A Prototype Autonomous Robot for Underwater Crime Scene Investigation and Emergency Response

Zuhayr Rymansaib
University of Bath
Bath, UK
z.rymansaib@bath.ac.uk

Benjamin Thomas
University of Bath
Bath, UK

Alfred Anthony Treloar
University of Bath
Bath, UK

Benjamin Metcalfe
University of Bath
Bath, UK

Peter Wilson
University of Bath
Bath, UK

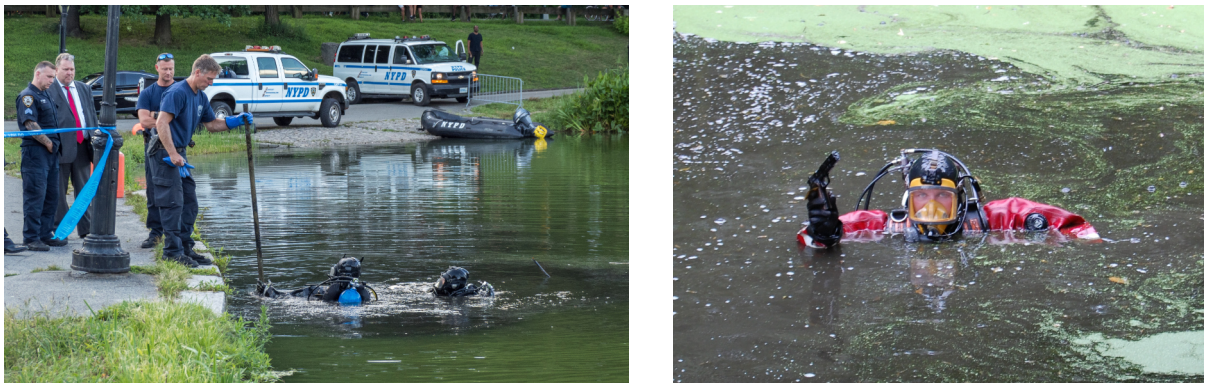
Alan Hunter
University of Bath
Bath, UK
a.j.hunter@bath.ac.uk

Abstract

Underwater crime scene investigation and emergency response are tasks typically carried out by divers constituting part of a specialist team. Operating in such dynamic environments, often with poor visibility and risk of concealed hazards, can be time consuming and dangerous. Autonomous uncrewed vessels with underwater acoustic imaging sensors have been used for similar purposes in other fields (e.g. hydrography, naval mine countermeasures, etc) but have not been adopted in this specific application domain.

The Police Robot for Inspection and Mapping of underwater Evidence (PRIME) is an autonomous uncrewed surface vessel (USV) that is being developed for this purpose. It is a novel application of existing robotic technology that is intended to be used within an end-to-end police and emergency underwater search process. It aims to enhance the effectiveness, efficiency, and safety of divers by autonomously locating and highlighting target objects or regions of interest, as well as benign regions, thereby reducing their time spent underwater. Side-scan imaging sonars are used to sense the underwater environment using techniques leveraged from the similar application domain of naval mine counter-measures. The system autonomously generates actionable intelligence in the form of simplified coverage and anomaly maps for easy interpretation by the dive team. These are communicated to shore in real-time and geo-referenced on satellite maps.

This paper details the PRIME system prototype and presents results from initial field experimentation. The prototype has been operated in various urban, shallow-water environments. The experimental results shown here were collected in Bristol Harbour (UK) with a water depth of approximately 5 m. In the experiment, a clothed mannequin resembling a human body was deployed on the muddy floor. Autonomous searches were executed and the body was detected successfully as an anomaly against the background, illustrating the feasibility and viability of the system as an autonomous robotic aid for locating missing persons in a representative, unstructured, and dynamic real-world environment.



(a) (b)

Figure 1: Examples of police underwater search operations and outcomes: (a) NYPD divers searching Harlem Meer following a murder in the area (McGrady, 2019); (b) officer from the North West Police Underwater Search & Marine Unit, UK, shown retrieving a discarded firearm (Hemans, 2012).

1 Introduction

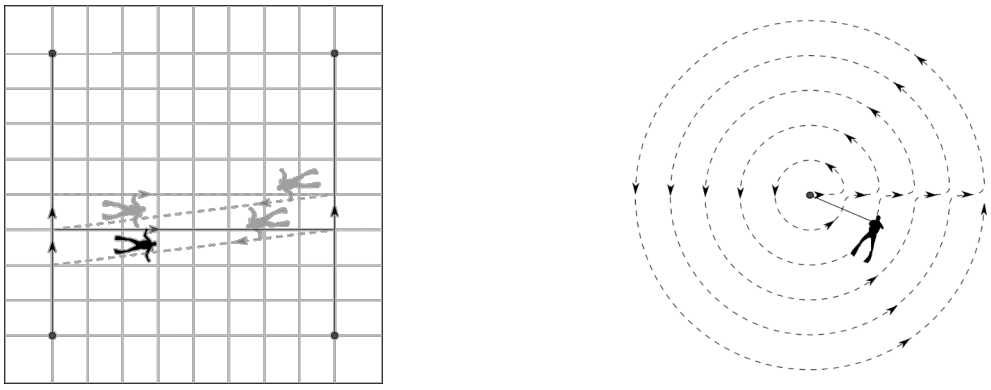
1.1 Underwater Search and Recovery

Police and emergency underwater search teams are generally comprised of highly skilled, specialist divers (Becker et al., 2013). In the UK, dive teams exist as part of civilian police forces, such as the Metropolitan Police Service (MPS) or North West Police Underwater Search & Marine Unit, as well as private groups such as Specialist Group International (SGI). Mission types vary, but typically involve: body recovery (e.g., in missing persons cases involving accidental or intentional death); finding and retrieving items of evidence (e.g., weapons or narcotics thrown into a body of water in an effort to be concealed or destroyed); and mitigating dangerous objects such as improvised explosive devices (IEDs). Some photographs from example cases are shown in Figure 1. The duration of a mission can vary from hours to weeks depending on the mission parameters and condition of the underwater environment. The environment is typically an inland waterway or lake with low to zero visibility and may be cluttered with obstructions or hazardous objects. Dive teams consist of specially trained members with a significant responsibility for the timely discovery, proper documentation, and recovery of any items considered to have evidentiary value (Erskine & Armstrong, 2021; Kelly, 2010; Wylie, 2019). Therefore, they are under very high demand due to their limited human and time resources. Searching for missing persons for example requires a considerable amount of human resources, with a single dive team consisting of a minimum of 3 to 5 personnel. Requirements vary depending on each case, and can sometimes require multiple teams, boats with associated crew, and cadaver dogs (Becker et al., 2013; Ruffell, 2014; Schultz et al., 2013).

1.2 Current Approaches and Problems

Modern robotic and acoustic imaging technologies exist that can aid dive teams in underwater search operations such as remotely operated vehicles (ROVs) scanning sonar devices, sub bottom profilers, and water penetrating radar (Becker et al., 2013; Decker, 2007; Erskine & Armstrong, 2021; Parker et al., 2010; Ruffell, 2014; Schultz et al., 2013). Towed side-scan sonar has been used to aid searches in missing persons cases (Erskine & Armstrong, 2021; Ruffell, 2014; Schultz et al., 2013). However, the use of such equipment can be troublesome. Firstly, tethered devices such as ROVs and towed sonar are challenging to operate due to the risk of entanglement in cluttered environments or with underwater foliage (Schultz et al., 2013), snagging when close to shore (Ruffell, 2014), as well as risk of damage to equipment or loss. Secondly, such devices require significant operator training and experience for optimal operation, as well as interpretation of the

gathered data, which can accumulate rapidly (Decker, 2007; Erskine & Armstrong, 2021; Schultz et al., 2013). In practice, this results in officers reverting to more familiar and trusted manual techniques involving tight, raster style search patterns such as those illustrated in Figure 2, which are very time consuming. These require that divers are tethered for systematic searching, as well as to ensure safety, with dives typically limited to durations of 15 to 20 minute shifts (Becker et al., 2013; Erskine & Armstrong, 2021; Ruffell, 2014). When employed appropriately, side-scan sonar is a highly effective tool in searching and locating of missing persons (Schultz et al., 2013). A requirement therefore exists for an autonomous system with the ability to rapidly survey underwater regions with minimal operator oversight, whilst automatically providing a simple interpretation of the collected data, thereby improving performance and safety.



(a) Linear (jackstay) search pattern (b) Circular search pattern

Figure 2: Illustrations of manual search patterns (Southwood, 2011). (a) A jackstay search pattern allows one or more divers to search large areas thoroughly and efficiently. The divers traverse a search line (jackstay) rigged on the bottom (left-right). When the diver(s) reach the end of the line, it is advanced by a certain distance (up-down) and the diver(s) search back along the line. (b) A circular search pattern can be efficient when the position of a target is approximately known. Diver(s) search in concentric circles around a central point where a rope is fixed. A constant radius is achieved by maintaining tension on the rope; the radius is extended by a certain distance each revolution by increasing the length of the line.

1.3 Current Research

Research on autonomous human body detection using sonar is limited, but similarities can be drawn with naval mine counter-measures (MCM) operations. Target features in sonar images can be enhanced with image processing techniques such as wavelet filtering (Hunter & van Vossen, 2014), whilst seafloor characterisation and complexity mapping are used to evaluate performance of automatic target recognition (ATR) algorithms (Fakiris et al., 2013; Geilhufe & Midtgaard, 2014; Williams, 2015). Effective scanning of a given area can be enhanced by using appropriate search patterns (Hunter et al., 2018), and scanning multiple times from different orientations can improve the detection and classification performance of ATR algorithms (J. Fawcett et al., 2010; Zerr et al., 1997). Although objects being searched for in MCM operations are typically sound-hard scatterers as opposed to a relatively soft human body, human bone has a similar reflective coefficient to sandstone and concrete (Erskine & Armstrong, 2021), allowing for detection using sonar based approaches. Autonomous detection of a mannequin, used as a proxy for a human target, has been demonstrated using a convolutional neural network (CNN) trained with multibeam sonar images (Nguyen et al., 2019) but these were gathered manually and processed offline.

Commercial and research platforms employing low cost, commercially available sonar devices have been used to accurately survey shallow water marine environments using towed or hull-mounted side-scan or multibeam sonar (Kaeser et al., 2013; Kebkal et al., 2014). Uncrewed surface vessels (USV) for autonomous long term, or large scale exploration operations have been developed based around vessels ranging from basic catamaran assemblies (Girdhar et al., 2011) to instrumentation of commercially available kayaks (Moulton et al., 2018).

1.4 System Requirements and Proposed Solution

There are both cost and operational barriers that impede the uptake of modern underwater acoustic imaging technologies. This has motivated the desire for a low-cost, untethered, and integrated system for aiding police divers during underwater crime scene and accident investigations that can operate autonomously and requires minimal operator training and intervention. It is further motivated by the potential to increase diver safety as a by-product of reducing their time in the water and by giving them prior intelligence on underwater hazards. The University of Bath has been working together with representatives within the UK police and emergency services to develop and deliver such a system.

The system requirements and use cases were determined following consultation with stakeholders. The key requirements are:

1. Effective and Safe – The system should enhance the search capabilities of the divers without causing interference or harm;
2. Simple – It should gather and interpret data autonomously before presenting meaningful and easily understandable information to the operator without need for specialist training;
3. Compact – It should fit easily inside a standard-issue police 4×4 sports utility vehicle (e.g., Mitsubishi Shogun) and be deployable by no more than two personnel;
4. Low Cost – It should cost roughly an order of magnitude less than existing ocean-grade solutions used by the commercial and military sectors, i.e., on the order of tens of thousands of pounds.

The system is intended to be used at times when a dive team is on site but would not normally enter the water. Possible time windows include during mission briefing and preparation (approximately 1 hour) and outside regular working hours (approximately 8 hours overnight). The first use case for consideration is the search for objects that resemble human bodies. In this context, resemblance relates mostly to the approximate dimensions of an object, since a body could be in any pose or it could be concealed. Moreover, it was identified as the simplest first step due to the large size of the target object in comparison to those in other use cases (i.e., weapons, packages, or IEDs).

Our prototype solution is the Police Robot for Inspection and Mapping of underwater Evidence (PRIME). The system hardware is based around a custom-built USV with off-the-shelf electric propulsion, sonar sensing, navigation, and computing hardware. The software is based upon the Robotics Operating System (ROS) (Quigley et al., 2009) using a mixture of existing open-source packages and custom-built nodes. PRIME can autonomously execute search patterns, acquire and process data, and communicate its gathered intelligence on the underwater environment to a computer on shore in real-time. It cannot currently avoid obstacles, but has a manual remote-control override for safety. The information communicated to shore is in the form of a simplified “heat map” that does not require significant training or experience to interpret. A satellite map of the surveyed area is overlaid with the cumulative sonar coverage area using a simple two-colour scale indicating regions in blue that are benign and featureless versus areas in red where the features could indicate the presence of a body. The originality lies in the real-time, autonomous generation of this actionable intelligence to aid human dive teams in missing persons scenarios, by indicating where to prioritise efforts during a search.

