Ocean Engineering 267 (2023) 113202

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Review

Object perception in underwater environments: a survey on sensors and sensing methodologies

Dinh Quang Huy [b,*], Nicholas Sadjoli [a,b], Abu Bakr Azam [b,c], Basman Elhadidi [d], Yiyu Cai [b,c,*], Gerald Seet [b]

[a] SAAB, Singapore
[b] SAAB-NTU Joint Lab, Nanyang Technological University, Singapore
[c] Energy Research Institute, Nanyang Technological University, Singapore
[d] School of Engineering and Digital Science, Nazarbayev University, Kazakhstan

ARTICLE INFO

Keywords:
Underwater robotic
Object perception
Turbid environment

ABSTRACT

Underwater robots play a critical role in the marine industry. Object perception is the foundation for the automatic operations of submerged vehicles in dynamic aquatic environments. However, underwater perception encounters multiple environmental challenges, including rapid light attenuation, light refraction, or backscattering effect. These problems reduce the sensing devices' signal-to-noise ratio (SNR), making underwater perception a complicated research topic. This paper describes the state-of-the-art sensing technologies and object perception techniques for underwater robots in different environmental conditions. Due to the current sensing modalities' various constraints and characteristics, we divide the perception ranges into close-range, medium-range, and long-range. We survey and describe recent advances for each perception range and suggest some potential future research directions worthy of investigating in this field.

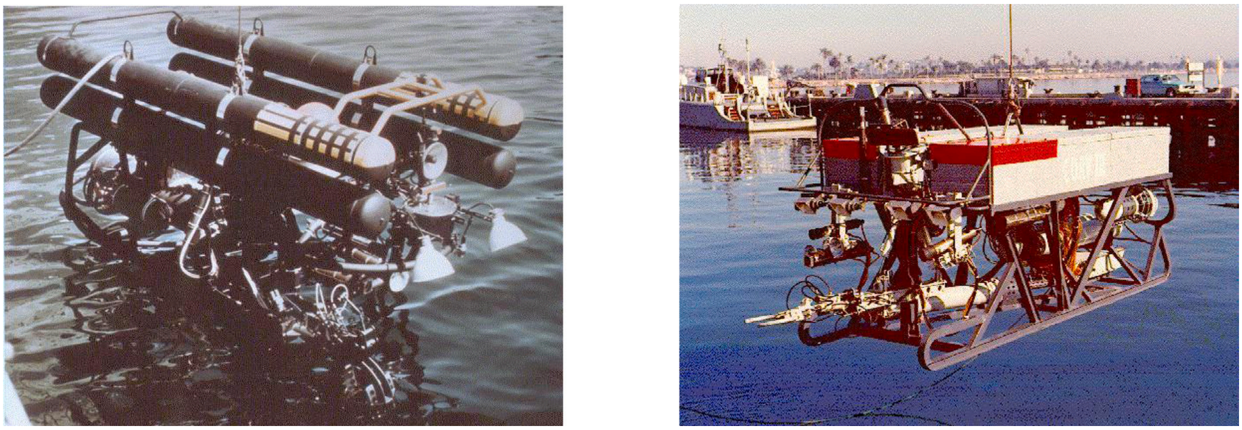
1. Introduction

Although subsea technology has progressed much, 80 percent of the ocean remains unexplored. The deep sea is the last unknown frontier on our planet for humanity. With the increasing oil price and decline in land deposits, there is a rising demand for offshore exploration and mining, infrastructure installation, maintenance, and repair. The development of electric vehicles recently also contributed to the increasing interest in extracting rare minerals in the seabed. Therefore, it is a critical mission to develop sustainable solutions to explore and exploit this massive, cold, dark, and intense environment. The biggest challenge of ocean exploration comes down to its physical properties. The harsh ocean is characterized by extremely low visibility, severe temperatures, and excessive pressure. It is difficult, dangerous, and costly to send human divers to such an environment. Therefore, underwater robotic platforms naturally become decent solutions for these tasks. These submersible vehicles provide robust solutions to survey, explore and exploit ocean areas unreachable by human divers. According to The Global Underwater Robotics Market report by Data Bridge Market Research, the submersible vehicle market was worth 2.685 Billion USD in 2020 and is projected to reach 6.719 Billion USD in 2028. Remotely Operated Vehicles (ROVs) and Autonomous Underwater Vehicles (AUVs) accounts for the two most prominent type of segment. The market is expected to grow at a Compound Annual Growth Rate (CAGR) of 12.15 percent from 2021 to 2028. The demand for underwater robots comes from four main areas, including military and defense, marine research, offshore industries, and rescue and repair services, in which offshore industry and military concerns are the key market drivers.

Underwater robots such as Remoted Operational Vehicles (ROVs) were initially developed for military missions. ROVs offer a telepresence robotic solution for the operator to perform submersible tasks while remaining comfortable on the water surface. The human–robot interaction is accomplished via an umbilical cable. In the past, many ROVs were deployed to perform multiple rescue and recovery tasks, including retrieval of torpedoes, mines, and weapons, such as the Cable Controlled Undersea Recovery Vehicle (CURV)-I (Fig. 1(a)) and its upgraded version CURV-III (Fig. 1(b)). Nowadays, many countries are investing heavily in underwater robotics for military applications. In 2021, The US Navy was allocated 1.76 Billion USD for research, development, testing, and evaluation of unmanned underwater vehicles. Military ROVs are usually deployed to observe and assess the situation in many submersed tasks before sending any military divers into the ocean. ROVs are also highly effective in helping military officer

* Corresponding authors.
E-mail addresses: qhdinh@ntu.edu.sg (Q.H Dinh), myycai@ntu.edu.sg (Y. Cai).

https://doi.org/10.1016/j.oceaneng.2022.113202
Received 16 August 2022; Received in revised form 28 October 2022; Accepted 16 November 2022
Available online 10 December 2022
0029-8018/



(a) CURV-I in a nuclear bomb retrieval mission (b) CURV-III in a human rescue mission

Fig. 1. Early development of underwater robots for military applications with the Cable Controlled Undersea Recovery Vehicles.
Source: Pike (1999).

(a) Swordfish (b) Bluefin-12S (c) Boeing Echo Ranger

Fig. 2. Different types of AUV.
Source: Autonomous Undersea Vehicle Application Center (AUVAC)

in the battle against contraband, where the smuggling products is hidden under the ship's hull while reducing the risks of using soldier divers. Other ROV's military tasks include security scanning of ports for anomaly objects and underwater criminal investigation. On the other hand, AUVs are the new weapons added to the naval force in the last two decades. The tetherless vehicles can move with higher movement speed than ROVs, have low energy consumption, and have reliable data capture capability. The platforms can operate in inaccessible ocean areas independently without receiving commands from an operator, making them perfect weapons for surveillance and reconnaissance tasks in controversial areas. The US Navy classifies military AUVs into four main types using their size and weight: man-portable, lightweight, heavyweight, and extra-large. Man-portable AUVs weigh around 10 to 50 kg and usually have a torpedo-shaped design. An example of this type of AUV is the Swordfish (Fig. 2(a)). Man-portable AUVs are usually launched from small vessels for reconnaissance, sea exploration, seabed mapping, and mine countermeasure missions. The robots can travel to a depth of 40 ft with speeds of up to 5 knots. Lightweight AUVs are slightly bigger than man-portable AUVs. They can weigh up to 277 kg and usually be deployed using a crane system. An example of this type of AUV is the 213-kg Bluefin-12S (Fig. 2(b)). Compared to man-portable AUVs, the vehicle features a longer operation time of 26 h and a higher payload with a slightly slower speed (3 knots). Mine detections, mine disposal operations, and countermeasure are their typical applications. Heavyweight AUVs can be 5000 kg to 10000 kg in weight. They are specially designed for longer and larger missions whose operating time can range from 40 h to 80 h. A famous example of this AUV class is the 5000-kg Boeing Echo Ranger (Fig. 2(c)). The vehicle can operate at a depth of 10000 ft to perform similar tasks as the smaller classes. Extra-large AUVs are vehicles that weigh more than 10000 kg. They can be considered as an autonomous version of the submarine. The main tasks of the massive mechanisms are to detect and destroy the submarine, support surface military ships, mine countermeasures, and strike missions. A typical extra-large example is the Cutthroat LSV-2 and Boeing's Orca.

