on a sphere and the SO(3) Fourier Transform (SOFT). This leads to a very robust registration method, i.e., it can cope with high noise in the sensor data and larger changes in viewpoints with hence, e.g., occlusions. The method is largely parameter-free. Only the grid-size $N = 2^k$ (typically $N = 64, 128, 256$) has to be chosen to trade accuracy with the deterministic, i.e., fixed computation time. An implementation of FS2D is openly available and released with this paper.

As a use-case of FS2D, we presented the registration of underwater scans with a mechanically rotating sonar and short-term motion compensation to generate scan images. Note that this is a quite challenging application scenario compared to the use of higher quality scans from, e.g., a Multi-Beam Echsounder (MBES) for which our registration method is of interest, too.

In a comparison with multiple standard, openly available methods, we demonstrated the robustness of FS2D. The trade-off is a higher computation time compared to the other methods. But even with a computation time of about 1.3 sec for the high accuracy version with $N = 256$, this is much less than the more 20 sec that it takes to record a scan. Furthermore, there are multiple options to accelerate the computation of FS2D. It is due to its regular, grid-based processing, e.g., well-suited for an implementation in CUDA, which is left for future work.

References

[1] L. Paull, S. Saeedi, M. Seto, and H. Li, "Auv navigation and localization: A review," *IEEE Journal of Oceanic Engineering*, vol. 39, no. 1, pp. 131–149, 2014.

[2] Q. Xi, T. Rauschenbach, and L. Daoliang, "Review of underwater machine vision technology and its applications," *Marine Technology Society Journal*, vol. 51, no. 1, pp. 75–97, 2017. [Online]. Available: https://www.ingentaconnect.com/content/mts/mtsj/2017/00000051/00000001/art00009https://doi.org/10.4031/MTSJ.51.1.8

[3] J. Horgan and D. Toal, *Computer Vision Applications in the Navigation of Unmanned Underwater Vehicles*. Intech, 2009. [Online]. Available: https://www.intechopen.com/books/underwater_vehicles/computer_vision_applications_in_the_navigation_of_unmanned_underwater_vehicles

[4] J. P. Marage and Y. Mori, *Sonar and Underwater Acoustics*. Wiley, 2013.

[5] X. Lurton, *An Introduction to Underwater Acoustics - Principles and Applications*, ser. Geophysical Sciences. Springer, 2010.

[6] R. P. Hodges, *Underwater Acoustics: Analysis, Design and Performance of Sonar*. Wiley, 2010.

[7] 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, Boston: De Gruyter, 2022, pp. 133–168. [Online]. Available: https://doi.org/10.1515/9783110714692-006

[8] L. Cheng, S. Chen, X. Liu, H. Xu, Y. Wu, M. Li, and Y. Chen, "Registration of laser scanning point clouds: A review," *Sensors*, vol. 18, no. 5, p. 1641, 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/5/1641

[9] J. Salvi, C. Matabosch, D. Fofi, and J. Forest, "A review of recent range image registration methods with accuracy evaluation," *Image and Vision Computing*, vol. 25, no. 5, pp. 578–596, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0262885606001594

[10] L. G. Brown, "A survey of image registration techniques," *ACM Computing surveys*, vol. 24, no. 4, pp. 325–376, 1992.

[11] Y. Xu, R. Zheng, S. Zhang, and M. Liu, "Robust inertial-aided underwater localization based on imaging sonar keyframes," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–12, 2022.

[12] 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 Systems Journal*, vol. 14, no. 3, pp. 3301–3312, 2020.

[13] A. Burguera, "Cluster-based scan matching for robust motion estimation and loop closing," in *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019, Conference Proceedings, pp. 2512–2517.

[14] J. Li, M. Kaess, R. M. Eustice, and M. Johnson-Roberson, "Pose-graph slam using forward-looking sonar," *IEEE Robotics and Automation Letters (RAL)*, vol. 3, no. 3, pp. 2330–2337, 2018.

[15] A. Burguera, "A novel approach to register sonar data for underwater robot localization," in *IEEE Intelligent Systems Conference (IntelliSys)*. IEEE, 2017, Conference Proceedings, pp. 1034–1043.

[16] N. Hurtos, D. Ribas, X. Cufí, Y. Petillot, and J. Salvi, "Fourier-based registration for robust forward-looking sonar mosaicing in low-visibility underwater environments," *Journal of Field Robotics*, vol. 32, no. 1, pp. 123–151, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21516

[17] A. Mallios, P. Ridao, D. Ribas, and E. Hernandez, "Scan matching slam in underwater environments," *Autonomous Robots*, vol. 36, no. 3, pp. 181–198, 2014. [Online]. Available: http://dx.doi.org/10.1007/s10514-013-9345-0

[18] N. Hurtos, S. Nagappa, X. Cufí, Y. Petillot, and J. Salvi, "Evaluation of registration methods on two-dimensional forward-looking sonar imagery," in *MTS/IEEE OCEANS*, 2013, Conference Proceedings, pp. 1–8.

[19] A. Mallios, P. Ridao, M. Carreras, and E. Hernandez, "Navigating and mapping with the sparus auv in a natural and unstructured underwater environment," in *OCEANS 2011*, 2011, Conference Proceedings, pp. 1–7.