This paper presents the design, implementation, field deployment, and testing of the PRIME system prototype. The interested reader can find comprehensive descriptions of the design and implementation details, covering the physical platform, sensing hardware, and electronics in Section 2 and control, perception, and autonomy algorithms and software in Section 3. These technical sections can be skipped without impacting on the interpretation of results from the field experiments, which are presented in Section 4. The outcomes are discussed in Section 5 and a final summary is provided in Section 6.

2 Hardware Design

2.1 Platform

A USV was selected as the robotic platform in preference over an uncrewed underwater vehicle (UUV). This is a well established autonomous vessel configuration, with the design developments and challenges well documented, (Z. Liu et al., 2016; Manley, 2008) and has been chosen for several reasons. A USV operating from the surface has easy accessto Wi-Fi and global navigation satellite systems (GNSS), which greatly simplifies communication and navigation. It is visible to the dive team and users of the waterway at all times and vice versa. It is simpler to deploy and recover, it does not need complicated trimming and ballasting, and requires lower standards of waterproofing.

The PRIME USV is custom-built and has been designed and developed over multiple iterations of rapid prototyping. It has a catamaran hull with dimensionsof approximately 1.2m in length, 0.6m in width, and 0.4m in height. The total weight of the vehicle and its typical payloads is around 20kg. Earlier iterations (PRIME-1, 2, and 3) explored variations on the hull design. These were constructed from pairs of foam pontoons, reinforced with coatings of epoxy and fibreglass. Aluminium sheet and extrusions were used to assemblethe two pontoons together and to provide a frame for attaching further hardware. Two T200 DC brushless thrusters (Blue Robotics Inc. (USA), 2021b) are mounted at the rear of the USV for propulsion. The differential drive provides good maneuverability and a maximum speed in still water of approximately 3 m/s.

The latest design iteration, PRIME-4, is shown in Figure 3 and the previous iteration, PRIME-3, is shown in Figure 10a. PRIME-4 has a moulded catamaran hull constructed from carbon fibre (C12 Composites Ltd. (UK), 2020). It travels at an operating speedof approximately 1.2 m/s. It also features hinged components to allow for hardware that is submerged during operation, such as the sensorsand thrusters, to be raised for easy transport and storage.

Modular payloads can be mounted to the aluminium frame, with the electronics and batteries being housed in IP68-rated waterproof enclosures. This allows flexibility for experimenting with different operational concepts. Although commercial USVs with similar capabilities readily exist, (Kebkal et al., 2014) a custom USV was built in the interest of research flexibility and cost saving.

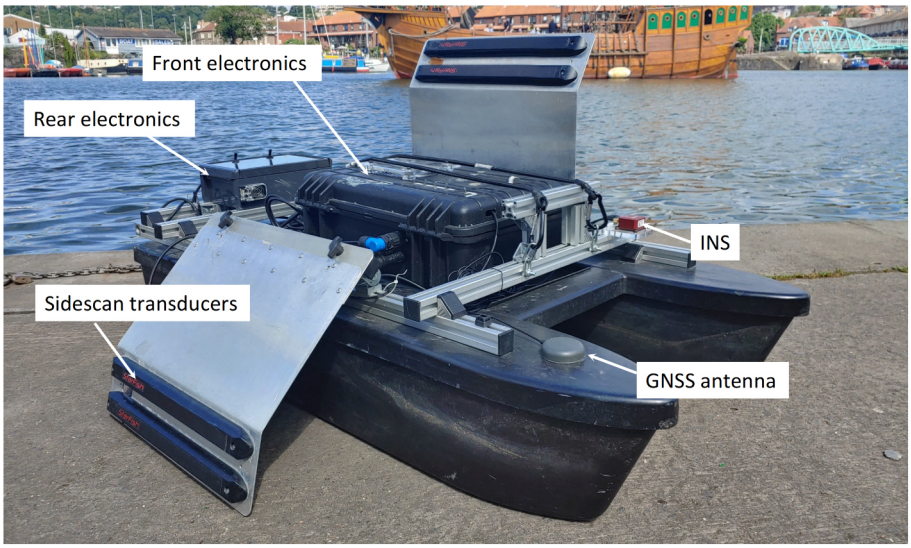
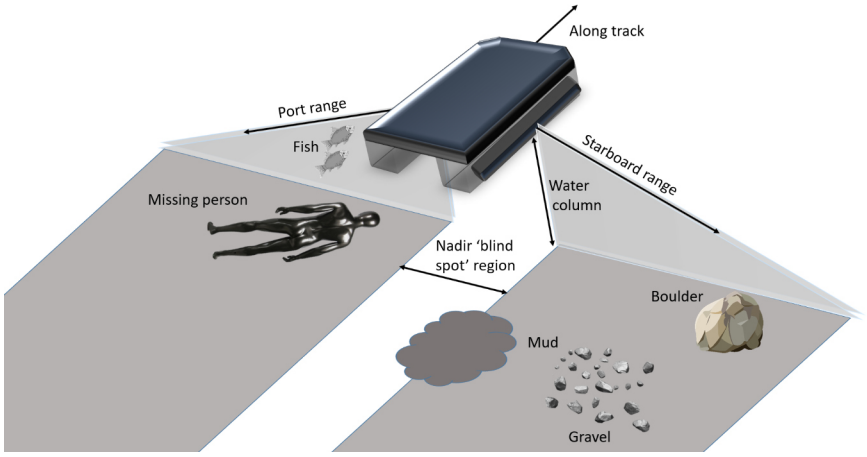
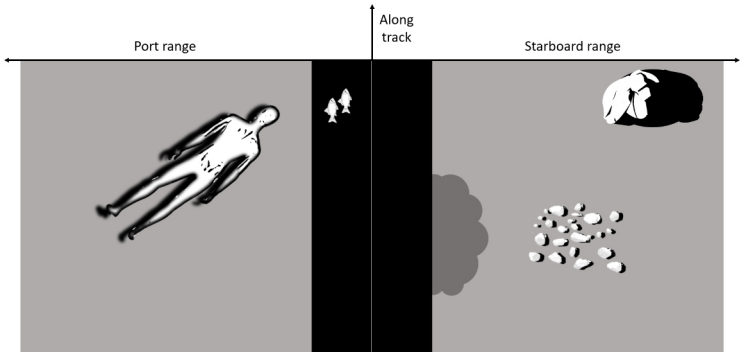


Figure 3: PRIME-4 USV with carbon-fibre hull, showing the sonar transducers partially deployed.



(a) Imaging geometry



(b) Raster image

Figure 4: A side-scan sonar makes acoustic reflectivity measurements from fan-shaped beams as the platform moves along a track. These are stacked to form a 2-D raster image. Changes in the composition of the floor appear as variations in the image intensity (e.g., stronger for hard materials like gravel and weaker for soft materials like mud). A low intensity strip at close range corresponds to the water column between the sensor and floor, including objects in the water column (e.g., fish). Objects on the floor appear as highlights followed by acoustic shadows and objects cast longer shadows with range (e.g., the boulder).

2.2 Underwater Acoustic Sensing

Side-scan sonar is one of the most commonly used underwater imaging technologies (Blondel, 2009). It was selected for use on PRIME due to its rapid area coverage rate and its cost advantage over more complicated multi-beam systems. It is intended for the initial wide-area survey of the underwater environment to expose areas of interest for the human dive team and / or to inform autonomous re-interrogation using other sonar types or other sensing modalities (e.g., optical, chemical, magnetic, or tactile) in the future.

Side-scan sonar operates by projecting an acoustic pulse or "ping" to the port and starboard sides of a platform at regular intervals as it moves along a nominally straight track at constant velocity. The acoustic beams are narrow in the direction of travel and wide in the vertical direction, as illustrated in Figure 4a. Thus, each ping measures the intensity of acoustic reflections within the beam cross-section as a function of acoustic travel time. These 1-D measurements are stacked to generate a 2-D raster image that is an orthographic projection of the 3-D underwater environment. The axes of the image correspond to the port and starboard ranges versus the position along the track, determined from the travel times and the known speed of sound in water $c \approx 1450 \mathrm{m/s}$. In the range axis, the image resolution is constant and is determined by the bandwidth of the acoustic signal. In the along-track axis, the resolution degrades with range and is determined by the horizontal beam-width. The vertical beam-width and declination of the transducers determines the observable swath of floor and, typically, this leads to a blind spot directly below the vehicle termed the nadir gap. In practice this gap can be filled using another sensor (e.g., a downward or forward-looking sonar or camera) or by conducting surveys with overlapping coverage (Hunter et al., 2018). The image formation process is illustrated in Figure 4.

Side-scan sonar images have the appearance of a top-down view of the floor with side illumination from the track towards the port and starboard directions. The nadir gap manifests as a strip of low reflectivity in the centre of the image, corresponding to the water volume between the vehicle and the first observable range to the floor. A particular characteristic of side-scan sonar images is that an object sitting proud on the floor appears as a highlight due to the reflection from the object, followed by an acoustic shadow cast onto the floor behind. Importantly, the orthographic projection preserves the geometrical dimensions of the object. These image features are useful for object recognition.

PRIME is equipped with two side-scan sonars from Blueprint Subsea (UK): a Starfish-450 and a Starfish-990. These have centre frequencies of 450kHz and 990kHz and horizontal beamwidths of 0.5° and 0.3°, respectively. Both have vertical beamwidths of 60°. The lower frequency suffers less from attenuation and can propagate further with a maximum range of 100m at 450kHz compared to 35m at 990kHz. However, the higher frequency yields finer along-track resolution. The (constant) range resolution is $\Delta r = 7.5\mathrm{cm}$ in both bands, but the (range-varying) along-track resolution is given by $\Delta u_{450} \approx 0.0087 \times r$, at 450kHz compared to $\Delta u_{990} \approx 0.0052 \times r$ at 990kHz. The two sensors have complementary characteristics. Furthermore, there is future potential to use spectral differences between the images from each band to aid pattern recognition. During operation, the 450kHz and 990kHz transducers are deployed at depths of 20cm and 15cm below the water surface, respectively, and at a declination angle of 30°.

2.3 Navigation Sensing and Communications

PRIME uses an SBG Ellipse inertial navigation unit (France) with an IMU and external GNSS antenna, both of which are mounted externally, as shown in Figure 3. This gives a positioning accuracy to within 2m and provides accurate timing. The IMU produces pitch and roll measurements, but its main purpose on PRIME is to measure heading using its magnetic compass with accelerometer and gyroscope aiding. The heading accuracy is within 0.8°, but must calibrated with the platform fully powered and running to account for any electromagnetic interference (EMI) from platform electronics.

A Wi-Fi hub and antenna on the shore is used to facilitate communication between all of the computers in the system – both onboard the USV and on the shore for monitoring and control.

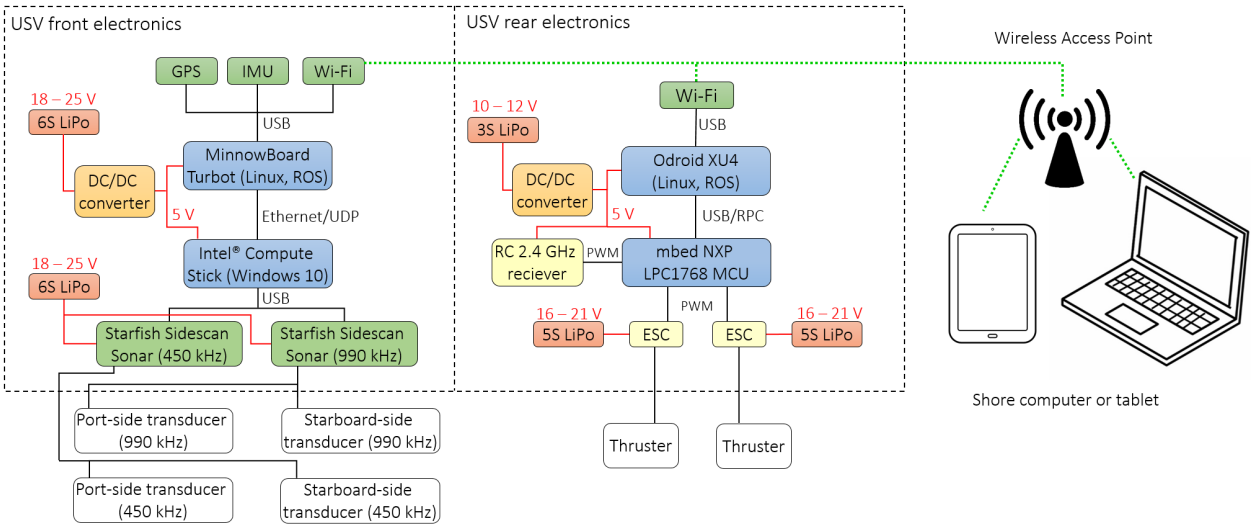


Figure 5: PRIME-4 hardware architecture.