The oil and gas industry is in massive demand of underwater robots. With the dwindling of these fossil fuels beneath the land, harvesting oil and gas on the ocean floor becomes indispensable. Almost 30 percent of current oil and gas production comes from the offshore industry. Since the location for extracting oil as gas can be extremely deep inside the ocean, it is risky and even impossible for human divers to perform site exploration, infrastructure inspections, intervention, and repairs. The divers sometimes must pull the equipment out of the water for repair, which is time-consuming and costly. On the other hand, robotic platforms for the offshore industry are designed to endure extreme pressure and temperature while maintaining a stable operation in numerous dynamic conditions. Most vehicles have optical, acoustic, and robust light systems to deal with poor visibility conditions. ROVs for the offshore industry can be categorized into two main classes: observational ROVs and work-class ROVs. Observational ROVs are small vehicles equipped only with cameras, sonars, and light systems. They can record and transmit real-time visual data for pure observation tasks such as surveying and inspection. Work-class ROVs are larger and more powerful, with additional sensors and even strong manipulators for lifting capabilities. ROVs and AUVs are becoming indispensable in the gas and oil industry. They can get involved in seabed surveying and acoustic mapping before and after cable and pipeline installation. They can also perform nondestructive inspections for the pipeline and

(a) Nexxus - A heavy work class ROV with intervention system.

(b) Nexxus ROV inspects and repairs the pipeline system damaged by corrosion.

Fig. 3. Example of using ROV for offshore infrastructure maintenance and repair.
Source: Oceaneering,Charalambides (2016))

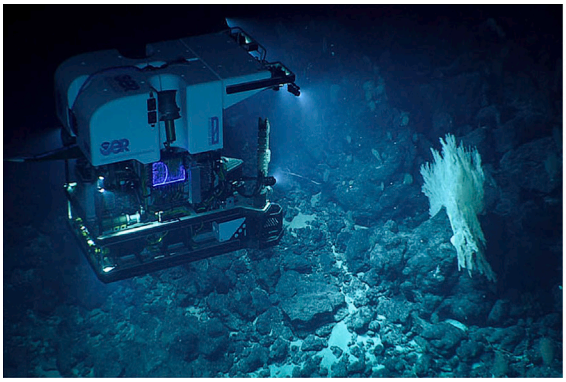
umbilical cable system or physical intervention tasks such as debris removal, valve actuation connection, and disconnection. These robots can work independently or with human divers in various functions, including oil survey, drill support, construction support, inspection, maintenance, and repair (see Fig. 3).

Marine research is another field that is significantly beneficial from the development of the underwater robot. Marine science provides human answers to a vast number of topics such as biodiversity, submerged physics, seabed topography, climate changes, carbon cycle, invasive and endangered species, or natural evolution. Marine research is constrained by the challenging underwater environment requiring sophisticated equipment for complex and time-consuming experiments. Underwater robots allow marine scientists to overcome these limitations and improve the quality and efficiency of this research. Many researchers use an ROV with a high-resolution camera and a long-duration battery to closely monitor and record invasive species' impact on the local species. Aquaculture researchers use ROVs to monitor the behavior of specific species, such as fish farms and sea cucumbers, in their habitats. One significant advantage of ROVs in species monitoring is the capability of minimizing unfavorable alteration of surrounding environments for the species. Many ROVs can be customized with additional sensors to perform oceanographic sampling to obtain samples from the seabed. Reef monitoring for coral bleaching is another example of using underwater robots for marine research. The robot can capture high-resolution images of the reef to be analyzed by marine scientists to identify any health problem. Sea floor survey using ROVs equipped with multi-beam sonar is also a very active research issue where a robot with beam-forming sonar is deployed to reconstruct the digital map of an area of the ocean floor. AUVs have also become more common in marine research with advanced sensing modalities such as high-resolution cameras, multi-beam sonar, MEMS (microelectromechanical system), and advanced acoustic positioning systems. Researchers can program the robot's trajectory before the mission, launch the vehicle, recover the robot after the journey and collect the data. There is no need for a surface vessel during the AUV's mission which is a much more cost-effective solution for researchers. AUVs with advanced sensors can travel deeper into the ocean, close to the seafloor while maintaining a close distance from underwater organisms. This feature dramatically enhances the quality of data captured by the robot, resulting in more efficient ocean explorations. One major disadvantage of AUVs usage is the difficulty retrieving the robot if a problem occurs during its missions (see Fig. 4).

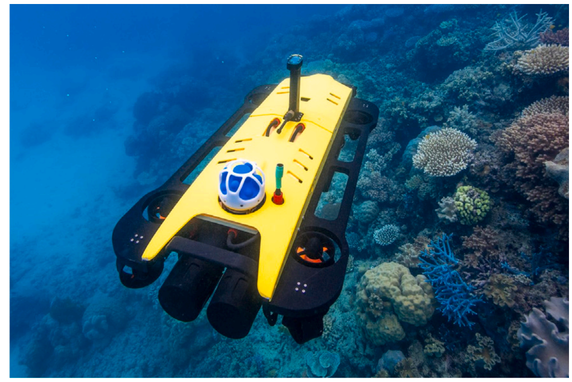
Underwater vehicles are excellent solutions for underwater search and rescue processes. These operations require urgent actions with a well-prepared and safe operational plan. The operations may happen in arduous submersed areas, which puts the rescuer in many life-threatening situations. The first step of any rescue operation is to identify if there is a chance to save a life. Well-trained and specialized divers must execute this process. However, there is always a risk of sending human divers to such cold, low visibility, and uncertain conditions. Underwater robots such as ROVs can be highly beneficial in such contexts. With the development of aquatic perception systems, many search and rescue operations send underwater robots to inspect, evaluate the situation, or locate the target or critical areas before sending any divers. The operational change reduces risk factors for the divers and saves them a lot of time and effort. Human divers also cannot stay underwater long and must take breaks between dives. An underwater robot can work longer and in deeper areas. Some robots can be equipped with a grasping arm to retrieve objects and pull them to the water's surface. One recent example of this application is the search and rescue campaign for the Malaysia Flight MH370. The US Navy deployed the Bluefin-21 AUV to search the seabed to find the missing airplane.

The development of underwater robotics depends on the advances of perception, planning, and control systems which correspond to the "see, think, act" cycle of a robot's operation. Perception is associated with a robot's capabilities to sense and understand its environment. The process requires various advanced sensors to capture environmental information, which is processed using sophisticated sensing algorithms to extract meaningful information. Underwater perception is crucial and much more complex compared to the terrestrial perception. The ocean environment is a highly dynamic unconstrained scenario. Therefore, the robot must possess an advanced perception system to improve situational awareness that can help it to comprehend and react to any uncertain events that can happen accidentally. Also, as most ocean areas still need to be explored and seawater attenuation effects limit large continent remote-sensing techniques, there is a lack of maps and data for deep ocean navigation comparing land navigation. The deployed autonomous robot relies on its perception sensors and sensing algorithms to accomplish essential tasks such as localization, mapping, navigation, or trajectory planning. Underwater perception is also vital for joint operations, requiring the robot to perceive and reason about the relative relationship between itself and its partners, sharing the same work for efficient collaboration results.

So, perception is a critical prerequisite for underwater robots, especially in deep-sea scenarios. Various sensors and sensing algorithms have been developed to achieve automation capability for submersible platforms. While there have been survey articles on navigation and localization (Jalal and Nasir, 2021), planning and control (Guo et al., 2021), obstacle avoidance (Cheng et al., 2021), object tracking (Kumar and Mondal, 2021), communication (Sozer et al., 2000), and machine learning (Christensen et al., 2022) for underwater robotics, a survey paper that cover a comprehensive understanding of underwater



(a) ROV Deep Discover for ocean exploration and research



(b) CoralAUV autonomously following reef contours and slopes

Fig. 4. Examples of using ROV and AUV for marine research.
Source: NOAA Office of Ocean Exploration and Research, Australian Institute of Marine Research.

modalities and sensing algorithms still needs to be included. This has motivated us to write this survey paper to provide an overview of state-of-the-art underwater perception technology. The article presents sensing modalities and algorithms using the perception range to offer an intuitive understanding of the methods. The rest of the paper is organized as follows. Section 2 explains the challenges of underwater perception in detail and summarizes the current state of the art of sensing devices. Section 3 describes methodologies for sensing algorithms. We provide some insights for future field developments in Section 4. We conclude the paper with a summary exemplifying the evolution of underwater robot perception and future results in Section 5.