[20] E. Hernandez, P. Ridao, D. Ribas, and A. Mallios, "Probabilistic sonar scan matching for an auv," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, 2009, Conference Proceedings, pp. 255–260. [Online]. Available: 10.1109/IROS.2009.5354656

[21] 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 *OCEANS 2009*, 2009, Conference Proceedings, pp. 1–6. [Online]. Available: 10.1109/OCEANSE.2009.5278219

[22] H. Bülow, M. Pfingsthorn, and A. Birk, "Using robust spectral registration for scan matching of sonar range data," in *7th Symposium on Intelligent Autonomous Vehicles (IAV), IFAC*. IFAC, 2010, Conference Proceedings.

[23] D. Ribas, P. Ridao, J. Neira, and J. D. Tardos, "Slam using an imaging sonar for partially structured underwater environments," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, P. Ridao, Ed., 2006, Conference Proceedings, pp. 5040–5045.

[24] X. Tong, Z. Ye, Y. Xu, S. Gao, H. Xie, Q. Du, S. Liu, X. Xu, S. Liu, K. Luan, and U. Stilla, "Image registration with fourier-based image correlation: A comprehensive review of developments and applications," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 10, pp. 4062–4081, 2019.

[25] B. Reddy and B. Chatterji, "An fft-based technique for translation, rotation, and scale-invariant image registration," *Image Processing, IEEE Transactions on*, vol. 5, no. 8, pp. 1266–1271, 1996.

[26] Q.-S. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of fourier-mellin transforms for image registration and recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 12, pp. 1156–1168, 1994.

[27] P. E. Anuta, "Spatial registration of multispectral and multitemporal digital imagery using fast fourier transform techniques," *IEEE Transactions on Geoscience Electronics*, vol. 8, no. 4, pp. 353–368, 1970.

[28] P. Kostelec and D. Rockmore, "Ffts on the rotation group," *Journal of Fourier Analysis and Applications*, vol. 14, no. 2, pp. 145–179, 2008. [Online]. Available: http://dx.doi.org/10.1007/s00041-008-9013-5DO-10.1007/s00041-008-9013-5

[29] H. Bülow and A. Birk, "Scale-free registrations in 3d: 7 degrees of freedom with fourier-mellin-soft transforms," *International Journal of Computer Vision (IJCV)*, vol. 126, no. 7, pp. 731–750, 2018.

[30] K. Zuiderveld, *Graphic Gems IV*. San Diego: Academic Press Professional, 1994.

[31] N. Yoder, "peakfinder(x0,sel,thresh,extrema,includeendpoints,interpolate) matlab central file exchange," Retrieved September 11, 2022. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/25500-peakfinder-x0-sel-thresh-extrema-includeendpoints-interpolate

[32] M. Pfingsthorn, S. Schwertfeger, H. Bülow, and A. Birk, "Maximum likelihood mapping with spectral image registration," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE Press, 2010, Conference Proceedings.

[33] M. Pfingsthorn and A. Birk, "Generalized graph slam: Solving local and global ambiguities through multimodal and hyperedge constraints," *International Journal of Robotics Research (IJRR)*, vol. 35, no. 16, pp. 601–630, 2016.

[34] ——, "Simultaneous localization and mapping (slam) with multimodal probability distributions," *The international Journal of Robotics Research*, vol. 32, no. 2, pp. 143–171, 2013. [Online]. Available: http://ijr.sagepub.com/content/early/2012/10/05/0278364912461540.abstract

[35] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp." vol. 2, no. 4, p. 435, 2009.

[36] N. Mellado, D. Aiger, and N. J. Mitra, "Super 4pcs fast global pointcloud registration via smart indexing," vol. 33, no. 5, pp. 205–215, 2014.

[37] T. Stoyanov, M. Magnusson, H. Andreasson, and A. J. Lilienthal, "Fast and accurate scan registration through minimization of the distance between compact 3d ndt representations," *The International Journal of Robotics Research*, vol. 31, no. 12, pp. 1377–1393, 2012.

[38] M. Magnusson, A. Lilienthal, and T. Duckett, "Scan registration for autonomous mining vehicles using 3d-ndt," *Journal of Field Robotics*, vol. 24, no. 10, pp. 803–827, 2007.

[39] D. Ribas, P. Ridao, J. D. Tardós, and J. Neira, "Underwater slam in man-made structured environments," *J. Field Robot.*, vol. 25, no. 11-12, pp. 898–921, 2008.

[40] "BlueROV2: The world's most affordable high-performance ROV." Blue Robotics, accessed September 2022. [Online]. Available: https://bluerobotics.com/store/rov/bluerov2

[41] "Ping360; a mechanical scanning sonar for navigation and imaging." Blue Robotics, accessed Septem-

ber 2022. [Online]. Available: https://bluerobotics.com/store/sensors-sonars-cameras/sonar/ping360-sonar-r1-rp/
[42] "DVL A50; the dvl-a50 is the world's smallest commercially available doppler velocity log." WaterLinked, accessed September 2022. [Online]. Available: https://store.waterlinked.com/product/dvl-a50/