2.4 Electronics and Computing

The hardware architecture is illustrated in Figure 5. The rear electronics box houses the control hardware and the front box houses the sensing and autonomy hardware.

The rear box contains two electronic speed controllers (ESC) (Blue Robotics Inc. (USA), 2021a) for driving the thrusters. These are interfaced through an Arm Mbed NXP LPC1768 microcontroller. The microcontroller arbitrates thruster control signals from either a 2.4 GHz remote-control receiver during emergency override by a human operator or via remote procedure calls over USB from an Odroid XU4 single board computer (SBC) during normal autonomous operation. The Odroid runs a Linux operating system (Ubuntu MATE 18.04) and acts as a bridge to the computing hardware in the front box over Wi-Fi. The reason for this separation is to electrically isolate the sensitive data acquisition hardware in the front box from the noisy thruster control hardware in the rear box, which was found to cause problematic interference during development of the system.

The front box contains the sonar data acquisition hardware and two SBCs. At the time of this work, software drivers for the Starfish side-scan sonars were only available for the Microsoft Windows operating system and this necessitated the use of an Intel Compute Stick SBC running Microsoft Windows 10. The Compute Stick interfaces with the Starfish hardware via USB and broadcasts the sonar data and settings via the user datagram protocol (UDP) over ethernet. A MinnowBoard Turbot Quad-Core SBC with Ubuntu Linux runs the core autonomy software. It executes the autonomous control of the USV, sending forward/backward thrust and left/right yaw velocity demands to the Odroid in the rear box. It also performs the sonar control and data processing and communication of status and data products to the computers on the shore.

The current prototype uses five lithium-polymer (LiPo) batteries for separately supplying the two thrusters (21 V fully charged for a specified operating range of 7 V to 20V), the two sonar data acquisition units (25 V for a range of 9 V to 28V), and the SBCs in the front and rear boxes (25 V and 12 V, respectively, converted to 5V).

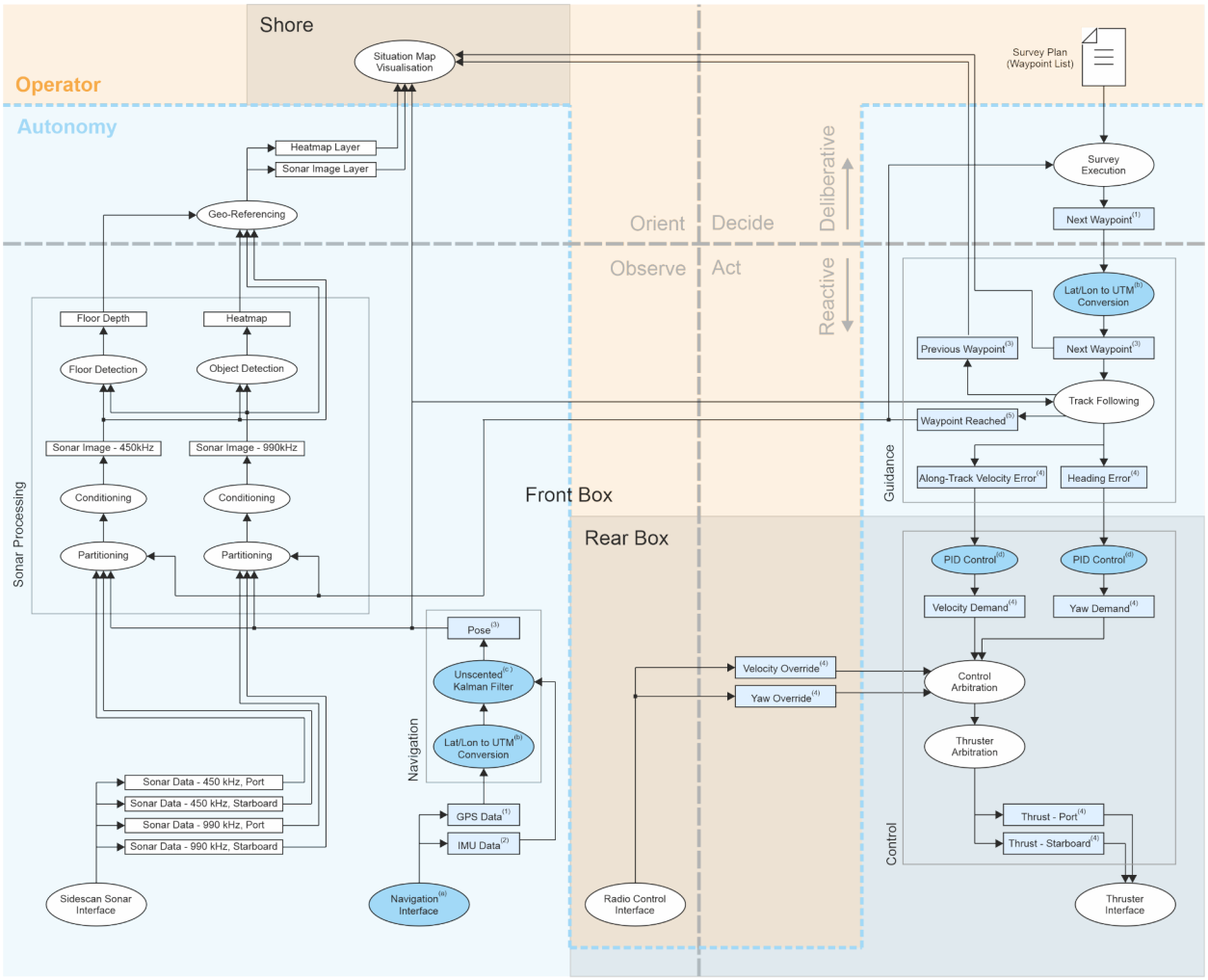


Figure 6: PRIME-4 software architecture, represented as a simplified ROS node graph. Ellipses represent nodes, rectangles represent topics, and arrows represent the publishing and subscribing relations; the nodes and topics in white are custom-built whereas the others are standard or from existing packages. Note that in this early prototype there is closed-loop reactive feedback and control, but the deliberative control is open-loop.

3 Software Design

PRIME uses the Robot Operating System (ROS) (Quigley et al., 2009). ROS runs distributed over Wi-Fi across the two SBCs on the USV as well as any number of monitoring and control computers on the shore. The majority of the ROS nodes run on the MinnowBoard SBC, with only a bridging node for relaying thruster demand signals running on the Odroid, and monitoring and control nodes on the shore. A simplified node graph is shown in Figure 6. The architecture consists of a combination of existing and custom built packages.

3.1 Autonomy

A basic level of autonomy has been implemented that allows the USV to conduct a pre-defined survey pattern, whilst simultaneously processing the sonar data and reporting the mission status and underwater intelligence to the shore in real-time.

The design is based on a layered hybrid deliberative / reactive architecture (Gat et al., 1997), with behaviours organised into the four broad categories of the observe-orient-decide-act (OODA) model (Boyd, 1987; Proud et al., 2003). In the reactive layer of the architecture, the USV acquires data from its sensors and processes these into useful data products. It uses some of these products directly for reactive control and some for higher-level perception. In the deliberative layer, the data products are fused into a useful *world model*. The represents the robot's situational awareness of itself and the environment, and is used to inform the deliberative planning and execution of the mission goals. Components of the world model relating to the mission status and, importantly, the underwater intelligence are communicated to human operators on shore.

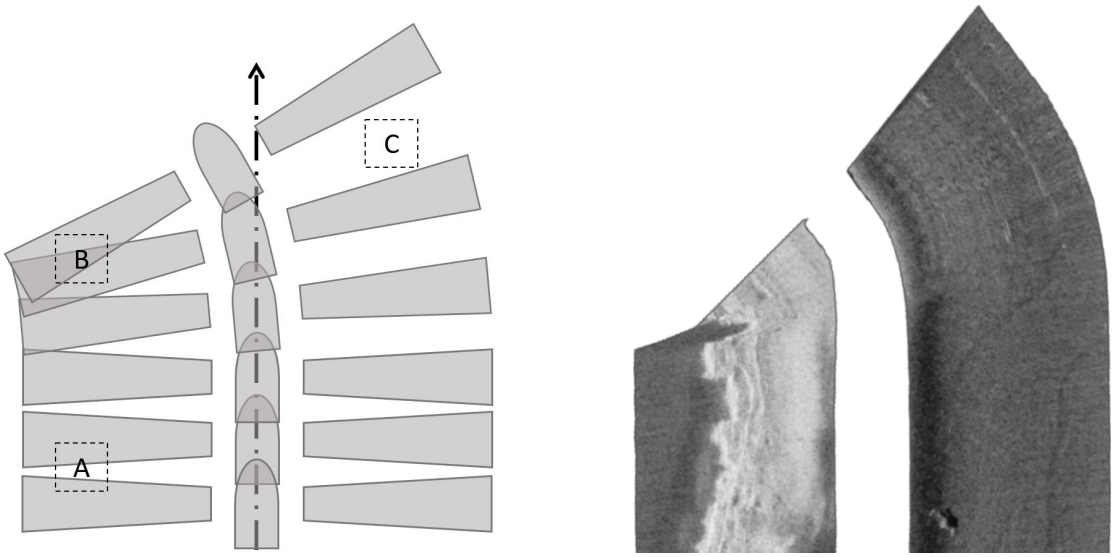
In the current implementation, the reactive layer contains behaviours relating to navigation, guidance, and thruster control for executing straight survey lines between waypoints and the acquisition, processing, and automated interpretation of sonar data. These are detailed in Sections 3.2 to 3.4. The deliberative layer contains a simple path planner with pre-defined survey waypoints and a basic world model comprised of a situation map with geo-referenced overlays for the various sonar data products, including the heat map for the human dive team. These are detailed in Section 3.6 and 3.7, respectively.

3.2 Guidance, Control, and Navigation

To ensure the collection of good-quality side-scan sonar data, the platform must travel in straight lines at a consistent speed. This speed depends mainly on the ping repetition frequency of the sonar, which varies depending on several factors such as water depth, maximum range, and reverberation from the environment (Blondel, 2009). Uniformity of the path ensures that the raster image is undistorted and slowness ensures that the acoustic beams from subsequent pings adequately sample the floor without leaving gaps. However, travelling too slowly makes control of the platform more difficult, particularly in currents or cross-winds. An illustrative example is shown in Figure 7 for a straight path followed by a turn, where the sonar image generated along the straight portion is better quality than the distorted and under-sampled portion during the turn.

The two thrusters mounted at the rear of the USV provide a differential drive. Activation of both thrusters equally in the same direction produces forward or reverse motion (surge) while dissimilar activation results in a change in orientation (yaw). The USV has an additional degree of freedom in the direction perpendicular to the vehicle heading (sway). This results in an under-actuated control system with more degrees of freedom than control inputs. Various line-of-sight (LOS) control algorithms have been developed for this problem (Chaos et al., 2009; Furfaro, 2012; T. Liu et al., 2017; Shin et al., 2017).

The guidance system achieves straight survey tracks using the LOS method of (Furfaro, 2012). Figure 8 illustrates the geometry for the LOS algorithm. The location of the USV is denoted by x_p and an ideal



(a) (b)

Figure 7: The effect of path linearity on side-scansonar image quality. During the straight portion of the track (A), the sonar beamsproduce a uniform raster coverageof the floor. However, during a turn the beams are bunched and spread on the inside and outside of the turn, respectively, leading to image distortion (B and C) and under-sampling (C).

straight path is defined between waypoints x_a and x_b. The perpendicular distance of the USV from the ideal path is given by

$$d = \left\| (x_a - x_p) - (x_a - x_p) \cdot \frac{x_a - x_b}{\| x_a - x_b \|} \frac{x_a - x_b}{\| x_a - x_b \|} \right\| , \tag{1}$$

and the closest point on the ideal path to the USV is given by

$$x_\perp = x_a + (x_p - x_a) \cdot (x_b - x_a) . \tag{2}$$

A *lookahead point*

$$x' = x_\perp + h \frac{x_b - x_a}{\| x_b - x_a \|}, \tag{3}$$

is introduced on the path at a distance h ahead of $x_\perp$, which advancesalong the ideal track as the USV progresses. A LOS vector

$$\hat{n}' = \frac{x' - x_p}{\| x' - x_p \|}, \tag{4}$$

is defined to provide a target heading for the USV. The heading error

$$\beta = \left\| \frac{\hat{n}' \times \hat{n}}{\| \hat{n}' \times \hat{n} \|} \cdot \hat{z} \right\| \arctan 2(\| \hat{n}' \times \hat{n} \|, \hat{n}' \cdot \hat{n}) \tag{5}$$

is established between the USV heading θ and the target heading θ', where arctan 2(y, x) is the four-quadrant inverse tangent function. The speederror

$$v' = v - \dot{x}_p \cdot \frac{x_b - x_a}{\| x_b - x_a \|} \tag{6}$$

is established as the difference between the desired target velocity v and the component of the USV velocity along the ideal path. The heading error β and speederror v' are inputs to proportional, integral, derivative

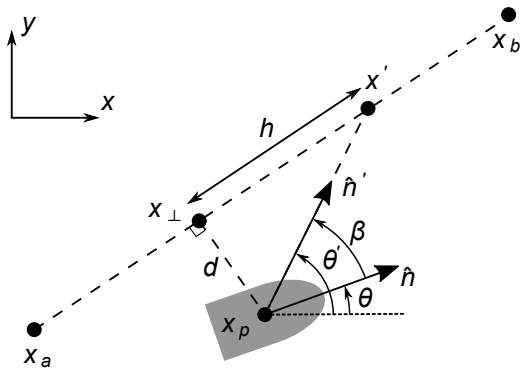


Figure 8: Geometry of the line-of-sight (LOS) control algorithm used by the USV to traverse straight tracks between waypoints.