2. Underwater perception

Underwater object perception can be defined as becoming aware of the surrounding objects in the environment via different types of sensors and sensing algorithms. This process allows the robot to detect, track, and recognize environmental things accurately. In this section, we first introduce the challenges for underwater robot perception. Then, we present a summary of state-of-the-art sensing modalities for different perception ranges.

2.1. Underwater perception challenges

2.1.1. Environmental challenges

The underwater environment strongly absorbs visible lights. The absorption process happens due to the collisions between photos and water molecules, which results in water heating. The interaction prevents visible lights from traveling further into the water, strongly attenuating light in the red and violet spectrum. The attenuation effect is reduced in the blue and green range. The absorption effect causes red and violet light to disappear at less than 7 m, yellow and orange color at 15 m, and blue and green color at around 30 m. The impact explains why seawater mostly has blue or green color. Additionally, underwater substances such as biological matter and dissolved organic matter can further attenuate light and tend to do so more in the blue wavelength. Other effects that reduce ambient light in underwater environments include light reflection, light refraction, and light diffusion. Light reflection involves the bouncing back of light at the water's surface, while light refraction occurs when light beams scatter through the water. Light diffusion occurs when light is bent away from the target when it travels through the water. These interaction mechanisms explain why most visible lights cannot shine intensely into the water, and the deeper the robot sinks into the water, the darker the surrounding environment (see Fig. 5).

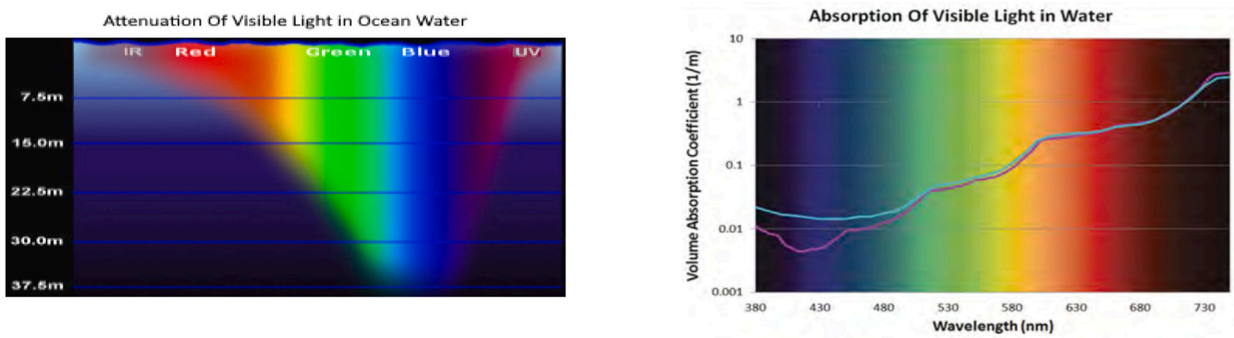
Many researchers address this low visibility issue using an active light system attached to the robotic platforms. The solution is inefficient due to two main reasons. First, the artificial light system also suffers from energy attenuation effects which prevent light from entering deeper layers of water. Second, the "back-scattering" effect frequently occurs in underwater imaging when the active light system illuminates the particles between the camera and the object or the open water space behind the object instead of the object. This phenomenon results in blurred and noisy images in many underwater optical systems. While it may help increase the illumination of the object, back-scattering also reduces the contrast between the object and the background. The problem is due to back-scatter along the illumination path, washing out the object and scattering the reflected light from the object, blurring the image. Increasing the total illumination of the object will not improve contrast in this case because the scattering scales with intensity and no net contrast increase will result. The effect limits the object detection distance in contrast-limited imaging applications like human vision or film. These perception problems worsen in a turbid environment as higher turbidity increases light attenuation and scattering intensity. The dense concentration of clay, silt, algae, and other organic matter in turbidity water can disable optical vision systems (see Fig. 6).

2.1.2. Underwater sensor limitations

Underwater contexts are significantly different from terrestrial environments. Many sensing modalities such as Light Detection and Ranging (LIDAR), Radio Detecting and Ranging (RADAR), or Global Positioning System (GPS) do not function underwater due to severe attenuation of electromagnetic waves, limiting the sensors and sensing techniques used for sub-sea perception. Optical- and acoustic-based sensors are two primary sensors for underwater perception that can operate in these conditions. However, the performance of these devices also suffers significantly due to some environmental factors. Typical examples of such parameters are ambient light and turbidity level.

Optical-based sensors can provide very high-resolution images. However, they can only capture object that is close to the sensor. Also, the image quality suffers in low visibility underwater conditions due to the lack of ambient lights. This limitation is further exacerbated in challenging turbid water environments, as shown in the example in Fig. 7(b). This decrease in range and visibility usually causes optical imaging sensors to be used for closer-range perception.

Meanwhile, acoustic-based sensors such as sonars are very robust in difficult water conditions since acoustic signals are not heavily attenuated underwater. However, acoustic signals have much lower bandwidth than optical signals, causing relatively lower resolution data. This limitation has recently been mitigated with the development



(a) Attenuation of light sources in ocean water.

(b) The absorption coefficient of light in different water conditions.

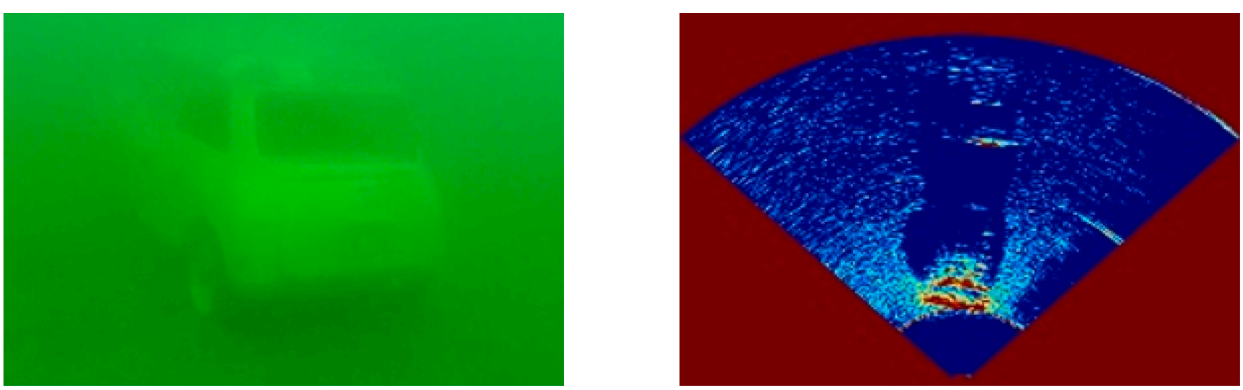
Fig. 5. Understanding light attenuation in underwater environments.
Source: Tom Morris, Fullerton College.

(a) Turbid water greatly reduces image quality

(b) Back-scattering effect in open area

(c) Light diffuses into the water

Fig. 6. Understanding light attenuation in underwater environments.
Source: Brent Durand, Underwater Photography Guide.



