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

Adaptive Heading for Perception-Aware Trajectory Following

Jonatan Scharff Willners[1], Sean Katagiri[1], Shida Xu[1], Tomasz Łuczyński[1], Joshua Roe[1] and Yvan Petillot[1]

Abstract— **This paper presents an adaptive heading approach for perception awareness during trajectory following. By adapting the heading of a robot to improve the feature tracking in the current mapped environment, the accuracy in localisation can be improved. This can have a significant advantage for autonomous operations in GPS-denied environments such as subsea or in caves. The aim of the proposed approach is to position the sensor used for perception and feature tracking in such a way that it; obtains a view that contains a good observation of the previously mapped environment, face forward along the direction of travel, reduces the change in heading and view the perceived environment along the surface's estimated normals. These 4 objectives create a weighted utility function that is used to find the most beneficial heading. The benefit is a system that improves feature tracking for simultaneous localisation and mapping (SLAM) while considering the safety of the robot by being aware of its surrounding. To sense the environment, a simulated sensor is discretised to a set of vertical rays based on the vertical field of view. The vertical rays are swept 360 degrees around a position to evaluate for a new heading. This allows for the simulated sensor data from ray casting to be reused and therefore reduces the computational load to find the heading which maximises the utility function. The paper is focused on holonomic robots capable of controlling the robot's heading or sensor orientation independently from the position. We present results and evaluation in a simulated environment where we show a great improvement in the SLAM's pose estimation. In addition, we endow an autonomous underwater vehicle (AUV) with the proposed approach during field trials and present the result in two different environments.**

I. Introduction and Background

During autonomous operation in GPS-denied environments such as indoors, in caves or sub-sea, the lack of an absolute positioning system creates an increased challenge during navigation. In many cases the environment is unknown prior to execution and online mapping and localisation such as SLAM is required to ensure a safe operation. While navigating autonomously in such an environment, it is important to move collision-free. A key element to this consists of both building an accurate map as well as being confident in your position within the obtained map. An important aspect of obtaining high accuracy in SLAM is to continuously have features to track. In sub-sea environments, visual systems such as cameras have a short range of vision due to environmental conditions. In addition, the underwater domain usually has a sparse amount of features to track, hence, making use of what is available can greatly improve navigation. The approach of taking the environment into account for improving localisation during operation is generally referred to as perception-aware planning or active perception. The concept of active perception was introduced by Bajcsy [1] to tackle the problem of not only what is being seen, but how to look at it. While sensors such as cameras are passively collecting data, we can actively put them in positions that can generate more informative data. For the last decade, perception-aware planning has mainly seen usage for aerial vehicles as quadrotors. Cosante et al. utilises the photometric and geometric information observed in the environment to find a trajectory that reduces the localisation uncertainty [2], [3]. The work in [4] actively explores unknown regions where the vehicle is planned to navigate through using a local planner. Falanga et al. present an approach where the goal for a quadrotor is to maintain a view of specific features as an objective for the control system [5]. Similarly, in [6] a set of a priori known static features are used while planning a path to maintain the view of at least one feature. In [7], a Fisher information field is used to select the heading along a path when the environment and its key points are previously known. A local history is used to keep track of features instead of fusing them into a global map in [8], focusing on fast and safe local operations. An environmental adaptive approach was presented by Quan et al. [9] which adjusts the velocity of a quadrotor based on the distance from the environment. Similarly, adjusting the velocity based on the distance to the environment on remotely operated vehicles (ROVs) has been used with omnidirectional cameras [10]. For terrestrial robots, a belief roadmap (BRM) was constructed based on the pose graph from SLAM [11], [12]. When possible, the robot will optimise its path in the BRM to reduce its uncertainty. This was based on the uncertainty assigned to the poses in the pose graph. For testing the approach, the terrestrial robot was equipped with a 3D laser, and hence calculating a heading is not required at the poses. The work of perception-aware planning in the sub-sea domain is sparse. Autonomous exploration [13], [14] contains elements of it, but the planning for exploring the environment is not focused on reducing the localisation error, but instead on observing new regions of the environment. We have previously presented work that addresses the localisation aspect for AUVs by endowing them with an autonomous relocalisation action when visual feature tracking is lost [15]. Xanthidis et al. presents a perception-aware planning framework aiming to track application-defined clusters of features for underwater robots [16], [17]. A field that shares elements of perception-aware planning is active SLAM. Stachniss et al. [18] uses an

*This work was supported by the EPSRC funded ORCA-HUB (EP/R026173/1) and EU METRICS (Grant agreement ID: 871252)

[1]J. Scharff Willners *(Corresponding author)*, S. Katagiri, S. Xu, T. Łuczyński, J. Roe and Y. Petillot are with the Institute of Sensors, Signals and Systems, Heriot-Watt University, Edinburgh, UK, {j.scharff_willners, s.katagiri, sx2000, t.luczynski, joshua.roe y.r.petillot}@hw.ac.uk

active loop-closure procedure by opportunistically searching for trajectories to revisit during previously explored areas for the SLAM. For underwater, active SLAM has been used with sonars for ship hull inspection [19] and during exploration [13].