(PID) controllers which operate the thruster inputs and hence control the USV speed and heading with forward/backward thrust and left/right yaw velocity demands.

Two PID controllers are used to control thrust and yaw independently. These were implemented from the open-source ROS Control framework (Chitta et al., 2017), as PI controllers using the following gain values for thrust, $K_{T,P} = 0.5, K_{T,I} = 0.2$, and yaw, $K_{Y,P} = 3.0, K_{Y,I} = 0.1$. The gains were selected through use of the autotune algorithm available within the ROS PID package, which is based on the Ziegler-Nichols method (Ellis, 2012). Autotuning was carried out whilst executing a demand trajectory along a 5m x 5m square. The process was repeated on different bodies of water, including a canal stream, a still lake in both calm and windy conditions, and a working harbour with passing traffic. The values obtained were then manually averaged to provide a common set of gain parameters across all environments tested. The output range for both controllers is normalised in the range −1 to 1.

As the USV moves along its path the location of the lookahead point x' advances, and the USV is guided as if being pulled towards the sliding lookahead point. The distance d can be adjusted by the choice of lookahead distance h, with a shorter lookahead reducing d and hence the path following error. A lookahead distance of $h = 2$ m was found to give good performance for path tracking in the environments tested.

In situations when the heading error is large, for example immediately after a waypoint has been reached, it is advantageous to prioritise heading error reduction over forward velocity. This is achieved by disabling the thrust controller and halving the yaw controller gains when the magnitude of the heading error β exceeds a user defined value σ, that is,

$$K'_T, K'_Y = \begin{cases} K_T, K_Y, & \|\beta\| \leq \sigma \\ 0, \frac{K_Y}{2}, & \|\beta\| > \sigma \end{cases}. \tag{7}$$

A value of $\sigma = 30^\circ$ was found to give good performance.

The PID controller outputs are fed to a thruster arbitration algorithm, which maps the thrust and yaw velocity demands to the port and starboard thruster ESC commands,

$$\begin{aligned} T_{\text{port}} &= \frac{\sqrt{a^2 + b^2}}{\sqrt{2}} \cdot \cos\left[\arctan 2(b,a) + \frac{\pi}{4}\right] \\ T_{\text{stbd}} &= \frac{\sqrt{a^2 + b^2}}{\sqrt{2}} \cdot \sin\left[\arctan 2(b,a) + \frac{\pi}{4}\right], \qquad a,b \in [-1,1] \end{aligned} \tag{8}$$

which are functions of the orthogonal yaw and thrust velocity demands, represented as a and b, respectively, and normalised between −1 to 1. The same algorithm is used to arbitrate control commands from the

manual override. This algorithm is based on a method for mapping orthogonal commands such as those from a joystick input to a differential drive, or skid-steer system, commonly used on tracked vehicles such as tanks (Taylor, 2010). It improves over basic addition/summation methods which can cause thrust commands of one side to saturate when turning, thus limiting the maximum yaw rate.

Navigation is necessary for providing feedback to the guidance and control system. For this purpose, state estimation of the USV's 2-D location and velocity is achieved using an unscented Kalman filter (UKF) with fusion of input from the GNSS, INS, and regularisation with an assumed motion model. A UKF implementation has been used from the open-source robot localisation package for ROS, which in general is more stable and accurate than an Extended Kalman Filter (EKF) (Moore & Stouch, 2016; Wan & Van Der Merwe, 2000). Other localisation algorithms exist, but these have not been evaluated (Dhariwal & Sukhatme, 2007).

3.3 Side-Scan Sonar Image Processing and Analysis

The side-scan sonars produce images from each track in two frequency bands on the port and starboard sides of the vehicle, i.e., 4 images per track. These are processed and analysed to generate simplified heat maps, which are fused over multiple tracks and geo-referenced into a world model for the situational awareness of the USV and human operators.

The reader is referred to Section 4.1 of the results to find representative examples of the sonar images, which will aid in understanding the sonar image processing and perception algorithms described in the following sections.

3.3.1 Image Conditioning

Raw side-scan images are typically generated with 8-bit greyscale pixel intensity values, where echo strength is represented in the range 0–255. Equivalently, this can be represented on a normalised scale of 0–1. Use of the entire range is rarely achieved, as the minimum value corresponds to the absence of any signal and the maximum value corresponds to saturation or clipping. This tends to result in a low contrast image, with pixel intensities not making full use of the available dynamic range. Contrast stretching has been used to expand this range thus enhancing image characteristics (Yang, 2006), as the initial stage of the image processing chain.

Consider a raw greyscale image

$$i_{\text{raw}}(u, r) \in [0, 1], \tag{9}$$

where u and r are along track and range directions, respectively, and pixel intensity values have been normalised between 0 and 1. The image after contrast stretching and normalisation is given by

$$i(u, r) = \begin{cases} 0, & i'(u, r) < 0 \\ 1, & i(' u, r) > 1 \\ i'(u, r) & \text{elsewhere,} \end{cases} \tag{10}$$

where

$$i'(u, r) = (i_{\text{raw}}(u, r) - I_{\text{L}} + I_0) / (I_{\text{U}} - I_{\text{L}}) . \tag{11}$$

The distribution of intensity values is shifted by I_0, rescaled between the lower and upper bounds I_{L} and I_{U}, and clipped to remain within the range 0–1. The shift is selected so that the expected median value of the resulting image lies in the middle of the range. Using lower and upper bounds corresponding to the 15th and 99th percentiles of the expected raw intensity values was found to provide good contrast with sufficient detail in the highlights, and good depth in the water column and shadows. It is common in some implementations of contrast stretching to compute the parameters dynamically from the histogram of raw pixel intensities (Young et al., 1998). However, fixed values of $I_0 = 0.2$, $I_{\text{L}} = 0.4$, and $I_{\text{U}} = 0.75$ have been used to ensure consistency when fusing results collected from multiple tracks.

3.4 Object Detection

For automated detection of targets, an approach based on 2D wavelet filtering and complexity mapping (Geilhufe & Midtgaard, 2014) is used. Wavelet filtering is applied to emphasise objects with scales that are consistent with the target body-like objects. The wavelet filtered images are then simplified using an image contrast metric to quantify the feature complexity at these scales. Regions of high complexity are thereby associated with high likelihood for the presence of target objects (Fakiris et al., 2013; Williams, 2015).

3.4.1 Wavelet Filtering

The images can be represented by their wavelet decompositions,

$$I(u, r, \mathbf{s}) = W\{i(u, r)\} \quad . \tag{12}$$

The operator $W\{\cdot\}$ performs the wavelet transform using a chosen mother wavelet function and produces a multi-resolution set of sub-images. These contain wavelet coefficients (weightings) corresponding to resampled image coordinates and are organised according to a hierarchy of wavelet scales and orientations,

$$\mathbf{s} = \left\{ \mathbf{s}^{(1)}, \mathbf{s}^{(2)}, \cdots, \mathbf{s}^{(N)} \right\}, \tag{13}$$

where

$$\mathbf{s}^{(n)} = \left\{ s_A^{(n)}, s_H^{(n)}, s_V^{(n)}, s_D^{(n)} \right\} . \tag{14}$$

The superscript $n \in [1, N]$ denotes the integer scale level, and the maximum level N (corresponding to the largest scales and lowest spatial frequencies) is limited by the number of pixels in the image and the chosen wavelet. The subscript $m \in \{A, H, V, D\}$ denotes the approximation (low-pass) coefficient and detail (high-pass) coefficients for the horizontal, vertical, and diagonal orientations, respectively (S. G. Mallat, 1989).

Wavelet filtering can be applied to retain only coefficients at scales consistent with the expected dimensions of the target object. Thus, a wavelet filtered image is obtained by applying a scale-dependent window function, followed by the inverse wavelet transform,

$$\tilde{i}(u, r) = W^{-1}\{I(u, r, \mathbf{s}) \cdot \hat{W}(\mathbf{s})\} \quad , \tag{15}$$

where $\hat{W}(\mathbf{s})$ is the window function.

A rectangular “band-pass” window,

$$\hat{W}(\mathbf{s}) = \begin{cases} 1, & n_1 \le n \le n_2 \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

is used to retain the scales between levels n_1 and n_2. The choice of wavelet function is not critical (F. Adamo et al., 2013), but the Daubechies 5-tap wavelet was found to provide satisfactory emphasis of the target objects in the side-scan imagery with cut-off values of $n_1 = 5$ and $n_2 = 7$. Wavelet transforms have been implemented using the PyWavelets library (Lee et al., 2019).

3.4.2 Image Complexity (“Heat”) Map

Regions of interest, containing potential target objects, are identified automatically by computing a complexity map from the wavelet filtered image. The root-mean-square (RMS) contrast metric has been used (Fortune et al., 2001; Peli, 1990) to quantify feature complexity, evaluated over a sliding rectangular window.

The resulting complexity map is given by

$$c(u,r) = \sqrt{\frac{\sum_{p,q} \tilde{i}(u + p\Delta u, r + q\Delta r)^2}{\sum_{p,q} \tilde{i}(u + p\Delta u, r + q\Delta r)}}, \qquad p \in \left[\frac{-P}{2}, \frac{P}{2}\right], q \in \left[\frac{-Q}{2}, \frac{Q}{2}\right], \tag{17}$$

This procedure is applied independently to both low and high frequency images and then averaged.

A window with dimensions $P = 48$ and $Q = 96$ was chosen with an aspect ratio of 2, roughly corresponding to the dimensions of the target object. Sliding steps of $\Delta P = 4$ and $\Delta Q = 8$ were used as a trade-off between smoothness of the resulting complexity map and computational load. These simplified data products are communicated to the end users instead of the more complicated sonar images. They are referred to as "heat maps" due to the choice of colour scale, which uses red to represent regions of high complexity, indicating likelihood of a target object, and blue for low complexity regions that are likely to be benign.

3.5 Floor Detection

The location of the water-floor boundary must be determined to facilitate removal of the water column and to estimate the depth for geo-referencing the sonar images and heat maps. Wavelet filtering has been used to isolate features of appropriate scale that are parallel with the path by adjusting the parameters of the procedure described in Section 3.4.1 accordingly. This is followed by edge enhancement, using the Sobel edge-detector (Sobel, 1970) and peak detection to determine the boundary, which is smoothed using a moving average filter. The wavelet filtered image is given by

$$\tilde{\tilde{i}}(u,r) = W^{-1}\left[I(u,r,\mathrm{s}) \cdot \tilde{W}(\mathrm{s})\right], \tag{18}$$

where

$$\tilde{W}(\mathrm{s}) = \begin{cases} 1, & m = \mathrm{V}, n \le n_{\mathrm{f}} \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

is a "high-pass" window that isolates only the vertical detail coefficients $s_V^{(n)}$ for scale levels n_{f} and below. Empirically, a Daubechies 5-tap wavelet and a cut-off scale of $n_{\mathrm{f}} = 4$ was found to provide satisfactory emphasis of the water-floor boundary. The edge-enhanced image is given by

$$\hat{i}(u,r) = S\left[\tilde{\tilde{i}}(u,r)\right], \tag{20}$$

where $S\{\cdot\}$ is an edge enhancement operator. The Sobel operator with a kernel size of 7 points was found to provide satisfactory enhancement of the water-floor boundary in this application, and was implemented via the functions available in the Python Open Computer Vision library (Bradski, 2000). The range to the boundary is estimated for each along-track position x by finding the onset of the peak in the edge-enhanced image

$$R_{\min}(u) = \arg\min_r \left[\hat{i}(u,r)\left(\frac{\mathrm{sign}(\rho(u) - r) + 1}{2}\right) - A\,\rho(u)\right] * h(u), \tag{21}$$

where

$$\rho(u) = \arg\max_r \left[\hat{i}\{u,r\}\right] \tag{22}$$

is the range at which the peak occurs, A is a fraction, and $h(u)$ is a smoothing filter. A value of $A = 0.6$ and a 30-point moving-average filter was found to work well.