(a) Object in turbid water captured by a camera

(b) Object in turbid water captured by a forward-looking sonar

Fig. 7. Limitations of underwater sensors.

of high-resolution imaging sonars such as multibeam FLS. However, the 2D image data produced are still of relatively low resolution for objects at longer ranges. Acoustic signal also has a low signal-to-noise ratio, making it easily suffer from environmental noise. Another limitation of acoustic-based imaging sensors is the loss of spatial information from 2D images. This limitation is mainly due to the processing used by the sonars, as visualized by the example of the FLS model seen in Fig. 8. Imaging sonars would detect points within its insonification region formed between its horizontal field-of-view - α, and vertical field-of-view - β. 3D information of objects is then rendered onto the 2D image plane shaded in blue. Such projection effectively removes the depth information h_1 and h_2 of the objects from the final produced image. In practice, this removal of depth causes underwater acoustic images of objects with complex shapes to lack clarity and distinctive features, as shown in Fig. 7(a) where a car submerged in a turbid lake is rendered to be an ambiguous 'blob' shape in the FLS image.

2.2. Underwater sensing modalities

Various types of sensor systems have been developed to help robots adapt to the underwater environment's challenging conditions. This surveying paper will review three primary sensors for underwater perception: sonars, cameras, and acoustic-based positioning systems.

Three types of sonar are prevalent for underwater applications: multibeam sonar, side-scan sonar, and mechanical sonar. This paper will not discuss mechanical sonar since it offers no advantages over the other sonar types. Side-scan sonar is ideal for search and rescue operations (SAR) since it is lightweight, deploy-friendly, and cost-effective. Side-scan sonar must be towed behind a boat or mounted at the side of a ship. Once the equipped boat moves, the device can generate a sonar image of the areas below, both to the left and right sides of the sonar. While side-scan sonars provide a convenient solution to scan a sizeable sea-bed area to produce bird-eye-view images, they

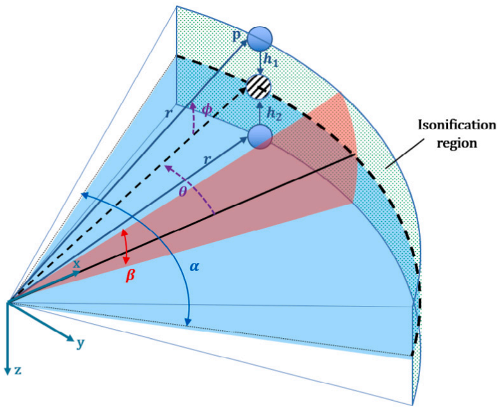


Fig. 8. FLS model visualizing its image processing. Projection causes depth information to be lost in the final image produced.

usually lack the object's details. The reason is that side-scan sonars manipulate low-frequency components and operate at a high altitude to quickly check a vast site, limiting their capability to scan small targets. Another drawback is their limited flexibility. The device must be moved continuously in one direction and stick to the search pattern during an operation, even if it detects the target. The operator cannot reorient the device during the procedure or risk losing area information. A multibeam sonar, on the other hand, focuses on scanning the fan-shaped area in front of it. The device can be pointed in different operations directions and produce sonar images even at stationary. The feature makes the device more flexible compared to side-scan sonars, which explains why multibeam sonars are very suitable for mounting on ROVs or AUVs for object perception tasks. Although multibeam sonar can only cover a smaller area, it can offer better resolution than side-scan when looking for small targets. Although integrating data from multibeam sonar is a bit harder than side-scan sonar, a multibeam sonar can generate a side-scan image by performing a structured search as side-scan sonar. Ideally, an underwater operation combines both sonars in submersible tasks. A side-scan sonar is deployed to cover a broad ocean area. Once a target is identified, an ROV or AUV equipped with multibeam sonar is sent to collect more detailed data. An example of different types of sonar is shown in Fig. 9.

Optical cameras for underwater applications are different from standard cameras. While traditional camera systems focus more on improving image quality by improving the sensor's resolution, underwater cameras turn to improve low-light sensitivity and video latency. The reason is that high resolution does not matter in low light and turbid water. Using a built-in light system from the robot will cause a back-scattering effect blinding the observer or robot's perception system with light reflected by many particles in the water. It is preferable to navigate the robot with natural light in such a situation. Generally, a camera that can generate HD video or images is enough for subsea contexts. Latency is another critical aspect of the underwater video stream. Low latency camera allows the robot or human operator to react timely and effectively to uncertain events, improving the robotic systems' responsiveness and robustness. Examples of underwater camera systems are shown in Fig. 10.

Acoustic-based positioning systems are methods developed to replace GPS in subsea contexts using acoustic-based sensors. These methods include Long Baseline (LBL), Short Baseline (SBL), and Ultra-short Baseline (USBL) Systems. Long-baseline (LBL) systems (Fig.) require a network of transponders mounted at the seafloor before operations. The network contains at least three transponders for proper positioning. This transponder system can track an ROV and AUV equipped with an interrogator. With accurate transponder positioning, LBL systems can achieve very high accuracy, less than 1 m or even 0.01 m. Short-baseline (SBL) systems use a set of three or more sonar transducers mounted at different locations of a surface vessel for positioning. The system accuracy depends on the distance between the transducers and the mounting method. It can achieve accuracy similar to the LBL system when the space is wide enough. Ultra-short-baseline (USBL) systems consist of a transceiver and a transponder. The transceiver is attached to a pole under a surface vessel while the transponder is equipped on the subsea equipment, divers, ROVs, or AUVs. An acoustic pulse from the transponder replies to an acoustic pulse emitted by the transceiver. The range between the transceiver and the transponder can be inferred using the time between transmissions. The USBL head also contains an array of transducers separated by a short baseline. The system uses a phase-differencing method to obtain the direction or bearing of the returned acoustic signal. These positioning systems are illustrated in Fig. 11.

ROVs and AUVs mainly use optical cameras and multibeam sonars for underwater perception tasks. Optical cameras are suitable for capturing objects with high resolution at short distances from the sensor. On the other hand, Multibeam sonars are solutions for object perception at a longer distance. We will summarize the utilization of these sensors in the current underwater perception tasks using Fig. 12, which demonstrates a current methodological framework for underwater perception task. First, suppose the robot is far from the target, as demonstrated by the blue circle. In that case, it can obtain the target's range and heading using an underwater positioning system, given that the system can track the location of the docking station. At a distance of approximately 30 m from the docking station (green circle), the robot enters a long-range scenario where the sonar sensor starts to see the docking target. However, due to the sonar limitation and the low SNR (signal-to-noise ratio), the target resolution is still meager, which makes it impossible to recognize the target using geometrical features such as the object's size or object's shape. The robot can identify the station's correct location using special features such as the target's motion or the target trajectory. Once the robot reaches the orange circle (medium range), the sonar and camera sensor can start seeing the target depending on the water condition. The robot could perform some data fusion algorithms by complementing optical and acoustic data. Finally, at a very close range (red circle), the perception tasks become more comfortable with the accessibility of high-resolution captured data from optical and acoustic sensors. The robotic platform can afford many sophisticated perception algorithms, such as 3D point cloud reconstruction and pose estimations. We summarize the sensor utilization for underwater perception in different ranges in Table 1. In the next section, we describe underwater perception algorithms of three different ranges, including long-range 5m – 30m, medium-range 1m – 5m, and close-range (<1m). We will skip the very long range 30m – 10km since the robot's perception at this range is mainly performed through an underwater positioning system such as SBL or USBL without using any sensing algorithms.

3. Underwater object perception

3.1. Long range object perception

While AUVs are very efficient and safe solutions for various underwater tasks, the limitations of equipped sensing modalities make underwater perception difficult. In long-range object perception, acoustic-based sensors are the only sensor type that provides reliable results, especially in turbid conditions. Early research focused on using classical sector scanning sonars with a fast update rate to perform perception tasks for AUV platforms. The sensor returns each resulting scan as an image display where the received acoustic energy from detected obstacles is converted into pixel brightness. These sequences of images will be the input to different perception algorithms. A completed algorithm for object perception using forward-looking sonar (FLS) in long-range contexts using sonar images includes multiple steps: data preprocessing, object detection, object tracking, and object classification. We will summarize some of the work in the field, focusing on different parts of the framework.