In this paper, we present an approach for adaptive heading to improve perception-awareness while following a trajectory. The next heading for a waypoint is selected using a custom simulated sensor in 360 degrees in the current perceived map of the environment and the current state of the system. The approach has been evaluated to show its advantage in simulation where ground truth is accessible. In addition, it has been used online during sub-sea operations. The presented approach is applicable for any domain as long as position and sensor orientation can be controlled independently. The result in this paper however will focus on the subsea domain and hence requires a robot with at least 4 degree of freedom (DoF) (x,y,z and yaw) of control.

The paper presents the following novelty and contributions:

- An approach for heading control that can be separated from path-planning, and hence work independently from path planners and arbitrary trajectories.
- The simulated sensor model swept 360 degrees around the position to find the heading with the highest utility. The approach allows for reusing the simulated sensor data multiple times for various headings which reduces computational effort.
- A utility function that selects the next heading based on the perceived map, the map's normals, the current state of the robot and the next waypoint.
- Extensive experiments in simulation and online operation in field trials show that the proposed approach leads to significant improvement in SLAM's pose estimation.

II. Method

We propose an approach to find the next heading of a robot based on simulating the sensor used for SLAM, by rotating a set of vertical rays representing the sensor's vertical field of view (FoV), in 360 degrees, divided into a discrete set of headings at the desired position. The rays are used to estimate what a set of discrete headings would observe in the currently mapped environment. The simulated sensor in addition to the current state and plan of the vehicle is added to a sum of weighted functions to determine which heading is most beneficial for the next waypoint of the robot.

A. Mapping and Simulated Perception

The observed environment by the SLAM system [20] is represented as a point cloud and an OctoMap [21]. An OctoMap is a sparse representation constituted by voxels as a 3-dimensional occupancy grid. The voxels in the OctoMap are initially marked as unknown and during operation, as more of the environment is observed, the space in the map is marked as occupied and the space between the sensor and its observations are marked as free. An OctoMap gives fast access to voxels and efficient ray casting due to the discrete representation of the environment.

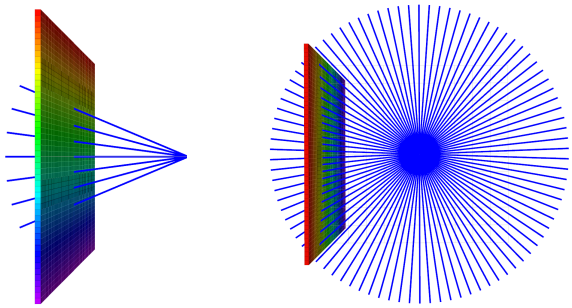


Fig. 1: Left: A set of vertical rays based on $r_{vFoV} = 7$ seen from the side. Right: Simulated rays based on $r_h = 90$ seen from above. A total of $r_{vFoV} * r_h$ rays are used to simulate the sensor in r_h headings. A sensor at a selected heading can be simulated using the vertical sets belonging to the horizontal headings within the selected headings horizontal field of view.

To simulate how the robot is expected to view the environment, we use ray casting within the OctoMap. However, instead of simulating the sensor's FoV as individual measurements at a pose ($\mathcal{P}$), we simulate the sensor in 360 degrees rotated around z at $\mathcal{P}$. This gives the opportunity to reuse the results from the same ray casting procedures for evaluating multiple headings at $\mathcal{P}$, by including the set of vertical rays which fall into the horizontal FoV of the robot's sensor. We define the number of discrete headings as r_h. For each heading, a vertical set of r_{vFoV} rays are used. Therefore in order to obtain data for the r_h headings, $r_{vFoV} * r_h = r_n$ rays are needed. In Fig. 1 (left) a set of vertical rays ($r_v = 7$) can be seen and in Fig. 1 (right) the set of headings are shown for $r_h = 90$. Simultaneously simulating the discrete set of headings for a sensor reduces the number of ray casting operations significantly as opposed to when simulating the sensor for the same number of headings (r_h times), which would lead to $r_{hFoV} * r_{vFoV} * r_h$ ray casting procedures. Therefore, the proposed approach reduces the number of ray castings required for the same result by r_{hFoV} times. While ray casting in a discrete representation as an OctoMap is fairly fast, it is nonetheless the most computational expensive procedure for finding the best heading. As the sensing range of different sensors and domains varies, either due to physical limitation (eg., laser or acoustics), or the minimum size where a feature would be distinguishable (such as optical cameras), the maximum distance it should be considered to be used should be accounted for, hence a maximum sensing distance is used. When ray casting is performed, the following information is stored to be used for the utility function to select the heading; if it hit a voxel (with the voxel's hash key) in the OctoMap or if it was a miss, and the angle towards the normal of the voxel. The angle of sensing compared to the normal of a surface has previously been used as a metric to evaluate the quality of the collected data of the perceived environment by Vidal et al. [14].