The approximate depth of the floor is then estimated along the track using the range to the nadir R_{min}, and the known sonar geometry

$$Z(u) = R_{\min}(u)\,\cos(\phi + \Delta\phi/2) + D, \tag{23}$$

where ϕ is the declination angle, $\Delta\phi$ is the vertical beamwidth, and D is the depth of the sonar transducer. A constant depth with range is assumed, i.e. a flat bottom. The depth estimate is used in the image geo-referencing process, described in Section 3.7.

3.6 Survey Planning and Execution

Survey plans are currently pre-defined by an operator, but executed autonomously by the USV. The plan is specified manually as a list of waypoints in longitude and latitude. The waypoints are converted to local Universal Transverse Mercator (UTM) coordinates before being passed on to the LOS controller described in Section 3.2. The LOS controller requires two waypoints, x_a and x_b, to define a track, as illustrated in Figure 8. Upon initialising the controller, the current position of the USV is taken as x_a (previous waypoint), and the first waypoint in the list is taken as x_b (current waypoint). The USV is guided towards x_b and, once reached, x_a is replaced by x_b, and x_b is replaced by the next waypoint in the list. A waypoint is considered to be reached when the distance to the waypoint $|x - x_b| < T$ is below a user-defined tolerance T, for which $T = 2$ m has been found to be a suitable value. The process repeats until the waypoint list is exhausted and the survey is complete.

A typical survey pattern is comprised of several "lawnmower" patterns of multiple straight tracks similar to a paired-track survey (Hunter et al., 2018), with each pattern oriented at a different fixed angle. Multiple orientations are used in order to observe the floor from different look angles, thus increasing the likelihood of detection (J. Fawcett et al., 2010; Zerr et al., 1997). More orientations with smaller angular increments provide a more thorough search at the cost of a longer survey duration and this leads to a trade-off.

3.7 Situation Map

A world model is created and maintained as the USV collects and processes data throughout its survey. This is stored in the form of a layered situation map, which is updated and communicated in real-time. It conveys information obtained from the survey (i.e., geo-referenced sonar images and heatmaps), the survey plan and execution status (i.e., waypoints, planned and executed tracks), and USV state information (i.e., position, speed, and heading), all overlaid on a satellite map. Currently, the situation map acts as a high-level interface to the operator. In the future, it will be the basis for feedback into a deliberative controller for task planning, sequencing, and execution.

3.7.1 Geo-referencing

The sonar images and corresponding heatmaps from each completed track are geo-referenced and presented as layers with adjustable transparency on the map. The geo-referencing is performed by splitting each image (and heatmap) into multiple along-track sections of length L from u_k to $u_{k+1} = u_k + L$. Each section is interpolated from local image coordinates (u, r) into the global coordinates (x, y) via a projective transform to build a geometric map (Burguera & Oliver, 2016). Thus, a geo-referenced image from a track comprised of K sections is given by

$$i(x, y) = \bigvee_{k=1}^{K} P_k\{i(u, r)\}\ , \tag{24}$$

where the projective transform for the kth section

$$P_k\{\cdot\} : u_k, v_k, u_{k+1}, v_{k+1} \rightarrow \eta_k, \xi_k, \eta_{k+1}, \xi_{k+1} \tag{25}$$

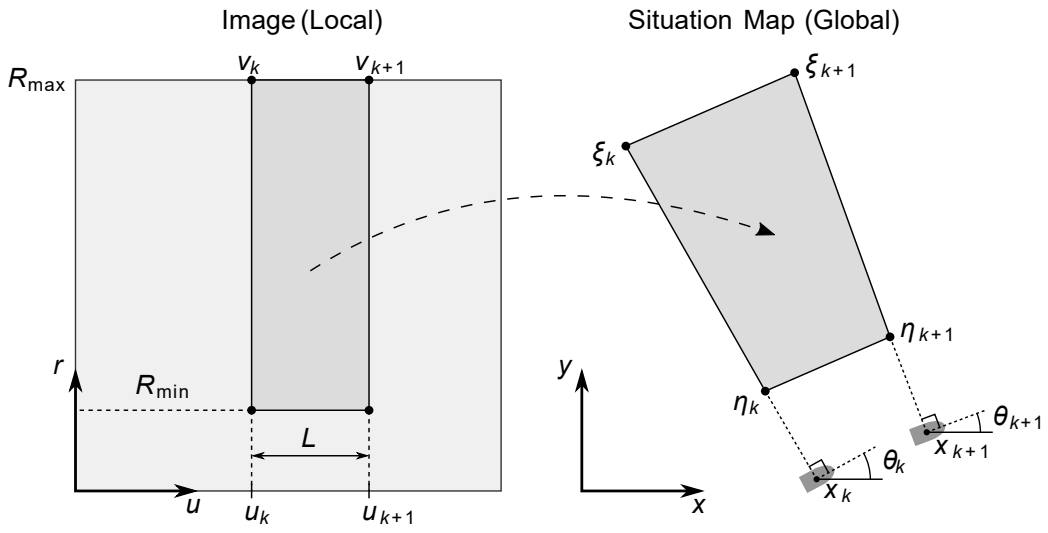


Figure 9: Geometry for geo-referencing one section of a sonar image, illustrated for the port side.

is defined by mapping its corners (i.e., the minimum and maximum ranges at the beginning and end of the section)

$$\mathrm{u}_k = \begin{bmatrix} u_k \\ R_{\min}(u_k) \end{bmatrix} \tag{26}$$

$$\mathrm{v}_k = \begin{bmatrix} u_k \\ R_{\max} \end{bmatrix} \tag{27}$$

to the appropriate locations on the floor

$$\eta_k = \mathrm{x}_k \pm Y_{\min}(u_k) \begin{bmatrix} \sin(\theta_k) \\ -\cos(\theta_k) \end{bmatrix} \tag{28}$$

$$\xi_k = \mathrm{x}_k \pm Y_{\max}(u_k) \begin{bmatrix} \sin(\theta_k) \\ -\cos(\theta_k) \end{bmatrix}, \tag{29}$$

where the + or − are used for the port and starboard sides, respectively. This mapping is illustrated in Figure 7 and described in Figure 9. The locations on the floor are determined from the measured position x_k and heading θ_k of the USV, together with the estimated distance along the ground to the observed water-floor interface

$$Y_{\min}(u) = \sqrt{R(u)^2 - Z(u)^2} \tag{30}$$

and the ground distance

$$Y_{\max}(u) = \sqrt{R_{\max}^2 - Z(u)^2} \tag{31}$$

to the maximum observed range $R_{\max}$.

Here, images are divided into sections of length $L = 1\,\mathrm{m}$. Sections that are within the tolerance distance T of the waypoint are also removed to exclude severely distorted projections that occur during turning, as demonstrated in Figure 7b.

3.7.2 Heatmap Fusion

The geo-referenced heatmaps produced from all of the tracks are collated at the end of the survey to fuse processed information which has been gathered from multiple orientations. This concept has been

shown to improve performance of ATR algorithms (J. Fawcett et al., 2010; Zerr et al., 1997). A simplified implementation is used here, by firstly applying a minimum threshold $C_{\min}$ to the individual heatmaps so that only excess values of significance are accumulated

$$C(x, y) = \begin{cases} 0, & c(x, y) < C_{\min} \\ c(x, y) - C_{\min}, & c(x, y) \geq C_{\min} \end{cases}. \tag{32}$$

These are then averaged across the tracks to produce an overall heatmap for the survey

$$H(x, y) = \frac{1}{M} \sum_{m}^{M} C_m(x, y) \tag{33}$$

where M is the number of tracks. A threshold of $C_{\min} = 0.075$ was found empirically to provide good target / background discrimination.

4 Field Deployment and Testing

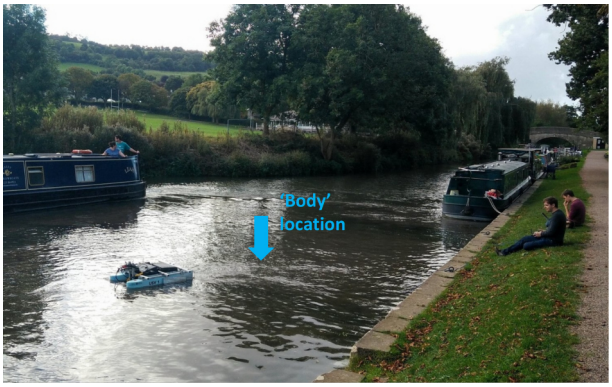
Field experimentation has been carried out at several different locations in the West of England, including the Kennet and Avon Canal at Bathampton (Figure 10a) and Bristol Harbour (Figure 10b). These were chosen as convenient, yet realistic locations for a police search operation. The water depth is approximately 1.5m at the Bathampton canal and approximately 5m at Bristol Harbour. The floor in both locations is a muddy sediment. The results detailed here are from an experiment conducted from Underfall Yard in Bristol Harbour, UK, in July 2021, shown in Figure 10b. This is a representative test location, with several human bodies retrieved from Bristol Harbour in recent years (Brock, 2019; Gillespie, 2021; Gogarty & Mercer, 2021).

A plastic mannequin approximately 1.8m tall has been used as a test target to resemble an adult human body. The mannequin has been clothed, as shown in Figure 10c, to provide a more realistic acoustic scattering signature. When deployed, it is weighted with a small sand-filled plastic weight causing it to fill with water and sink to the floor. A loose rope is also attached and tethered to the shore for recovery afterwards.

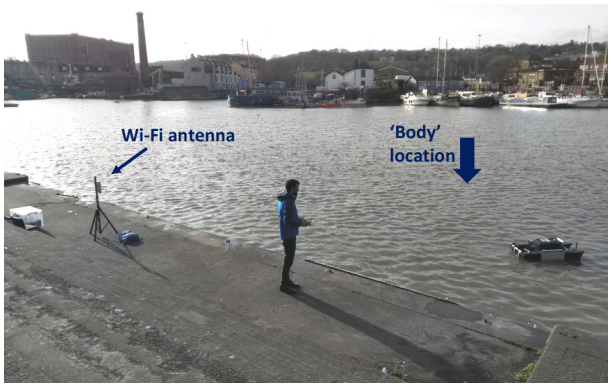
4.1 Representative Images from Bristol Harbour

Figure 11 shows an example of the raw and conditioned images produced from a representative track during the field experiment in Bristol harbour. The images can be interpreted as explained in Section 2.2 and details of the conditioning are given in Section 3.3.1. The conditioned image in Figure 11d is the result of contrast-stretching. It has a sharper water-floor boundary, clearer textures, and the target object is more prominent. The target is clearly visible on the port side as a highlight and shadow of the expected dimensions. The floor is at a depth of approximately 4m and is predominantly soft sediment with a moderate and consistent level of reflectivity. A sparse distribution of small, bright features can be seen against the floor and these are likely caused by hard debris on the floor or possibly individual fish. The harbour wall can be observed at longer range on the starboard side as a strong, roughly linear highlight. The region behind the wall does not present any meaningful features, as it relates to areas the sonar signal cannot physically reach. A region of floor with strong reflectivity extends in front of the wall and this is caused by a change in material properties from soft sediment to hard foundations.

Two bubble plumes can be seen in different parts of the images. The plume in the port-side water column has been generated by the USV thrusters from a previous track. The other plume, which cuts across both images, was generated by a larger vessel that passed through the survey area earlier. The plumes are somewhat less pronounced in the high-frequency, 990kHz image, most likely due to the size of the air bubbles more closely matching the acoustic wavelength at 450kHz. A shoal of fish appears near the nadir gap on the starboard



(a) Bathampton Canal, with earlier PRIME-3 prototype



(b) Bristol Harbour, with latest PRIME-4 prototype

(c) Clothed mannequin test target

Figure 10: Selection of photographs from the various field experiments.

side in the high-frequency image, while this has been mostly obscured in the low-frequency image by the bubble plume.

The two frequency bands provide complementary image properties. The image resolution and contrast is higher at 990kHz due to the more directive beam compared to 450kHz. This can be seen in the zoomed images of the target in Figure 11. On the other hand, the higher frequencies are more heavily attenuated and this limits their operating range. Consequently, system noise can be seen appearing at longer range in the 990 kHz image (i.e., beyond approximately 12m). The noise is exaggerated with range because a time-varying gain is applied to compensate for the acoustic spreading and (frequency-dependent) attenuation loss.