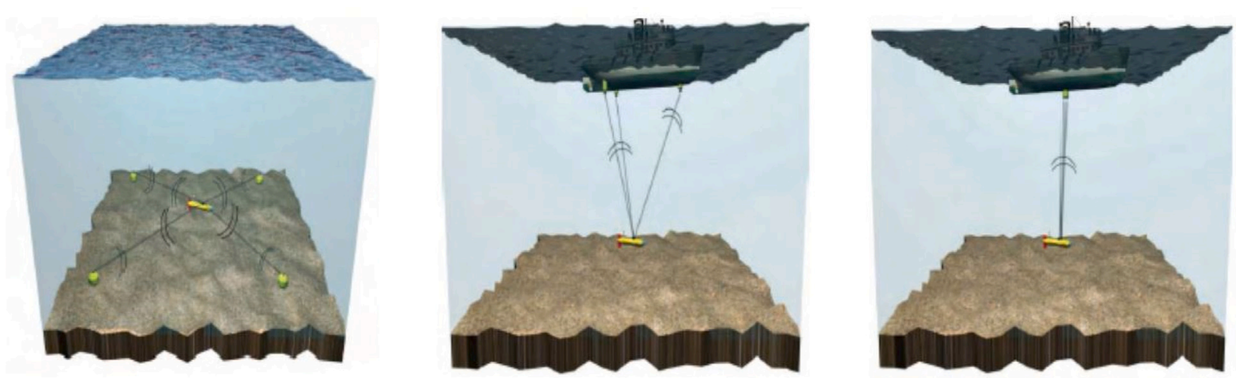
Fig. 9. Different types of sonar for underwater tasks.
Source: Tritech.

(a) DTR-100Z-ADF developed for ROVs and AUVs by Teledyne.

(b) SV-14IR Underwater camera developed by Shark Marine Technology.

(c) The Micro IP Camera from Reach Robotics.

Fig. 10. Some examples of optical cameras developed for underwater applications.
Source: Nautic Expo, Reach Robotics



(a) An example of LBL System.

(b) An example of SBL System.

(c) An example of USBL System.

Fig. 11. Different underwater positioning system.
Source: Xiang et al. (2009).

Table 1
A summary of sensors for robot's underwater perception at different ranges.

Range	Sensing modalities	Available data	Robot perception capability
30m – 10km	USBL, Acoustic modems	Range, heading	Approach the target using heading information
5m – 30m	Sonar	Acoustic Data	Detection, tracking and recognition using motion features
1m – 5m	Sonar, Camera	Acoustic Data, Optical Data	Detection, tracking and recognition using data fusion approaches
<1m	Sonar, Camera	Acoustic Data, Optical Data	Pose estimation, 3D point cloud reconstruction

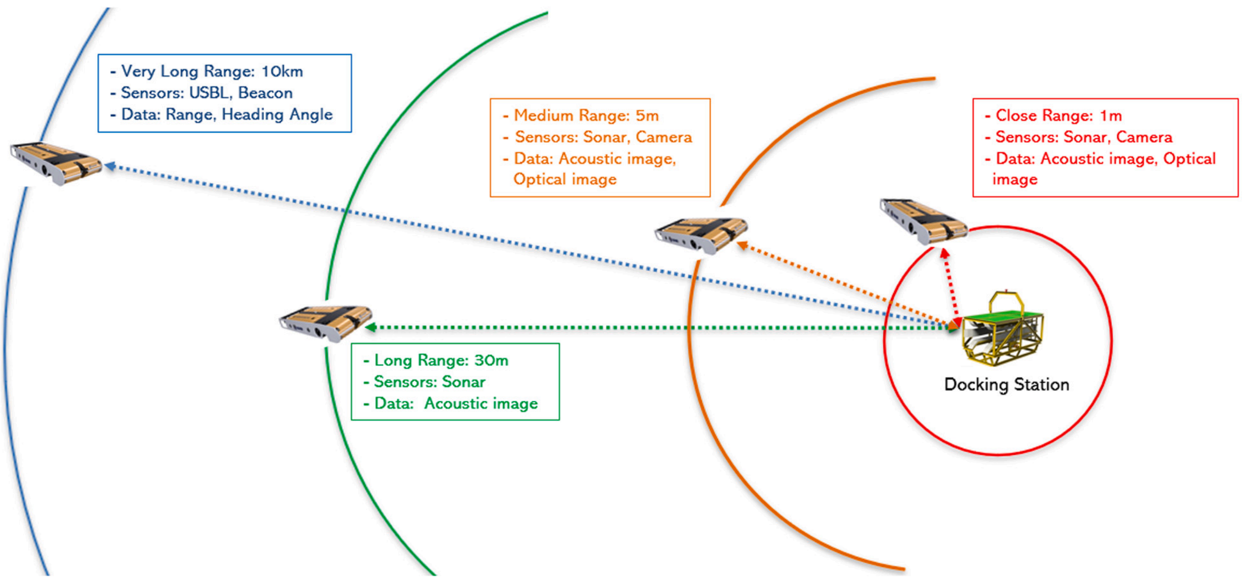


Fig. 12. Automatic docking process demonstration with different ranges.

3.1.1. Data preprocessing

Data preprocessing deals with image quality enhancement, removing or reducing interference, and speckle noise to increase the signal-to-noise ratio (SNR). Multiple approaches have been adopted to improve sonar images' quality, including using spatial domain or image domain information such as interpolation techniques including nearest-neighbor interpolation, bilinear interpolation, and bicubic interpolation. Median filtering (Lane and Stoner, 1994) is also a widely used method for enhancing sonar images. While time-domain approaches are simple to implement, they tend to smooth out edges due to the averaging process. Therefore, transform domain approaches such as Fast Fourier Transform (Lane et al., 1998a), Discrete Wavelet Transform (Sharmila et al., 2013; Piao et al., 2007), Stationary Wavelet Transform (Ravisankar et al., 2018) or a fusion of both DWT-SWT (Rajamohan et al., 2018; Demirel and Anbarjafari, 2011) are proposed for improving sonar image quality. Wavelet-based methods are more effective in preserving image details since they can decompose the original image into different frequency sub-bands, such as high-frequency and low-frequency. The enhancement techniques can be performed separately on these sub-bands. Similarly, Wavelet information can be combined with other domains for hybrid schemes. Some typical examples are combining Wavelet Domain Processing with New Edge-Directed Interpolation (NEDI) (Liu et al., 2015), Wavelet Transform with Karhunen–Loeve (K–L) Transform (Priyadharsini et al., 2015), or Wavelet transform with Discrete Cosine Transform (DCT) (Lama et al., 2016). Machine learning-based methodology for sonar image enhancement is also an active research area. Among them, image super-resolution based on sparse coding (Rajamohan and Rajendran, 2019; Ma et al., 2011) and image enhancement using Generative Adversarial Network (GAN) (Sung et al., 2018b; Rixon Fuchs et al., 2019) are extensively studied in recent years. Image enhancement is an ill-posed problem where the algorithm must calculate the intensity values of unknown pixels given the noisy image information. GAN is particularly useful for dealing with this problem. A GAN's structure highlights the deployment of two sub-network, including a generator and a discriminator. The generator network enhances a noisy image and produces a high-quality result. On the other hand, the discriminator will determine if a generator's output is an actual high-resolution image or just a synthetic one. The process will end when the discriminator cannot differentiate the generated images from the real ones. Sparse coding-based approaches (Wu et al., 2018; Park et al., 2019) can also recover high-resolution images from low-resolution images using sparse representation. These methods usually require training two dictionaries from high-resolution and low-resolution image patches. While machine learning-based methods can provide favorable results, they usually require high-quality datasets for good training results.