B. Selection of Heading

When a new waypoint is issued in the current trajectory we apply the presented algorithm to select the heading for the robot to take at the next waypoint. To select the orientation the described 360 degrees simulation of the sensor is performed at the waypoint to get the data for each ray. To evaluate the orientation of the robot using the simulated rays, each heading is selected and the rays which are within the sensor's FoV are selected to be evaluated using a utility function for this particular heading. The heading which maximises the utility function (1) is then selected as the next heading for the robot, as in (2), where H is the set of heading evaluated.

$$h_{score} = \mathcal{R}w_r + \mathcal{N}w_n + \mathcal{F}w_f + \mathcal{D}w_d \tag{1}$$

$$new\_heading = \arg\max_{h \in H}(h_{score}) \tag{2}$$

The utility function, (1), is based on the following 4 objectives:

- **Ray hits ($\mathcal{R}$):** In the SLAM system, each observation of a landmark in the camera frame can provide a constraint to the pose estimation. During the pose optimisation, more landmark observations can help SLAM filter out outliers caused by incorrect feature matching and improve the accuracy, but insufficient observations may cause the pose estimation to be dominated by the noise. The goal of the ray hits score is to increase the number of landmark observations SLAM obtains from the camera frame. The utility score for the simulated hits of the sensor is based on the average of the number of hits and the number of unique hits as in (3), where $r_{FoV} = r_{hFoV} * r_{vFoV}$ is the total number of rays for one heading. The unique hits refer to the number of unique voxels seen among the ray castings considered.

$$\mathcal{R} = \frac{hits + unique\_hits}{2 * r_{FoV}} \tag{3}$$

  $\mathcal{R}$ results in a score in the range $[0.0, 1.0]$. Using unique hits in addition to the total hits is due to the possibility of a heading would be in close proximity to a part of the occupied map, it could have all rays hit the same voxel, which would not lead to seeing a diverse part of the environment which could result in poor pose estimation due to low amount of matched features. However, at a certain point, the number of unique features to track is likely to be high enough to perform visual odometry with high accuracy, hence if $unique\_hits > \mathcal{R}_{th}$, $\mathcal{R}$ will be set to 1.0.
- **Face normals ($\mathcal{N}$):** For each ray, the angle towards the normal of the voxel hit is calculated. The average ($\mathcal{N}_{avg}$) of all of these ray-normal angles is then calculated.

$$\mathcal{N} = \frac{\pi - \mathcal{N}_{avg}}{\pi} \tag{4}$$

  As each ray-normal angle has a value in the range $[0, \pi]$ then $\mathcal{N}_{avg}$ has a value in the same range and the score for the average angle to normal, $\mathcal{N}$, is in the range of $[0.0, 1.0]$. Hence, when all rays are facing exactly towards the corresponding normals a score of 1.0 is achieved, and a score of 0.0 if all rays are facing in the same direction as the normals.
- **Face next ($\mathcal{F}$):** The closer a heading is to face the next waypoint the better the score, as this is facing along the direction traveled and hence beneficial for observing obstacles to avoid. If facing straight towards the next waypoint a score of 1.0 is achieved. The score is linearly decreasing to 0.0 when facing straight away from the next waypoint.
- **Heading difference ($\mathcal{D}$):** The difference between the current orientation and the evaluated heading, calculated and set to a score in the range $[0.0, 1.0]$, where 1.0 means no heading difference compared to the current state and 0.0 means facing the opposite direction as the current state of the platform. The component for heading difference reduces the sudden movements in orientation in cases where the SLAM has mapped the environment surrounding the robot at multiple sides, it will prefer to maintain an orientation similar to its current instead of shifting which parts of the environment to focus on.

Each individual score is in the range $[0.0, 1.0]$ and the total score for a heading is the weighted sum between these as in (1). The weights are normalised so that their sum is 1.0. Hence, the total utility function is in the range of $[0.0, 1.0]$ where the higher the score the more beneficial the heading is considered to be. For each of the discrete headings, the one which results in the highest score, (1) as in (2), is selected as the next waypoint's heading.