There are some features in the images that do not correspond to the underwater environment but are artefacts of the system. A low-intensity strip is visible on the floor at close range on both sides in the 990 kHz image. This is caused by a null and sidelobe of the vertical sonar beam, which is narrower at the higher frequency. These beam patterns can be compensated using intensity correction techniques (Burguera & Oliver, 2014) but, in this system, the region is covered adequately by the low frequency band. Despite efforts to minimise electrical interference, some cross-talk from out-of-band harmonics generated at 450kHz is present. This manifests as a narrow noisy feature in the water column of the 990 kHz images that is symmetrical about the port and starboard sides.

The wavelet filtered result of the 990kHz image is shown in Figure 12a. It demonstrates enhancement of features corresponding to the target, whilst other details such as changes in the floor texture have been reduced. Some non-target features are still retained, however, such as the shoal of fish due to its similar dimensions to the target. The subsequent heat map is shown in Figure 12b. While the heat map has correctly highlighted the target object, it has also produced a false positive for the shoal of fish and the harbour wall. Furthermore, it has resulted in a false negative in the view from a different track in Figure 12e due to an unfavourable sonar viewing angle as well as another false positive. These issues are mitigated by fusing the information gathered over multiple views from the different tracks, as illustrated in the next section.

4.2 Autonomous Survey in Bristol Harbour

A pre-planned survey pattern was defined over a rectangular region of 50m × 30m, with its longest edge adjacent to and aligned with the shore. Survey waypoints were arranged within the region to form a sequence of six overlapping “lawnmower” patterns of multiple tracks, with an average pattern length of 550m each. The spacing between tracks was 2.5m to provide overlap and nadir coverage. Each pattern was rotated by steps of 15° in the range from 0°–90° to provide multiple views of the floor (either from the port or starboard side of the vehicle) distributed uniformly over the full range of 360°. The cumulative length of the survey was 3.3km with a duration of 45 minutes at the nominal vehicle speed of approximately 1.2m/s. The target was deployed approximately 20m from the shore and 10m from the South-West edge of the survey area.

The full survey plan of overlapping patterns is shown in Figure 13a. A single pattern from this plan is shown in Figure 13b alongside the actual path executed by the USV. Some overshoot and recovery can be seen at the ends of the tracks and there is also an offset of approximately 1m caused by the LOS controller overcoming the water current. Figures 13c and 13d show two other patterns from the survey with the emphasized tracks corresponding to the example images used in Figures 11 to 12.

The USV executed the plan and communicated the situation map to shore in real-time. Figure 14 shows several views of the situation map throughout the survey.

The geo-referenced 990kHz side-scan sonar images from one of the tracks (Figure 13c) is shown in Figure 14a, where it is overlaid as a layer on the satellite view. The features from the harbour wall and foundations (shown in Figure 11 and described in Section 4.1) can be seen to align well with the satellite view. Moreover, the target can be observed at the known deployment location. The corresponding heatmap layer (Figure 12b) is shown in 14b with the target highlighted clearly in red. However, the harbour wall has also been identified

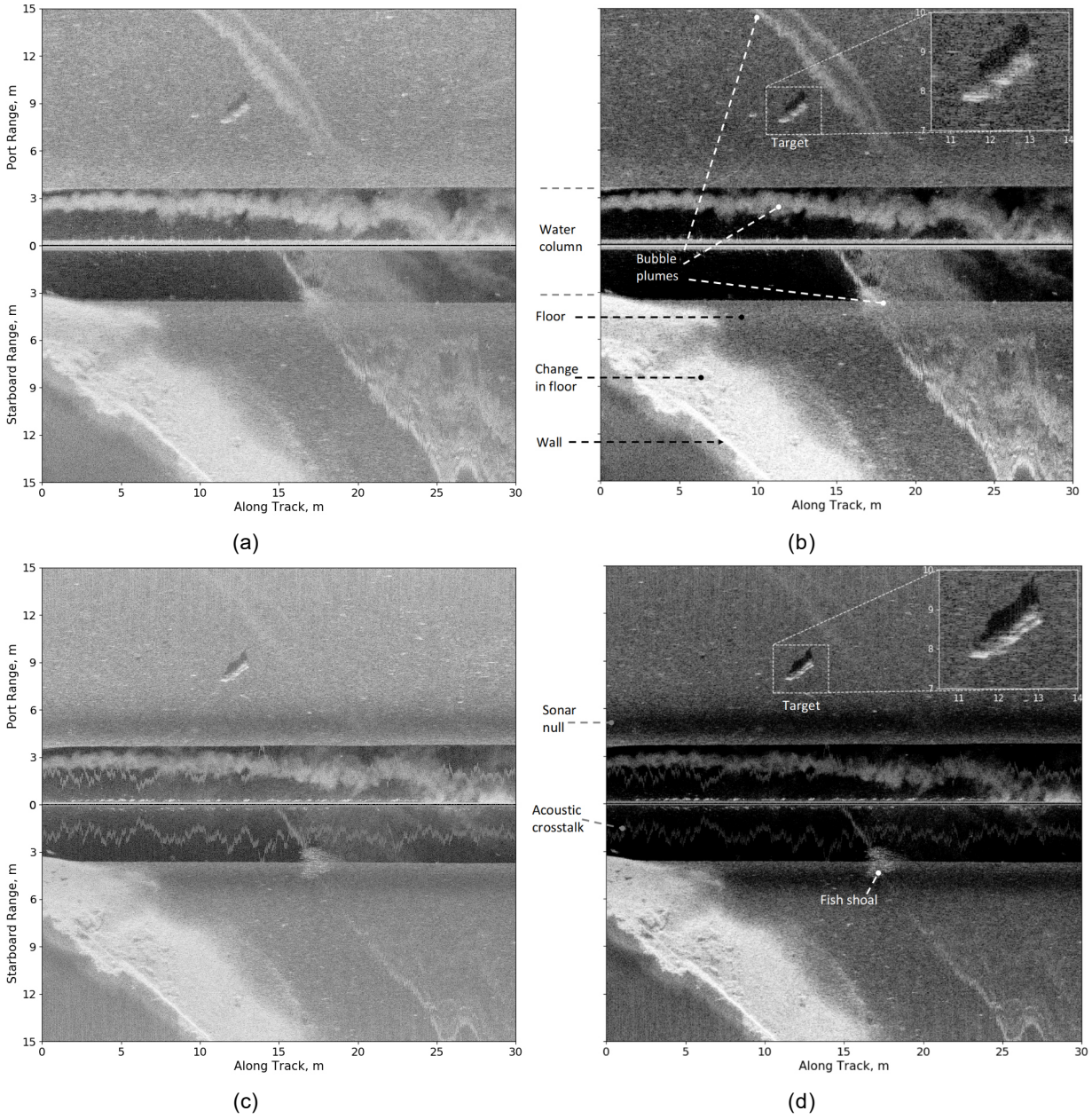


Figure 11: (a, b) Low and (c, d) high frequency side-scan images, before and after conditioning. The mannequin target is indicated by the zoomed inset figures.

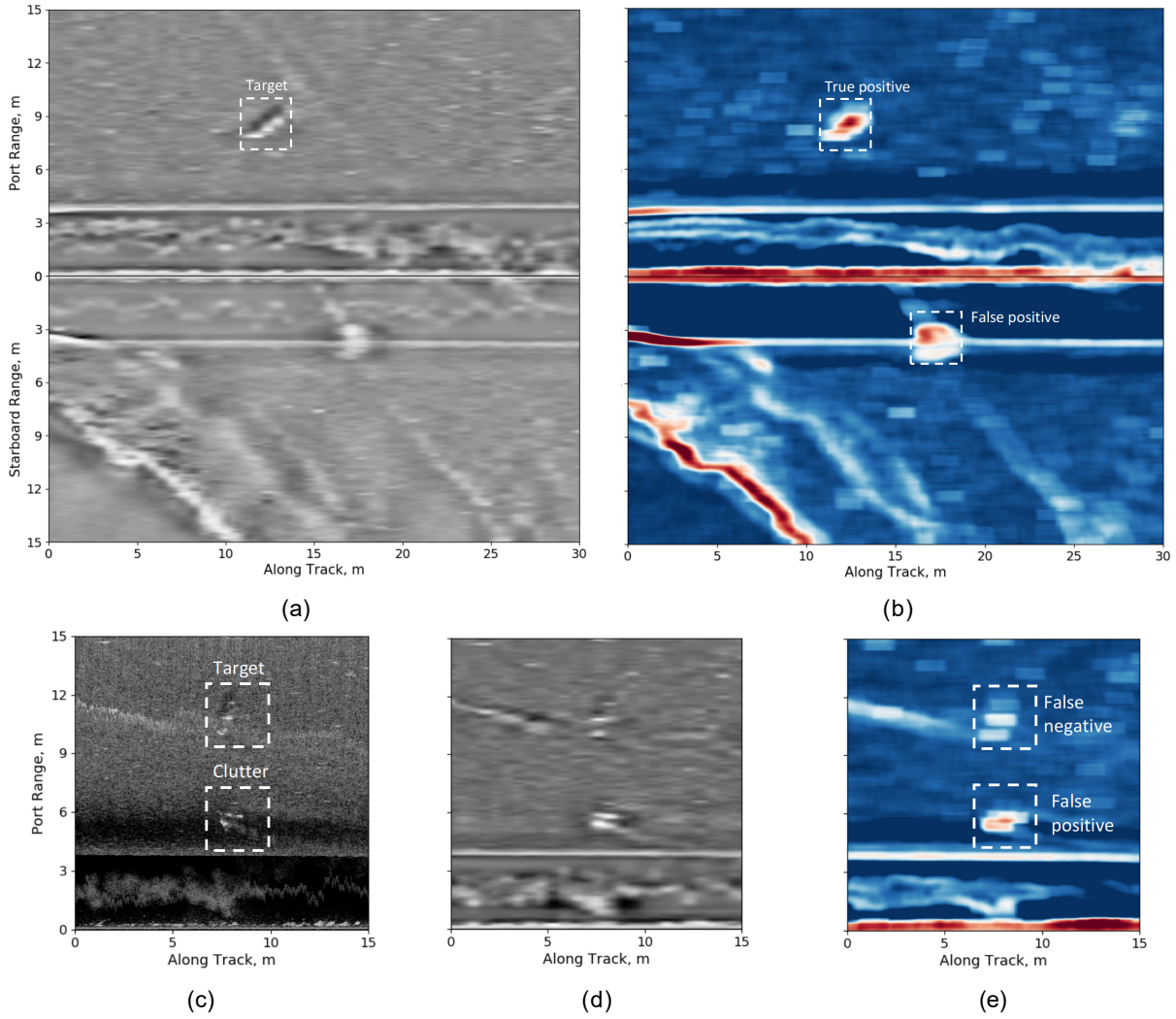
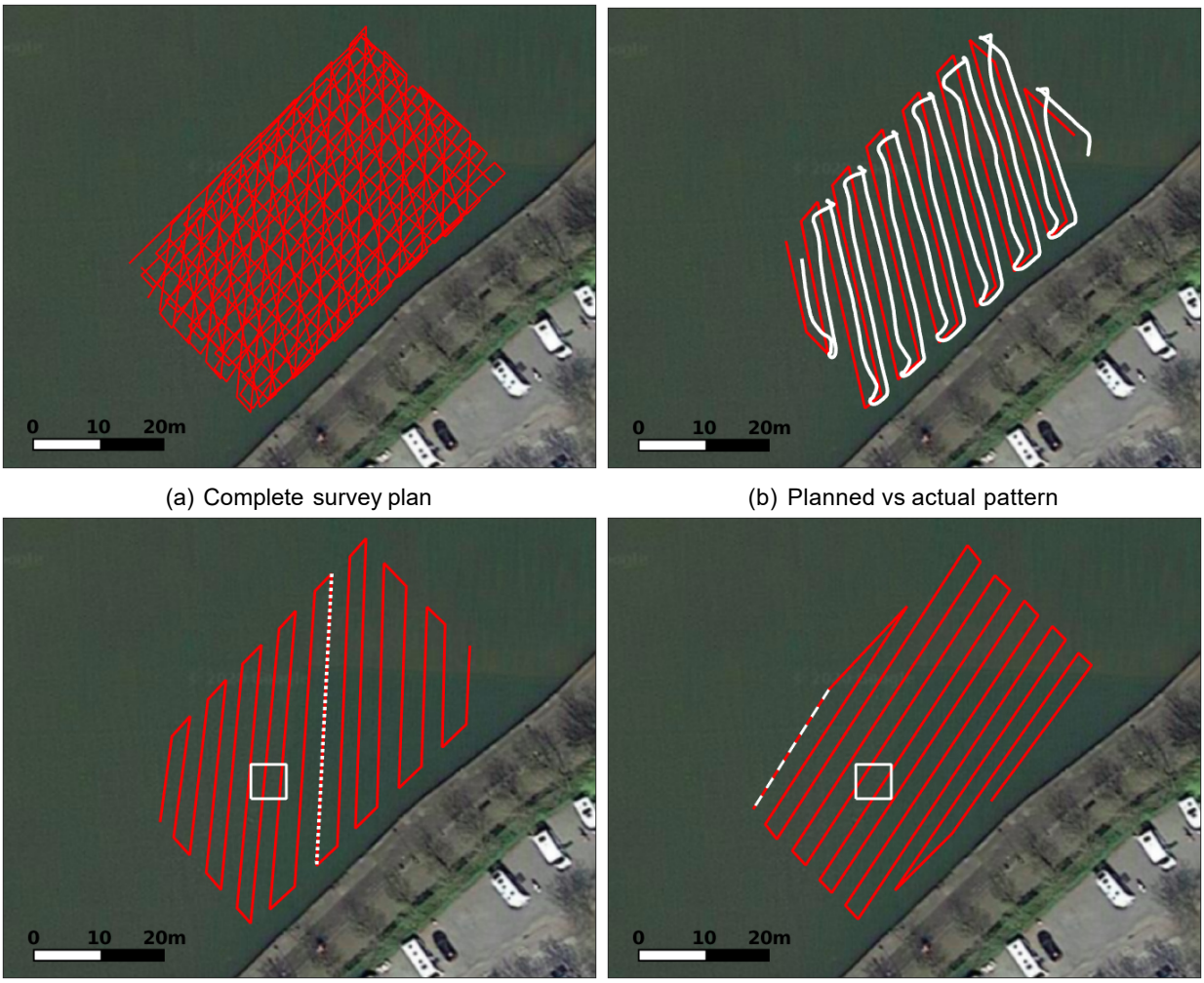


Figure 12: Wavelet filtered images in (a) and (d) with corresponding heatmaps in (b) and (e), respectively, resulting in a mixture of good and poor detection performance. The blue-white-red colour scale ranges from 0.075 to 0.2.