3.1.2. Object detection

Object detection in long-range contexts focuses on separating underwater objects from the environmental background and noise. Long-range objects appear very small in sonar images with low resolution. Also, object appearances change significantly between frames due to the view angle changes between objects and AUVs' sonar. Thus, detecting the targets using standard geometrical features such as the object's shape and size is challenging. Different kind of methodologies has been developed for this problem. For example, interframe difference (Weng et al., 2010; Chantler and Stoner, 1997) is a simple method to extract the moving objects from the static background. While the method is easy to implement, it not only reduces the object details but also cannot detect slow-moving objects and is not very effective against random noise. Background elimination is another method that estimates the background model using information from multiple frames. This method attempts to estimate the background model using information from multiple frames (Cui et al., 2018). A probability value of each pixel is calculated to determine if it belongs to the foreground or background. Objects can also be segmented from the background using image processing techniques such as adaptive thresholding (Ji et al., 2021), Otsu's method (Yuan et al., 2016; Modalavalasa, 2012), or double segmentation (Petillot et al., 2001). These methods are not very effective in the case of noisy environments where there is a big overlapping in intensity distribution between the noise and the objects. Frequency approaches can also be used to assume that the objects are moving with different frequencies compared to the noisy components and the static background. For example, Lane et al. (1998a) proposed calculating Fast Fourier Transform for a sequence of sonar frames. Frequency thresholding can then be applied to separate the dynamic objects from the static background. Constant False Alarm Rate (CFAR) detectors (Gandhi and Kassam, 1988) are also very popular in sonar image object detection. These CFAR-based methods assume that the background clutter or the moving objects follow a particular probability distribution such as Rayleigh (Maussang et al., 2007), Gaussian (Gao et al., 2009), K-distribution (Abu, 2020), or Weibull (Li et al., 2017). An adaptive threshold can be derived from the reverberation noise in the background using these noise models. The pixel under test is compared with the threshold for object detection. If the noise model is accurate, the CFAR detector can guarantee the objects' detection despite the object shapes. Different CFAR-based approaches have been developed with different adaptive thresholding strategies for trade-offs. While cell averaging CFAR (CA-CFAR) (Acosta and Villar, 2015; Weiss, 1982) can provide optimal results in uniform clutter environments, its performance quickly degrades in the presence of interference. Greater of CFAR (GO-CA-CFAR) (Kronauge and Rohling, 2013) performance is consistent for the clutter transition region, but it fails to detect two closely spaced targets. On the contrary, the smallest cell averaging (SO-CA-CFAR) (Yan et al., 2015) can detect two closely separated targets. However, the detection quality reduces significantly in the case of more than two close targets or there are targets in both leading and lagging windows. Finally, learning-based technology is also considered for solving this problem. However, while these methodologies are very efficient in detecting objects in images, they face a big challenge in long-range scenarios due to the extremely low resolution of the targets, which greatly reduces the number of features extracted from the images. As a result, machine learning-based methods are generally applied to detect underwater objects using optical camera data or high-resolution sonars in close-range scenarios. However, there are still some attempts to use neural networks to train a set of geometric, statistical, and textured-based features such as area size and contrast, shape moment-based features to detect the low-resolution objects (Perry and Guan, 2004; Jie and Pingbo, 2020).

3.1.3. Object tracking

Object tracking is the technique of associating or matching detections between consecutive frames. In long-range perception settings, several factors complicate the tracking problem. One of the critical reasons is the lack of tracking features of these low-resolution miniature objects. Also, there is no consistency in objects' appearances between frames due to the relative positions and orientation changes between the objects and the AUV's sonar. The robustness of the tracking algorithm can also degrade significantly due to occlusion effects when multiple neighboring observations merge or split. Furthermore, detections can also disappear and return randomly because of the shadowing effects or sidelobe pickup at short ranges effects. Finally, since sonar imagery suffers terribly from noisy turbid environments despite denoising algorithms, it is difficult for the tracking algorithm to differentiate between the small objects and the surrounding noisy components. Consequently, many research projects decided to replace traditional geometrical features with other features such as motion estimates for predicting objects' positions in consecutive frames during the tracking process. Optical flow-based methods naturally become the technique for tracking multiple objects in sector-scan sonars (Lane et al., 1998a,b). These methods can calculate the motion vector for objects using the recorded sonar data in consecutive image frames. While they can provide motion information at pixel-level accuracy, optical flow-related techniques usually require extensive computational power that small AUVs cannot afford. Two different categories of filtering algorithms subsequently solved the problem. The first set of algorithms assigns an independent single-target stochastic filter for each object in the scene. The tracking paths can be generated by combining the prediction data with the measurement data for each target. These algorithms usually require accurate sensor measurements to improve the data association process. Tracking algorithms using Kalman filter (Ruiz et al., 1999), extended Kalman filter (Petillot et al., 2001) or Particle filter (Doucet et al., 2003) are the most common representations. For example, different versions for object tracking with Kalman filter were developed for tracking moving objects in the open water column (Trucco et al., 2000), static objects lying on the seabed (Quidu et al., 2010), both dynamic and still objects (Karoui et al., 2015), or for map matching (Mallios et al., 2014). While tracking the underwater object in long-range scenes with the Kalman filter showed promising results, there are still some limitations. The state transition and observation models of tracking underwater objects can be extremely non-linear, breaking the Kalman filter assumption. Tracking with Particle filter (Zhang et al., 2020) is proposed for solving this problem. By representing the posterior density of the object state using many particles, the Particle filter can improve the object position tracking accuracy using the sonar measurements. However, the algorithm also suffers from high computational complexity. Recently, some research has been trying to solve this problem, such as using particle swamp optimization (Wang et al., 2018). The second set of algorithms involves using multiple-target filtering instead of single-target filtering to recursively estimate the number and state of multiple underwater targets given multiple sonar observations. These algorithms are very robust in long-range scenarios with false alarms and miss-detections by propagating the target's intensity in temporal dimensions instead of the total multi-target posterior density. Many results are obtained with performances that vary depending on clutter, noise, and model uncertainties. The sequential Monte Carlo PHD filter (SMC-PHD) (Vo et al., 2003) not only suffers from high computational particles but also from the added inaccuracy from clustering techniques to estimate object states. Gaussian Mixture PHD (GM-PHD) (Vo and Ma, 2006) fixed the inaccuracy problem by replacing the clustering algorithm with Kalman filter equations. However, the approach constraints the tracking problem with linear dynamical models. A solution for this problem is to implement the non-linear version of the Kalman filter, such as the Unscented Kalman filter (UK-PHD) (Melzi and Ouldali, 2011) or the Gaussian Particle Implementations of the PHD filter (Granstrom et al., 2012).

3.1.4. Object classification

Finally, the object's type can be realized by analyzing the tracked data. Different methodologies have been proposed for this task. Intra-frame features (derived from a single scan) can be extracted for classification. A typical example is to use a feature set including geometrical shape, size, and intensity level proposed by Lane and Stoner (1994). Thresholding values are then calculated using these features for classifying objects. Similarly, Dos Santos et al. (2017) introduced a set of 10-dimensional feature vectors to be extracted from each detected object in the scene. K-nearest Neighbor, Support Vector Machine, and Random Trees are three classifiers used for comparing purposes. While these proposed methods are straightforward, they usually suffer from low-resolution object intra-frame features and cluttered backgrounds in long-range contexts. Chantler and Stoner (1997) improve the feature selection process by introducing inter-frame features or temporal features (derived from a sequence of scans). Dai et al. (1995) further improved the method by combining intra-frame and inter-frame features to identify the correct class of the object with a linear discriminant function. Inter-frame features are not only discriminative between different objects but also between objects and noise. Therefore, the combined feature methods are very robust to noisy environments. Typical temporal features are statistical measures such as mean value, mean contrast, and variance. While hand-crafted features have shown promising results in long-range object classification and recognition, they are time-consuming, unoptimized, and require expert knowledge. Therefore, deep learning-based methods have recently become the new research trend for underwater object classification. Again, the major challenge for applying machine learning techniques is the availability of standardized datasets. Those available provide minimal options in the number of samples, object types, or object resolutions. For example, Fuchs et al. (2018a) combines transfer learning with a limited amount of training data for sonar object recognition. A Convolutional Neural Network (CNN) model is trained for categorizing five classes: fish, hull, pole, stone, and swimmer. The results are compared with other learning-based methods and classical hand-crafted feature approaches. Phung et al. (2019) proposes a deep learning architecture consisting of a CNN and a hierarchical Gaussian Process classifier. They solve the data scarcity problem using a Generative Adversarial Network to generate extra sonar snapshots. Some researchers also focus on using deep learning technology to classify objects lying on the seabed for different underwater tasks. For example, Williams (2016) introduces using CNN for training a system to classify objects deployed on the seafloor, including mine-shape objects, calibrated rocks, and other artificial objects. This work unfolds the immense power of deep learning in long-range object classification with a big jump in performance compared to traditional hand-crafted methods. Other similar projects include using deep learning for mine inspection proposed by Denos et al. (2017) or using a pre-trained Mask R-CNN for boulder object segmentation and recognition (Christensen et al., 2021). We summarize our review for long range object perception in Table 2

3.2. Medium range object perception

Even if the AUVs can track and recognize multiple objects at a long distance, they still need to move closer to the target to verify the classification results or to collect more data for other purposes, such as object inspections or pose estimations. There are mainly two sensing modalities for the robots to work with in medium-range scenarios, including optical-based sensors and acoustic-based sensors. This section will summarize articles addressing fusion-based methods using data from multiple perception devices and sensors. Fusion-based methodologies combine data from various "close-range" and "long-range" perception devices, where each device complements the other. However, the available data can no longer qualify for both ranges these devices were initially designed for. Instead, they qualify for a range that is comfortably in between both. The following are recent articles that strongly align with this concept.