III. Results

We extend the framework for autonomous operations previously described in [22] with the proposed adaptive heading approach. The framework, which is used in both the simulated comparison and for the operations during field trials, is based on Robot Operating System (ROS) [23]. For path planning informed-RRT* (I-RRT*) [24] is implemented using Open Motion Planning Library (OMPL) [25] with Flexible Collision Library (FCL) [26] used for collision detection. The path-planning is performed in 3D (x, y, z). The autonomous navigation framework implemented verifies that the current trajectory is valid every second. If a new addition to the sensed map would obstruct the current path, a replanning procedure is initiated, leading to continuous collision-free operation in unknown environments. In the case that a sensed object is too close to the current position the vehicle will autonomously move away to a safe position. The goal waypoint(s) were issued manually. For the evaluation, both in simulation and in field the following parameters were used: The weights before normalisation are set to $w_r = 4$, $w_n = 1$, $w_f = 3$, $w_h = 3$ and the number of ray hits threshold as $\mathcal{R}_{th} = 0.3$. The weights were obtained from manual tuning

to improve performance, with the choice that feature tracking (ray hits) should be the priority and highest weight among the four objectives.

A. Simulated comparison

We evaluate the presented approach for adaptive heading selection when following a trajectory in a simulated environment to compare the pose estimation with the ground truth. We show how the approach compares to following the same trajectory while facing forward in the direction of travel or keeping the orientation of the goal along the trajectory. The comparison is performed using UUV Simulator [27] and the SLAM system from Xu et al. [20] which fuse the Doppler Velocity Log (DVL), depth sensor and Inertial Measurement Unit (IMU) with the visual information from a stereo camera setup. As the DVL measures relative motion, it contains a significant amount of noise compared to the sensors which are capable of measuring absolute values, and hence we add a Gaussian noise with a standard deviation of 0.01 metres per second.

In the comparison we evaluate moving along the same trajectory between the same start and goal pose around a structure, following the same waypoints, with 3 different methods for selecting the heading: 1) facing the direction of travel (Face forward, see Fig. 2c) 2) keeping the goal waypoint's heading during the path (Lock goal, see Fig. 2d), and 3) the proposed approach for perception-aware heading selection (see Fig. 2e). The structure, which is a blowout preventer, and the AUV can be seen in Fig. 2a. The trajectory consists of 3 waypoints, and the resulting path from the proposed approach can be seen in Fig. 2b. During the simulation, the maximum ray casting range is set to 10 metres. The error metrics in Table I and Fig. 2b-2e were generated using [28].

TABLE I: Simulated Localisation Comparison

	Face forward	Lock goal	Proposed
Translation Error [m]			
Max	0.87	0.96	0.33
Mean	0.25	0.29	0.12
RMSE	0.28	0.33	0.14
Rotational Error [deg]			
Max	3.91	4.31	2.81
Mean	0.90	2.38	0.70
RMSE	1.02	2.56	0.94

As seen in Table I, the proposed approach achieves a significantly improved pose estimation, both in translation and rotation, compared to Face forward and Lock goal.

B. Experimental Evaluation

For experimental evaluation in water, we deploy the framework with the described approach on a BlueROV2 (seen in Fig. 3) [22]. The platform is equipped with a stereo-camera setup utilising the work described in [29], [30] to enhance the images before used with the SLAM. In addition the BlueROV2 is equipped with a DVL, IMU and a depth sensor which are fused with the SLAM [20]. We present results from two environments for the experimental evaluation; an indoor tank and a water-filled quarry. The maximum ray casting range is set to 3 metres.

During underwater operation, there is no access to ground truth or absolute positioning, hence, an evaluation of localisation error is not possible without specifically designed scenarios using markers or external measurements such as motion capture system (small region of operation) or acoustic systems (low position accuracy). Instead, to show the tracking capabilities we look at the number of matched features between consecutive camera frames. The maximum number of features for the SLAM to register in a single frame is set to 500. The matched features for the different environments can be seen in Table II for the tank and III for the quarry.