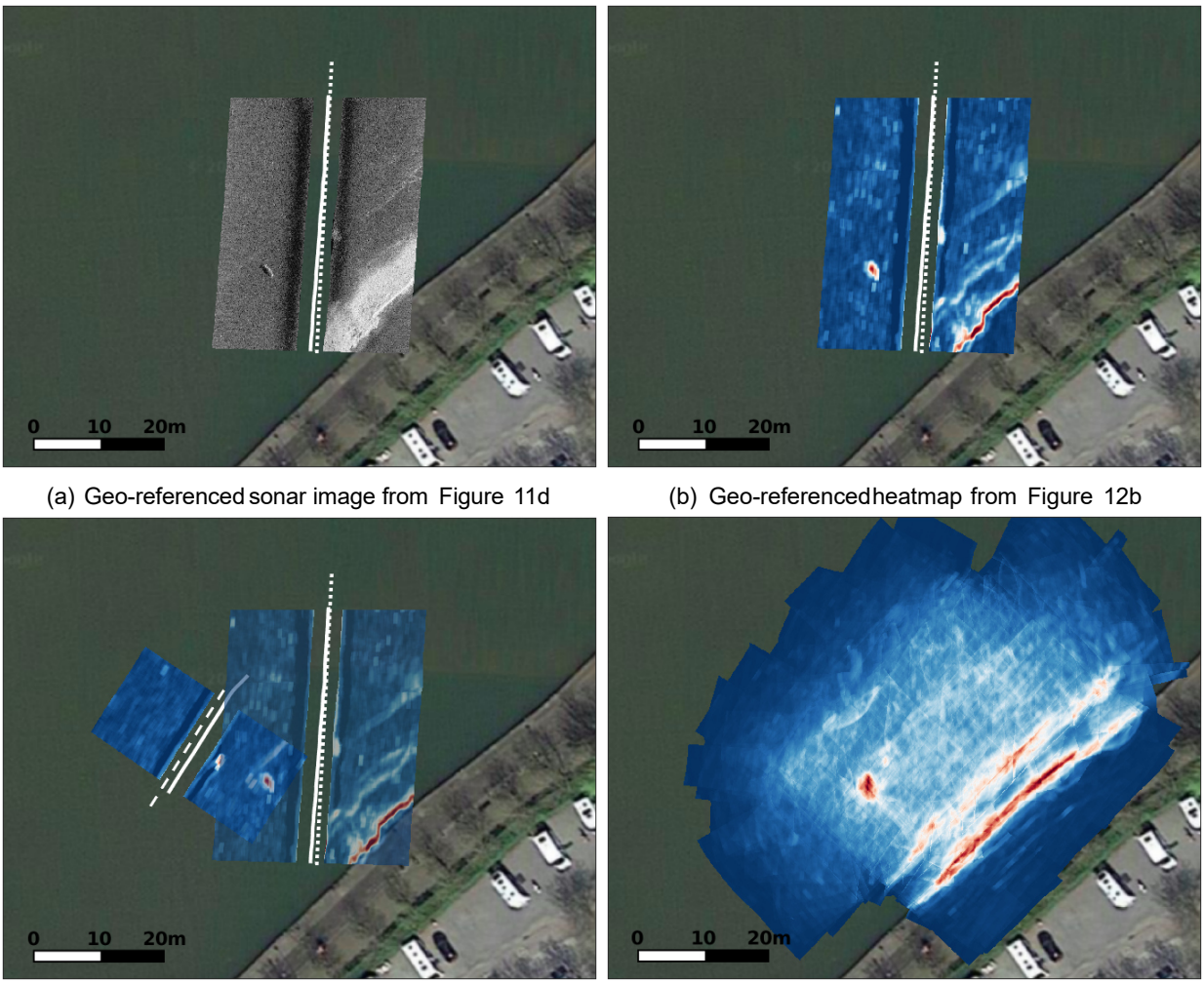
(a) Complete survey plan

(b) Planned vs actual pattern

(c) Planned track in Figures 14a, 14b, and 14c

(d) Planned track in Figure 14c

Figure 13: Survey plan and execution during the autonomous survey in Bristol Harbour, covering a 50m × 30m search area with 6 overlapping "lawnmower" patterns with 2.5m inter-track spacing. The actual path taken (b) shows some offset due to water current. The white boxes in (c, d) indicate the target location.



(a) Geo-referenced sonar image from Figure 11d

(b) Geo-referenced heatmap from Figure 12b

(c) Overlaid heatmap from Figure 12e

(d) Fusion of all geo-referenced heatmaps

Figure 14: Situation map through various stages of the autonomous survey in Bristol Harbour. The dashed white lines show planned paths and the solid white lines show actual paths. The red-white-blue colour scales in (b,c) and (d) range from 0.075–0.2 and 0–0.015, respectively.

as a false positive. The heatmap from another track (Figure 13d) is shown overlaid simultaneously in Figure 14c. In this case (Figure 12e) the target has not been well identified and there is another false positive. The imperfections from the individual tracks motivates fusing the heatmaps from multiple tracks.

The final fusion of heatmaps from all of the tracks was generated at the end of the survey and is shown in Figure 14d. The accumulation of true positives has led to a clear indication of the target location and the various false negatives and positives have been suppressed through averaging. This is the key output, produced autonomously in real time by the system, that provides actionable intelligence to the dive team.

5 Discussion

Results from field testing of PRIME have shown that the current prototype can autonomously navigate a region defined by GNSS waypoints, map the underwater space with side-scan sonar sensors, and detect and localise a human body shaped object located on the floor. The system presents actionable intelligence

to the user in a simple and meaningful manner as shown in Figure 14d. The fusion of results from six survey patterns as illustrated in Figure 13, with an averagepattern length of 550m each, and total length of 3.3km was produced in under an hour. In contrast, a single pattern would take over four and a half hours if carried out manually by a professional diver swimming at a moderate speed of 0.5m/s (Wojtk´ow & Nikodem, 2017). The heatmap effectively shows the target location, as well as benign regions, allowing for the area to be more efficiently searched by prioritising regions of interest. This outcome supports the feasibility, viability, and utility of using an autonomous robotic aid for police and emergencyunderwater search operations. Additional prior information of a search region increases the likelihood of a successful outcome, and minimising the time spent in the water inherently increasesdiver safety (Erskine & Armstrong, 2021). In caseswhere a site requires revisiting, the georeferenced heatmap produced in Figure 16d allows for the search region and features within it to be accurately located.

Further development is required to produce a more robust and reliable system. For instance, the waypoint list is generated manually for a given search area and time-frame, thus requiring the operator to understand the sonar coverage requirements and geography. Ideally, the search patterns would be generated automatically, taking into account land-water boundaries and water depths, as well as user-defined mission parameters such as areas of interest and exclusion zones. Furthermore, the system currently lacks the capability to autonomously detect and avoid obstacles and thus requires monitoring by a user to implement a manual override when necessary. This problem is a well documented area of research (Mousazadeh et al., 2018; Polvara et al., 2018; Wu et al., 2017), and further tests are required to evaluate which approach is most suitable for PRIME.

The path following controller described in Section 3.2 has been found to be robust enough to give sufficiently smooth, straight paths that provide good quality, undistorted imagery. However, since deviations from a linear track causesimage distortion, in the future it may be desirable to include correction methods that compensate for vessel motion e.g. (Blondel, 2009; Burguera & Oliver, 2016) to ensure robustness under heavy currents or winds. Open source versions of several of these processing techniques have been made available (Buscombe, 2017).

A pragmatic approach has been taken to improve target detection quality and lessen the impact of false positives and negatives whereby multiple sonar images are analysed, taken from different positions and orientations, and also from two frequency bands. However, performance has not been evaluated in heavily cluttered environments, which may result in increased false positives during object detection using the current methods described in Section 3.4. Nonetheless, the experiments shown in Section 4 prove the system concept, as an end-to-end solution in a representative environment. Future work will evaluate more diverse sites and scenarios(e.g. varying water depths, bottom types, and more cluttered environments). Autonomous identification of smaller objects such as weaponsor IEDs, and discrimination between multiple object types, including potential hazards (e.g., sharp objects, entanglement risks, etc) will require a more sophisticated sensing and / or machine learning approach. However, any algorithms developed in the future can easily integrated by virtue of the modular nature of ROS. Similarly, additional sensorscan be added to improve target detection and classification capabilities.

Some further refinement is necessaryfor an end product. The current power electronics setup uses multiple batteries running at different voltages powering components independently. While this has been acceptable during development, it would be impractical in an operational system. In future prototypes, all modules will share a single battery bank (e.g., 6S LiPos at 25V), battery management system, and regulated power rails of 5V and 18V that is accessiblevia a common charging port but electrically isolated to avoid interface.

The current communications architecture relies on a single Wi-Fi network set up from a router situated on the shore. This single-network architecture is simple and has been convenient for development, but its weaknessis that critical communication between the onboard computers is disrupted when the channel is blocked, e.g., when the vehicle travels out of range or is occluded. Future improvement to the communication architecture will use separate bridged networks onboard the USV and on the shore, and compatibility with cellular networks. Furthermore, in operational missions, it will be crucial to use encrypted networks (and

onboard memory) for cyber security and data protection.

6 Summary

An autonomous robotic concept for aiding police and emergency service divers has been introduced and demonstrated. This paper has documented the design of the PRIME USV prototype and its field deployment in a representative real-world environment for the relevant use case of body recovery. The experimental results have shown that the system can autonomously produce actionable intelligence on the underwater situation and communicate this to the dive team in real-time (i.e., a simple geo-referenced map demarcating areas of interest). This has utility for enhancing the efficiency of underwater search operations by potentially reducing the time wasted manually searching empty areas. Furthermore, it can enhance safety in the same way by reducing the time that divers need to spend in the potentially dangerous underwater environment. These outcomes support the feasibility and viability of the concept.

This is the groundwork for a future operational concept in human-machine cooperation within the police and emergency services. Reaching this vision will require further research and development to elevate the prototype to a higher level of autonomy and technology readiness. To this end, future work will explore the use of more advanced perception and control algorithms (e.g., deep learning, active learning, etc.) to improve the quality of situational awareness and autonomy. It is expected that this will enable more sophisticated decision-making that allows the system to carry out more challenging tasks, adapt to uncertain environments and unexpected situations, and be trusted to operate safely and effectively with minimal training and oversight.

Acknowledgements

The authors are grateful to the Engineering and Physical Sciences Research Council (EPSRC) for funding this work. They also thank the University of Bath Mechanical Engineering technicians, including Mike Linham, for their assistance in building the various prototypes.

References

Becker, R. F., Nordby, S. H., & Jon, J. (2013). *Underwater forensic investigation*. CRC Press.

Blondel, P. (2009). *The Handbook of Sidescan Sonar*. Springer Berlin Heidelberg.