The primary challenge of medium-range object perception is the issue of data fusion. For example, images obtained from a FLS and that of an optical camera do not see a given object in the same way. Hence, more than a mere superimposition may be required to take advantage of both imaging modalities. Therefore, due to the extra effort, such fusion methods are often relegated to navigation instead of perception.

Most articles combine various sonar devices like Profile Sonar (PS) with FLS. Some combine optical cameras with other sensors like an IMU (Inertial Measurement Unit) for improved calibration, thereby improving its range, while some use it with sonar for similar purposes. The selected articles can be broadly classified into two categories, fusion with different types of Sonars and fusion amongst different types of sonars only and fusion of other sensors, including sonars.

3.2.1. Fusion-based methodologies using sonars only

Different types of Sonars can be combined for improved perception.

Different sonar devices. Joe and Yu (2018) came up with a combination of an ALMS and a MSIS to realize the height of the observed objects. The proposed perception workflow consists of three steps, including point cloud data generation from both devices, scan matching of both point clouds, and slope and sensor drift correction. Joe et al. (2019b) recommends a combination of a PS, which provides reliable data in the vertical plane, and a Forward Scanning Sonar (FSS), which provides reliable data in the horizontal plane for 3D mapping. The PS mainly provides elevation information to complement the FSS data. While the data obtained from a PS can be of a larger range, it is narrower than that of an FSS. The former issue is used to build an occupancy grid in which data fusion effectively carves out areas where no object exists. In contrast, the latter issue is used to construct a particle filter since the data acquired in a PS will only be visible in an FSS after some time. Joe et al. (2019a) analyses the possible positions of the Mechanical Scanning Imaging Sonar (MSIS) to work effectively with an Acoustic Lens-based Multi-beam Sonar (ALMS). They identified three positions concerning the AUV, as shown in Fig. 13.

- *Position i* is fixed at the bottom at an angle of 60°
- *Position ii* is fixed perpendicular to the bottom
- *Position iii* is the same as *position ii* but with the ability to relatively rotate to the yaw-axis

Position i provided a view like ALMS, but with poorer quality. *Position ii* provided height information but was not useful due to the narrow field of view. Therefore, *position iii* helped extend the field of view by rotating the MSIS itself, which could complement the data obtained by an ALMS. Using similar concepts, Joe et al. (2020) explored the combination of an FLS and PS, with the major difference being the use of Monte-Carlo experiments to correct the FLS readings, which uses the probable elevation angles as particles for weight computation. Joe et al. (2021) studied the fusion of an FLS and a PS for 3D point cloud reconstruction in which line-based feature matching and principal component analysis were applied to the generated sub-maps.

Multi-frequency sonars. Tamsett et al. (2016, 2019) took a step further by suggesting a custom triple-frequency side-scan sonar system (114, 256, and 410 kHz) to obtain colorful maps of seabed in Scotland. Using multiple frequencies apart from one not only provides vibrant maps due to various levels of absorption and backscattering of sound waves, but it also reveals about 2.5 times more information about the seabed based on information entropy than a single frequency. The amplitude of reflected signals is converted to a value in the negative BGR color space with simple mathematical equations. Additionally, using a user-supervised texture mapping approach, elements of a map can be classified based on different objects, for example, sand and rock types. We summarize our review for sonar-only fusion-based methods in Table 4

3.2.2. Fusion of other sensors, including sonar

A more promising and reasonable trend is the use of multi-modal perception methods where one modality complements the other. Some of the articles focus on fusing sensors, sometimes excluding Sonars.

Sensor fusion without sonar. Perception leading to navigation and mapping elaborates on sensor fusion algorithms for calibrated sensors, leading to better perception and navigation tasks. Jakuba et al. (2010) proposed using optical cameras coupled with sensors like magnetic compasses and depth sensors to improve navigation and hence 3D reconstruction of seabeds. The corrections derived from SLAM can be used to improve the dead-reckoning accuracy of magnetic compasses while also helping validate vehicle depth measurements. As AUVs continue to be used in various water conditions, cameras must be constantly calibrated to obtain helpful information other than video feed. Anwer et al. (2017) uses a custom waterproof housing for a camera which accounts for the various refraction effects in water. The camera here is a combination of optical and near-infrared cameras which provides color and depth of the perceived environment, which also helps in better point cloud reconstruction. Bongiorno et al. (2018)'s work focuses on combining hyperspectral imaging with optical stereo imaging for seabed modeling. The stereo camera can enhance the in-water hyperspectral imaging data using image co-registration. Thus, the system can mitigate the various attenuation effects that regular hyperspectral imaging faces. The resulting maps are of a spatial resolution of 30 cm. Additionally, using support vector machines, the maps can be further classified into elements like sand, coral, etc. Gu et al. (2019) introduces a Camera-IMU calibration method for Monocular Visual-Inertial SLAM. The IMU and camera are linked with the help of intrinsic and extrinsic calibration methods unique to any camera. Along with other calculations, this ensures full calibration once initialized. The technique helps in the extended use of cameras for applications like Simultaneous Localization and Mapping, despite changing environments. We summarize our review of fusion-based methods for sensor calibration in Table 3

Table 2
References for long range object perception section.

Reference	Tracking technology	Performance analysis	Validation
Chantler and Stoner (1997)	Sector-Scan Sonar	Data pre-processing and object detection involve a combination of fast median filtering, thresholding, and region-growing. The object tracking is manually performed, while the object classification is achieved using a discriminant function with a set of temporal feature measures.	In-field Experiment
Lane et al. (1998a)	Sector-Scan Sonar	Dynamic objects are obtained by applying a Fast Fourier Transform method on a sequence of frames. The algorithm calculates the motion estimation of dynamic objects using optical flow methodology. Data association can be obtained by using a tracking tree to determine the best estimates of the object tracks.	In-field Experiment
Zhang et al. (2020)	Forward-Looking Sonar	Sonar data is pre-processed using median filtering and region-growing segmentation methodologies. An optimal set of features is determined using a combination of GRNN and search procedures. Object tracking is performed by combining an improved Gaussian Particle Filter with the feature set. Object classification can be obtained using a GRNN.	In-field Experiment
Quidu et al. (2012)	Forward-Looking Sonar	Object detection is performed using a simple goodness-of-fit test. Object tracking algorithms utilize multiple Kalman Filters for multiple objects. The state equation takes navigational data as inputs. Data association is achieved using a validation test and a validation gate derived from the innovation term of the Kalman filter.	In-field Experiment
Petillot et al. (2001)	Multibeam Forward-Looking Sonar	Objects are detected using a double layer segmentation algorithm. At the same time, object tracking involves using Kalman Filter with a state vector composing the position, the area of the object, and their associated first and second derivatives. A path planning algorithm is introduced using a nonlinear programming technique based on a CSG (constructive solid geometry) representation of the obstacles.	In-field Experiment
Perry and Guan (2004)	Sector-Scan Sonar	Data pre-processing is performed by cleaning the imagery using temporal averaging with motion compensation. Object candidates are segmented out using a threshold value. A first-stage multi-layer perceptron system is trained to reject obvious false alarms. Object tracking is performed using a bank of Kalman Filters. A second-stage recurrent neural network is used to process feature sequences and improve the object detection result.	In-field Experiment
Karoui et al. (2015)	Forward-Looking Sonar	Object detection is performed using a hierarchical detection procedure to detect various target signatures using echo analysis and constant false-alarm rate (CFAR). Target tracking is achieved in the Cartesian coordinate using Kalman Filter. Object association is performed using the joint probabilistic data association filter.	In-field Experiment
Wang et al. (2018)	Forward-Looking Sonar	Introduces an adaptive particle swarm optimization (APSO) algorithm for tracking multiple underwater objects, achieving higher tracking accuracy and faster tracking speed than traditional methodologies such as Kalman Filter.	In-field Experiment
Dos Santos et al. (2017)	Forward-Looking Sonar	Object detection involves thresholding, pixel searching, and intensity peak analysis. A set of geometric features is extracted for each of the detected objects. Object classification is performed using three methods: Support Vector Machine, K-Nearest Neighbors, and Random Trees. The results show that K-Nearest Neighbors achieve the highest performance for the given set of features.	In-field Experiment
Fuchs et al. (2018a)	Forward-Looking Sonar	Data-preprocessing and object detection are the same as proposed by Dos Santos et al. (2017). Object classification is performed using transfer learning using ResNet50 architecture. The algorithm uses the ARACATI dataset from a marina environment to categorize five classes of objects. The results show that transfer learning can be used as a good classifier for underwater object classification.	In-field Experiment
Christensen et al. (2021)	Side-Scan Sonar	This work present an automatic boulder detection system using deep learning methodology. The feature of the input side-scan sonar image can be extracted automatically using a Feature Pyramid Network (FPN). The feature map is the input for a region proposal network (RPN). The region proposals are refined using the box and score prediction head. The network also contains a mask prediction head to perform pixel-level classification for proposed regions.	In-field Experiment