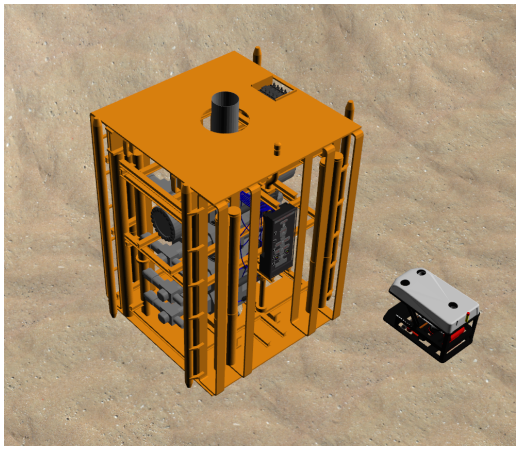
1) Scenario Tank: The first scenario for in water trials is an indoor tank (measuring approximately 12x12x3 metres) which has an artificial rectangular structure placed in it. We present the result from the following 3 scenarios that can be seen in Fig. 4a and are described below:

- **Fig. 4aA :** Traversing along the wall and back, and then partly around the structure. This scenario shows that while moving along a wall, the heading is slightly angled from the wall towards the direction of the next waypoint. Tests without using the score based on face next ($\mathcal{F}$) performed poorly on this scenario, as the robot maintained what was previously mapped in its view, hence preferring to looking more backwards than forward. Facing away from the direction of travelling increases the chances of colliding when operating in unknown environments. The structure had been mapped prior to the this execution, showing how the approach make use of previously seen structure where applicable.
- **Fig. 4aB**[1] **:** Moving to the other side of the artificial structure. The vehicle maintains the structure in view until the utility function's score to look forward is higher than maintaining the view of the mapped environment. In Table IIB, the matched features described are til the point where the distance between the robot and the structure is longer than the max distance for ray-casting. At this point, it can be seen in the figure how the proposed approach prefers to face towards the next waypoint, which here is in open water and hence no features to track.
- **Fig. 4aC :** Moving around the structure which is unknown prior to the execution. The robot maintains a good view of the structure throughout the whole operation and has, as can be seen in Table IIC, a high number of matched features.

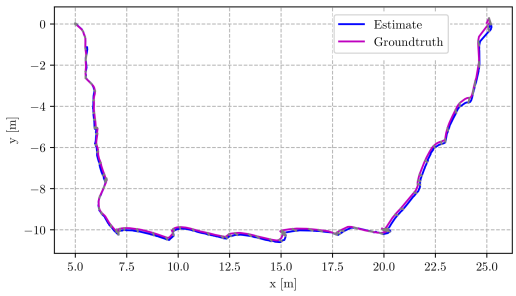
TABLE II: Features matched per frame in Tank

Scenario	min	average	max
Fig. 4aA	23	123	342
Fig. 4aB	25	145	263
Fig. 4aC	27	162	323

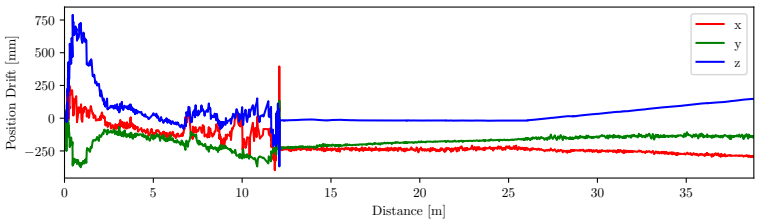
[1] https://youtu.be/33_gtpgdw60



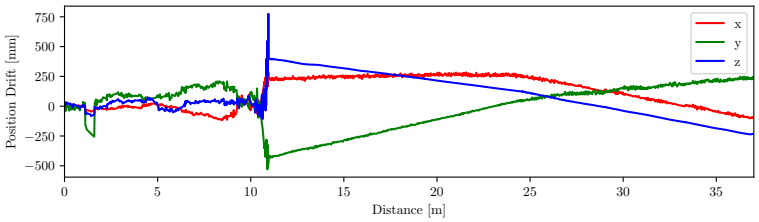
(a) Simulation environment



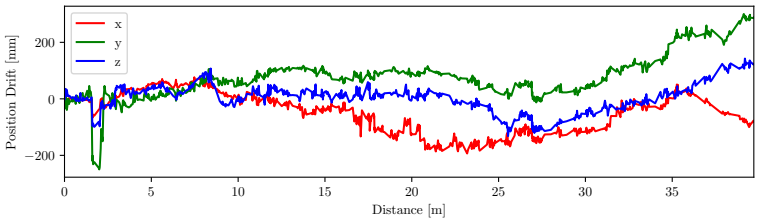
(b) Trajectory (proposed approach)



(c) Face forward



(d) Lock goal



(e) Proposed approach / Adaptive heading

Fig. 2: In (a) the simulated environment can be seen, where an AUV follows a path as in (b) around a blowout preventer, the wave pattern in the trajectory is due to non-optimised PID controller. The translation error relative to the ground truth of the compared approaches can be seen in (c)-(e). (c) : Face next waypoint/Face forward. (d) Lock goal : keeping the final waypoints heading. (e) The proposed perception aware approach. (c) and (d) loose features to track around 10-12 meters and are after this purely relying on dead reckoning (DR) leading to a worse pose estimation over time.