Blue Robotics Inc. (USA). (2021a). *Basic ESC*. Retrieved June 27, 2021, from https:/ / bluerobotics.com/ store/thrusters/speed-controllers/besc30-r3/

Blue Robotics Inc. (USA). (2021b). *T200 thruster*. Retrieved June 26, 2021, from https://bluerobotics.com/ store/thrusters/t100-t200-thrusters/t200-thruster-r2-rp/

Boyd, J. A. (1987). *A discourse on winning and losing* (G. T. Hammond, Ed.). Air University Press.

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Brock, A. (2019). Body of teenage boy discovered at bristol harbour after 'tragic accident'. *Bristol Live*. Retrieved September 17, 2022, from https:/ / www.bristolpost .co.uk/ news/ bristol-news/ body-teenage-boy-discovered-bristol-3192936

Burguera, A., & Oliver, G. (2014). Intensity correction of side-scan sonar images. *19th IEEE International Conference on Emerging Technologies and Factory Automation, ETFA 2014*, 52–55.

Burguera, A., & Oliver, G. (2016). High-resolution underwater mapping using side-scan sonar. *PLoS ONE*, *11*(1), 1–41.

Buscombe, D. (2017). Shallow water benthic imaging and substrate characterization using recreational-grade sidescan-sonar. *Environmental Modelling and Software*, *89*, 1–18.

C12 Composites Ltd. (UK). (2020). *C12 Composites Ltd.* Retrieved June 26, 2021, from http://www.c12cfd.co.uk/

Chaos, D., Moreno, D., Aranda, J., & De La Cruz, J. M. (2009). *A real-time control for path following of an USV* (Vol. 42). IFAC.

Chitta, S., Marder-Eppstein, E., Meeussen, W., Pradeep, V., Rodr´ıguezTsouroukdissian, A., Bohren, J., Coleman, D., Magyar, B., Raiola, G., Lu¨dtke, M., & Fern´andezPerdomo, E. (2017). Ros_control: A generic and simple control framework for ros. *The Journal of Open Source Software*. https:/ / doi.org/10.21105/joss.00456

Decker, T. (2007). Sophisticated sonar helps solve drowning cases. *The Columbus Dispatcj*. Retrieved July 18, 2022, from https:/ / eu.dispatch.com/ story/ news/ 2007/ 07/ 05/ sophisticated-sonar-helps-solve-drowning/23917495007/

Dhariwal, A., & Sukhatme, G. S. (2007). Experiments in robotic boat localization. *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1702–1708.

Ellis, G. (2012). Chapter 6 - four types of controllers. In G. Ellis (Ed.), *Control system design guide (fourth edition)* (Fourth Edition, pp. 97–119). Butterworth-Heinemann. https:/ / doi.org/ https:/ / doi.org/10.1016/B978-0-12-385920-4.00006-0

Erskine, K. L., & Armstrong, E. J. (2021). *Water-related death investigation: Practical methodsand forensic applications*. CRC Press.

F. Adamo, G. Andria, F. Attivissimo, A. Lanzolla, & M. Spadavecchia.(2013). A comparative study on mother wavelet selection in ultrasound image denoising. *Measurement*, *46*(8), 2447–2456.

Fakiris, E., Williams, D., Couillard, M., & Fox, W. (2013). Sea-floor acoustic anisotropy and complexity assessmenttowards prediction of ATR performance. *Proc. Int. Conf. Exhib. Underwater Acoust.*, (January 2013), 1277–1284.

Fortune, S., Hayes, M., & Gough, P. (2001). Statistical autofocus of synthetic aperture sonar images using image contrast optimisation. *MTS/IEEE Oceans 2001. An Ocean Odyssey. Conference Proceedings (IEEE Cat. No. 01CH37295)*, *1*, 163–169.

Furfaro, T. C. (2012). *A modular guidance, navigation and control system for unmanned surface vehicles* (Publication No. May) [Masters]. Florida Atlantic University.

Gat, E., Bonnasso, R. P., Murphy, R., & Press, A. (1997). On three-layer architectures. *Artificial Intelligence and Mobile Robots*, 195–210.

Geilhufe, M., & Midtgaard, Ø. (2014). Quantifying the complexity in sonar images for MCM performance estimation. *2nd International Conference and Exhibition on Underwater Acoustics*, 1041–1048.

Gillespie, T. (2021). Olisa odukwe: Body found in search for missing bristol university student. *Sky News*. Retrieved September 17, 2022, from https:/ / news.sky.com/ story/ olisa-odukwe-body-found-in-search-for-missing-bristol-university-student-12297335

Girdhar, Y., Xu, A., Dey, B. B., Meghjani, M., Shkurti, F., Rekleitis, I., & Dudek, G. (2011). Mare: Marine autonomous robotic explorer. *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 5048–5053.

Gogarty, C., & Mercer, J. (2021). Bristol harbour searchends as body found. *Bristol Live*. Retrieved September 17, 2022, from https:/ / www.bristolpost.co.uk/ news/ bristol-news/ bristol-harbourside-search-bristol-live-5525154

Hemans, M. (2012). *North west police underwater search & marine unit firearm find*. Retrieved November 19, 2019, from https:/ / en.wikipedia.org/ w/ index.php?title= File:North%7B%5C_%7DWest%7B%5C_%7DPolice%7B%5C_%7DUnderwater%7B%5C_%7DSearch%7B%5C_%7D%7B%5C%%7D26%7B%5C_%7DMarine%7B%5C_%7DUnit%7B%5C_%7Dfirearm%7B%5C_%7Dfind.JPG

Hunter, A. J., & van Vossen, R. (2014). Sonar target enhancement by shrinkage of incoherent wavelet coefficients. *The Journal of the Acoustical Society of America*, *135*(1), 262–268.

Hunter, A. J., Connors, W. A., & Dugelay, S. (2018). An operational concept for correcting navigation drift during sonar surveys of the seafloor. *IEEE Journal of Oceanic Engineering*, *43*(4), 913–926.

J. Fawcett, V. Myers, D. Hopkin, A. Crawford, M. Couillard, & B. Zerr. (2010). Multiaspect classification of sidescansonar images: Four different approachesto fusing single-aspect information. *IEEE J. Ocean. Eng.*, *35*(4), 863–876.

Kaeser, A. J., Litts, T. L., & Tracy, T. W. (2013). Using low-cost side-scansonar for benthic mapping throughout the Lower Flint River, Georgia, USA. *River Researchand Applications*, *29*(5), 634–644.

Kebkal, K., Glushko, I., Tietz, T., Bannasch, R., Kebkal, O., Komar, M., & Yakovlev, S. (2014). Sonobot-an autonomous unmanned surface vehicle for hydrographic surveys, hydroacoustic communication

and positioning in tasks of underwater acoustic surveillance and monitoring. *Proceedingsof the 2nd International Conference and Exhibition on Underwater Acoustics (UA2014), JS Papadakis and L. Bjørnø, Eds*, 1353–1362.

Kelly, J. (2010). The hidden world of police divers. *BBC News Magazine*. Retrieved March 11, 2021, from http://news.bbc.co.uk/1/hi/magazine/8727869.stm

Lee, G., Gommers, R., Waselewski, F., Wohlfahrt, K., & O'Leary, A. (2019). Pywavelets: A python package for wavelet analysis. *Journal of Open Source Software*, *4*(36), 1237.

Liu, T., Dong, Z., Du, H., Song, L., & Mao, Y. (2017). Path following control of the underactuated USV based on the improved line-of-sight guidance algorithm. *Polish Maritime Research*, *24*(1), 3–11.

Liu, Z., Zhang, Y., Yu, X., & Yuan, C. (2016). Unmanned surface vehicles: An overview of developments and challenges. *Annual Reviews in Control*, *41*, 71–93.

Manley, J. E. (2008). Unmanned surface vehicles, 15 years of development. *OCEANS 2008*, 1–4.

McGrady, R. (2019). *NYPD SCUBA in Harlem Meer*. Retrieved November 19, 2019, from https:/ / en. wikipedia.org/ wiki/ File:NYPD%7B%5C_%7DSCUBA%7B%5C_%7Din%7B%5C_%7DHarlem%7B%5C_%7DMeer%7B%5C_%7D(01681).jpg

Moore, T., & Stouch, D. (2016). A generalizedextendedKalman filter implementation for the robot operating system. *Advances in Intelligent Systems and Computing*, *302*, 335–348.

Moulton, J., Karapetyan, N., Bukhsbaum, S., McKinney, C., Malebary, S., Sophocleous, G., Li, A. Q., & Rekleitis, I. (2018). An autonomous surface vehicle for long term operations. *OCEANS 2018 MTS/IEEE Charleston*, 1–10.

Mousazadeh, H., Jafarbiglu, H., Abdolmaleki, H., Omrani, E., Monhaseri, F., Abdollahzadeh, M.-r., Mohammadi-Aghdam, A., Kiapei, A., Salmani-Zakaria, Y., & Makhsoos, A. (2018). Developing a navigation, guidance and obstacle avoidance algorithm for an unmanned surface vehicle (usv) by algorithms fusion. *Ocean Engineering*, *159*, 56–65.

Nguyen, H.-T., Lee, E.-H., & Lee, S. (2019). Study on the classification performance of underwater sonar image classification based on convolutional neural networks for detecting a submerged human body. *Sensors*, *20*(1), 94.

Parker, R., Ruffell, A., Hughes, D., & Pringle, J. (2010). Geophysics and the search of freshwater bodies: A review. *Science & Justice*, *50*(3), 141–149.

Peli, E. (1990). Contrast in complex images. *JOSA A*, *7*(10), 2032–2040.

Polvara, R., Sharma, S., Wan, J., Manning, A., & Sutton, R. (2018). Obstacle avoidance approaches for autonomous navigation of unmanned surface vehicles. *The Journal of Navigation*, *71*(1), 241–256.

Proud, R. W., Hart, J. J., & Mrozinski, R. B. (2003). Methods for determining the level of autonomy to design into a human spaceflight vehicle: A function specific approach.

Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R., & Ng, A. (2009). ROS: an open-sourceRobot Operating System. *ICRA 2009*.

Ruffell, A. (2014). Lacustrine flow (divers, side scan sonar, hydrogeology, water penetrating radar) used to understand the location of a drowned person. *Journal of hydrology*, *513*, 164–168.

S. G. Mallat. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, *11*(7), 674–693.

Schultz, J. J., Healy, C. A., Parker, K., & Lowers, B. (2013). Detecting submergedobjects: The application of side scan sonar to forensic contexts. *Forensic science international*, *231*(1-3), 306–316.

Shin, J., Kwak, D. J., & Lee, Y. I. (2017). Adaptive path-following control for an unmanned surface vessel using an identified dynamic model. *IEEE/ASME Transactions on Mechatronics*, *22*(3), 1143–1153.

Sobel, I. E. (1970). *Camera models and machine perception*. Stanford University.

Southwood, P. (2011). *Underwater searches*.Retrieved December 12, 2019, from https://commons.wikimedia.org/wiki/File:Circular%7B%5C_%7Dsearch%7B%5C_%7Dpattern.png

Taylor, D. (2010). *Using Diamond Coordinates to Power a Differential Drive* (tech. rep.) (http://latiful.hayat.web.id/wp-content/uploads/2020/06/Using-Diamond-Coordinates-to-Power-a-Differential-Drive.pdf). rChordata LLC.

Wan, E. A., & Van Der Merwe, R. (2000). The unscentedkalman filter for nonlinear estimation. *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, 153–158.

Williams, D. P. (2015). Fast unsupervised seafloor characterization in sonar imagery using lacunarity. *IEEE transactions on Geoscience and Remote Sensing*, *53*(11), 6022–6034.

Wojtk´ow, M., & Nikodem, A. (2017). Biomechanics of diving: The influence of the swimming speed on the kinematics of lower limbs of professional divers. *Acta of bioengineering and biomechanics*, *19*(4).

Wu, P., Xie, S., Liu, H., Li, M., Li, H., Peng, Y., Li, X., & Luo, J. (2017). Autonomous obstacle avoidance of an unmanned surface vehicle based on cooperative manoeuvring. *Industrial robot: an international journal*.

Wylie, I. (2019). Bring up the bodies: On the beat with manchester's underwater police. *The Guardian*. Retrieved March 11, 2021, from https:/ / www.theguardian.com/ cities/ 2019/jul/ 26/ bring-up-the-bodies-on-the-beat-with-manchesters-underwater-police

Yang, C.-C. (2006). Image enhancement by modified contrast-stretching manipulation. *Optics & Laser Technology*, *38*(3), 196–201.https://doi.org/https://doi.org/10.1016/j.optlastec.2004.11.009

Young, I. T., Gerbrands, J. J., & Van Vliet, L. J. (1998). Fundamentals of image processing, 45–46.

Zerr, B., Stage, B., & Guerrero, A. (1997). *Automatic target classification using multiple sidescan sonar images of different orientations* (tech. rep.). NATO, SACLANT Undersea Research Centre.