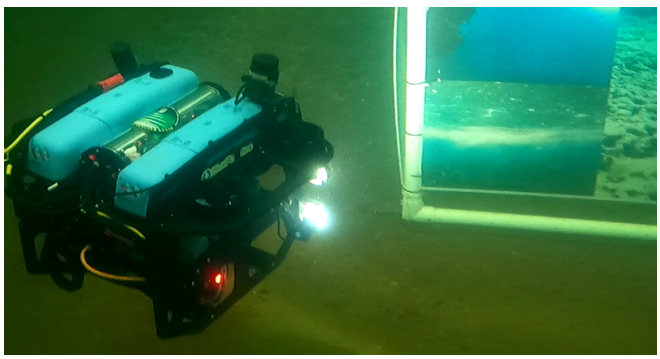
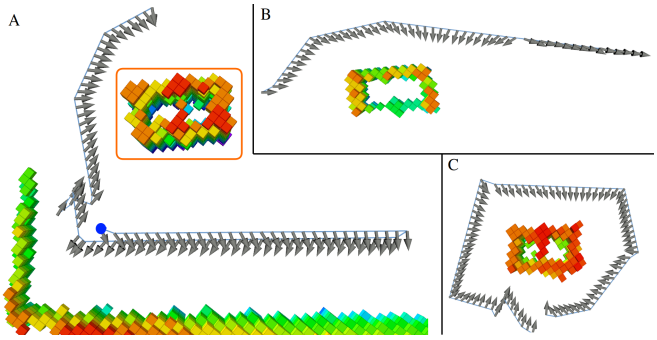


Fig. 3: For experimental evaluation, a BlueROV2, with a custom sensor and computational payload, is used in a tank with an artificial structure.

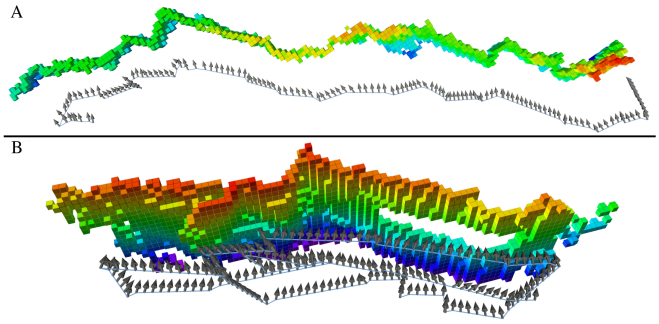
The experimental evaluation in the 3 different scenarios in the tank shows a heading adapting to maintain the view of the surroundings at all times. The matched features in Table II shows the vehicle is able to maintain tracking of features continuously throughout the operation. The exception is in Fig. 4aB where the vehicle moves away from the mapped environment, to open water, and hence the adaptive heading will prioritise to maintain a view in the direction of travel to increase the awareness of the area it is planning to move through and hence reduce the chance of colliding with objects that might exist in this region. The frames with a lower amount of matched features are generally due to relatively fast motions in yaw, increasing the motion blur.

2) Quarry: The second environment is an abandoned water-filled quarry. This demonstrates the adaptive heading in a larger-scale environment. Compared to the tank, the quarry does not have any freestanding obstacles to move around, instead, the walls are used. The two evaluated scenarios can be seen in Fig. 4b and are described below:

- **Fig. 4bA :** A 35 meters wide section where the robot follows the wall at a fixed depth. Similar to Scenario Tank A (Fig. 4aA) the vehicle keeps a heading slightly angled towards the direction of the next waypoint. The vehicle started at the right side of the image.



(a) Results from an indoor tank with an artificial structure. In scenario A), the structure (marked inside the orange rectangle) is known prior to the mission but the wall is not known, the robot started in at the blue circle. In B) and C) the environment is unknown prior to the start. The arrows shows the selected heading at each waypoint.



(b) Results from the quarry. In A) the robot follows the wall for 35 metres at the same depth. In B) The vehicle surveys a 4.5x16 metres region of the wall at different depths.

Fig. 4: Scenarios used with a BlueROV2.

- **Fig. 4bB[2] :** Similar section as in Quarry A, a 4.5 high and 16 metres wide. But instead of using a fixed depth, the robot varies to cover more of the environment in depth. For both Quarry A and B, the results are similar in terms of selected heading as in Fig. 4aA while following the walls.

The number of matched features can be seen in Table III. In Fig. 4bA, a total of 2 camera frames (total camera frames: 6584) did not contain enough matched features to perform pose estimation, and in Fig. 4bB 3 frames out of 7153 camera frames. A reason for not being able to perform feature matching on the few frames in these scenarios was due to rapid movement close to the wall causing motion blur, which can be corrected by decreasing the exposure time on the cameras.

TABLE III: Features matched per frame in Quarry

Scenario	average	max
Fig. 4bA	199	365
Fig. 4bB	225	370

[2]https://youtu.be/x2dezN7UaYw

The average and max matched features in the image frames in the quarry is higher than in the tank, a reason to this is that the difference in the orientation between consecutive frames for the AUV is rather low, as it is inspecting a wall. Less orientational motions means less motion blur occurred compared to the scenarios in the tank. In addition the field trials in the quarry was on a sunny day with good illumination and water conditions.

IV. Conclusion

This paper presents an approach for heading selection when a robot is following a trajectory. The selected heading has multiple benefits, such as observing prior seen parts of the environment, while taking into account the current state, next waypoint as well as preferring to face the normals of the sensed environment. We show in a simulated environment how the approach achieves a higher pose estimation accuracy than looking forward in the direction of travelling or maintaining the final orientation of a trajectory. In addition, we show data obtained from real scenarios in a controlled environment with an artificial structure as well as during inspections of a quarry wall. The work has been focused on the underwater domain, but has no restrictions based on this and can be used for terrestrial, aerial or other desired domains.

Future Work

The current approach incorporates a single sensor based on a camera. For future development, adding the option to include an arbitrary amount of sensors and of various types could improve the system's pose estimation and environmental awareness even further. The weights that determine which heading that is selected are currently manually specified, these should potentially be customised to the scenario, as such having an adaptive or learning approach for setting the weights depending on the environment and the sensor could improve the heading selection.

References

[1] R. Bajcsy, "Active Perception," *Proceedings of the IEEE*, vol. 76, no. 8, p. 9661005, 1988.

[2] G. Costante, C. Forster, J. Delmerico, P. Valigi, and D. Scaramuzza, "Perception-aware Path Planning," *arXiv:1605.04151*, pp. 1–16, 2016.

[3] G. Costante, J. Delmerico, M. Werlberger, P. Valigi, and D. Scaramuzza, "Exploiting photometric information for planning under uncertainty," *Springer Proceedings in Advanced Robotics*, vol. 2, no. December 2017, pp. 107–124, 2018.

[4] B. Zhou, J. Pan, F. Gao, and S. Shen, "RAPTOR: Robust and Perception-Aware Trajectory Replanning for Quadrotor Fast Flight," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1992–2009, 2021.

[5] D. Falanga, P. Foehn, P. Lu, and D. Scaramuzza, "PAMPC: Perception-Aware Model Predictive Control for Quadrotors," *IEEE International Conference on Intelligent Robots and Systems*, pp. 5200–5207, 2018.

[6] I. Spasojevic, V. Murali, and S. Karaman, "Perception-aware time optimal path parameterization for quadrotors," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 3213–3219, 2020.

[7] Z. Zhang and D. Scaramuzza, "Fisher Information Field: an Efficient and Differentiable Map for Perception-aware Planning," *arXiv preprint arXiv:2008.03324*, pp. 1–18, 2020. [Online]. Available: http://arxiv.org/abs/2008.03324

[8] P. R. Florence, J. Carter, J. Ware, and R. Tedrake, "NanoMap: Fast, Uncertainty-Aware Proximity Queries with Lazy Search over Local 3D Data," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. d, no. c, pp. 7631–7638, 2018.

[9] L. Quan, Z. Zhang, X. Zhong, C. Xu, and F. Gao, "EVA-Planner: Environmental Adaptive Quadrotor Planning," *IEEE International Conference on Robotics and Automation*, no. Icra, pp. 398–404, 11 2020. [Online]. Available: http://arxiv.org/abs/2011.04246

[10] E. Ochoa, N. Gracias, K. Istenic, R. Garcia, J. Bosch, and P. Cieslak, "Allowing untrained scientists to safely pilot ROVs: Early collision detection and avoidance using omnidirectional vision," *2020 Global Oceans 2020: Singapore - U.S. Gulf Coast*, 2020.

[11] E. H. Teniente, R. Valencia, and J. Andrade-Cetto, "Dense outdoor 3D mapping and navigation with Pose SLAM," *Workshop de Robótica Experimental*, pp. 567 – 572, 2012. [Online]. Available: http://upcommons.upc.edu/e-prints/handle/2117/15642

[12] R. Valencia, M. Morta, J. Andrade-Cetto, and J. M. Porta, "Planning reliable paths with pose SLAM," *IEEE Transactions on Robotics*, vol. 29, no. 4, pp. 1050–1059, 2013.

[13] N. Palomeras, M. Carreras, and J. Andrade-Cetto, "Active SLAM for autonomous underwater exploration," *Remote Sensing*, vol. 11, no. 23, pp. 1–19, 2019.

[14] E. Vidal, N. Palomeras, K. Istenič, N. Gracias, and M. Carreras, "Multisensor online 3D view planning for autonomous underwater exploration," *Journal of Field Robotics*, vol. 37, no. 6, pp. 1123–1147, 2020.

[15] J. S. Willners, Y. Carreno, S. Xu, T. Łuczyński, S. Katagiri, J. Roe, . Pairet, Y. Petillot, and S. Wang, "Robust Underwater SLAM using Autonomous Relocalisation," *IFAC-PapersOnLine*, vol. 54, no. 16, pp. 273–280, 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2405896321015068

[16] M. Xanthidis, M. Kalaitzakis, N. Karapetyan, J. Johnson, N. Vitzilaios, J. M. O'Kane, and I. Rekleitis, "AquaVis: A Perception-Aware Autonomous Navigation Framework for Underwater Vehicles," *IEEE IROS*, pp. 5410–5417, 2021. [Online]. Available: http://arxiv.org/abs/2110.01646

[17] M. Xanthidis, J. Johnson, J. M. O'kane, and I. Rekleitis, "Towards Multi-Camera Active Perception of Underwater Structures," *Proc. Advanced Marine Robotics Technical Committee Workshop on Active Perception at IEEE International Conference on Robotics and Automation Perception at IEEE International Conference on Robotics and Automation*, 2021.

[18] C. Stachniss, D. Hähnel, and W. Burgard, "Exploration with active loop-closing for FastSLAM," *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 2, pp. 1505–1510, 2004.

[19] S. M. Chaves, A. Kim, E. Galceran, and R. M. Eustice, "Opportunistic sampling-based active visual SLAM for underwater inspection," *Autonomous Robots*, vol. 40, no. 7, pp. 1245–1265, 2016.

[20] S. Xu, T. Luczynski, J. S. Willners, Z. Hong, K. Zhang, Y. R. Petillot, and S. Wang, "Underwater Visual Acoustic SLAM with Extrinsic Calibration," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2021.

[21] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, vol. 34, no. 3, pp. 189–206, 2013.

[22] J. S. Willners, I. Carlucho, T. Łuczynski, S. Katagiri, L. Chandler, J. Roe, D. Stephens, S. Xu, Y. Carreno, . Pairet, C. Barbalata, Y. Petillot, and S. Wang, "From market-ready ROVs to low-cost AUVs," *Oceans 2021*, pp. 1–7, 2021.

[23] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Ng, "ROS : an open-source Robot Operating System," *ICRA Workshop on Open Source Software*, no. Figure 1, 2009.

[24] J. D. Gammell, S. S. Srinivasa, and T. D. Barfoot, "Informed RRT*: Optimal sampling-based path planning focused via direct sampling of an admissible ellipsoidal heuristic," *IEEE International Conference on Intelligent Robots and Systems*, pp. 2997–3004, 2014.

[25] I. A. Sucan, M. Moll, and L. E. Kavraki, "The Open Motion Planning Library," *IEEE Robotics & Automation Magazine*, vol. 19, no. 4, pp. 72–82, 12 2012. [Online]. Available: https://ieeexplore.ieee.org/document/6377468/

[26] J. Pan, S. Chitta, and D. Manocha, "FCL: A general purpose library for collision and proximity queries," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 3859–3866, 2012.

[27] M. M. M. Manhães, S. A. Scherer, M. Voss, L. R. Douat, and T. Rauschenbach, "UUV Simulator: A Gazebo-based package for underwater intervention and multi-robot simulation," *OCEANS 2016 MTS/IEEE Monterey, OCE 2016*, 2016.

[28] Z. Zhang and D. Scaramuzza, "A Tutorial on Quantitative Trajectory Evaluation for Visual(-Inertial) Odometry," *IEEE International Conference on Intelligent Robots and Systems*, pp. 7244–7251, 2018.

[29] T. Łuczyński, M. Pfingsthorn, and A. Birk, "The Pinax-model for accurate and efficient refraction correction of underwater cameras in flat-pane housings," *Ocean Engineering*, vol. 133, no. March, pp. 9–22, 2017.

[30] T. Łuczyński, P. Łuczyński, L. Pehle, M. Wirsum, and A. Birk, "Model based design of a stereo vision system for intelligent deep-sea operations," *Measurement: Journal of the International Measurement Confederation*, vol. 144, pp. 298–310, 2019.