Sensor fusion with sonar. Bruno et al. (2015) combines optical stereo vision and 3D acoustic data in real-time using extrinsic calibration matrices. The resulting opto-acoustic data is robust to water turbidity effects. The data is represented as point clouds, hence algorithms like the iterative close-point algorithm is also possible, apart from the image processing steps required, e.g. denoising sonar data, conversion of images to point clouds etc. Lagudi et al. (2016) further validated this approach by real-time experiments. Raaj et al. (2016a) proposed a

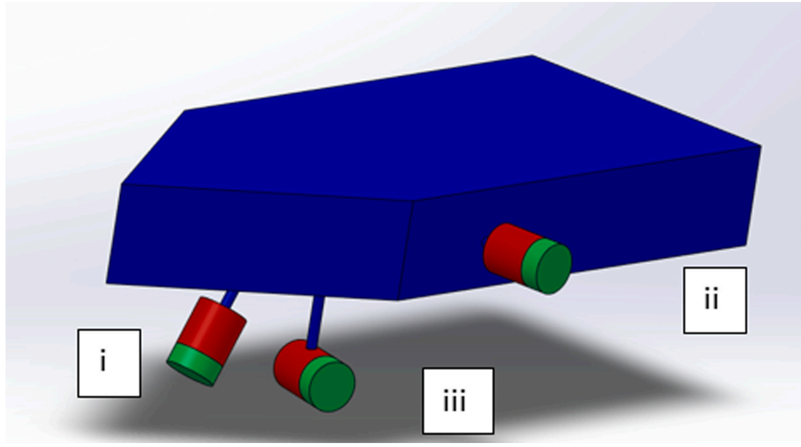


Fig. 13. Three Proposed Positions (Position 2 is shifted to the side instead of the bottom for better view).

Table 3
References for sonar-only fusion-based methods.

Reference	Tracking Technology	Performance Analysis	Validation
Joe and Yu (2018)	ALMS and MSIS (Sonar), Point Cloud Generation	Normalized Distribution Transform is used to combine data from short-ranged ALMS and Middle Ranged MSIS	Controlled Real-world testing
Joe et al. (2019b)	FSS and PS (Sonar)	Mainly for elevation angle, but a PS was used to obtain data that the FLS usually would not, uses concepts like particle filters and occupancy grids	Simulations and Controlled Real-world testing
Joe et al. (2020)	FLS and PS (Sonar)	3D reconstruction using an FLS and PS, using Monte Carlo Experiments	Simulations and Controlled Real-world testing
Joe et al. (2021)	FLS and PS (Sonar)	3D reconstruction using an FLS and PS, using concepts like PCA and Line Feature Matching	Controlled Real-world testing
Tamsett et al. (2016, 2019)	3 Frequency color Sonar	Fusing data of Sonar with different frequencies can give a better detail of the seabed, identification of individual objects using pattern matching gives better results than using single frequencies	Real World Testing

novel workflow for fusing data from an FLS and an optical camera using particle filters. The particle filter tracker utilizes extrinsic calibration between the camera and the sonar and vehicular odometry data, which is more straightforward than other data association methods. Its robustness is based on the fact that tracking can proceed even if either of the devices produces noisy data. Collings et al. (2020)'s preliminary work showed that using a Lidar with a side-scan sonar together was more beneficial than using either alone, as the combined data was able to differentiate the seabed better. By fusing single monocular camera data with single beam echosounder data, Roznere and Li (2020) can greatly improve the depth estimation to the point that it can be used for SLAM, image scaling and even image enhancement. After extrinsic calibration, the acoustic pulses emitted from the echosounder will be like a cone. When projected on an optical image from the camera, it will look like a circle. Similarly, Chemisky et al. (2021) proposed the Opto-Acoustic fusing algorithm underwater target mapping. Unlike optical cameras, the principle is that sonars have high ranges, but the acoustic images generated cannot be used to identify objects sufficiently. Using a seven-parameter Helmert transformation, they can calculate the acoustic sensor center relative to the optical camera used for data fusion. A multi-triplane spherical target is used for calibration as it is visible in visual and acoustic images.

The focus of several articles nowadays is on sensor fusion for navigation purposes. Works like the following have a primary imaging sensor coupled with several non-imaging sensors to optimize the underwater vehicle's trajectory. Rahman et al. (2018) introduced the method for combining a stereo camera, an IMU, a depth sensor, and a sonar for effective navigation using SLAM. The video feed from the camera is improved using a contrast-limited adaptive histogram equalization filter to work in challenging environments. Each sensor used in this setup has unique capabilities and issues. Hence a cost equation accounting for all possible errors like reprojection errors in cameras and drift in IMUs is used. Such a setup can help in better SLAM as it is targeted mainly for pose estimation and loop closure. Similarly, Cheng et al. (2022) in their article utilize the images obtained from an FLS to create point clouds in real-time, using image processing operations like adaptive thresholding based on averages and distance. This reduces the excess data obtained largely due to noise. The resulting point cloud data is fused with sensors like a DVL and IMU to perform simultaneous localization and mapping. Sadjoli et al. (2021) is an article that uses an orthogonal arrangement of two FLS devices for imaging. This is used for object recognition and pose estimation, which is further discussed in the following section. While works similar to this, whereby a fusion of imaging sensors is used for further characterization of targets, it can also be easily used for improved localization. Additionally, the use of sonars with multiple frequencies but in a side-scan configuration shows the usefulness compared to single-frequency data. However, this approach has yet to be explored in forward-looking sonars, which are typical of higher frequencies, up to 1.2MHz compared to a maximum of 480 KHz used in Tamsett et al. (2019)'s work.

Table 4
References for fusion with other sensors, including sonar.

Reference	Tracking technology	Performance analysis	Validation
Jakuba et al. (2010)	Camera, Magnetic Compass, Depth Sensor	Visual Augmented SLAM helps in validating magnetic compass and depth sensor readings, resulting in better 3D reconstruction	Real-world testing
Bongiorno et al. (2018)	Spectrometer and Optical Camera	Spectroscopy and optical images are used to map and differentiate (with the help of machine learning) different elements of a seabed	Real-world testing
Bruno et al. (2015) and Lagudi et al. (2016)	Stereo Camera and 3D Acoustic Camera	Data fusion using extrinsic calibration matrices which helps in producing detailed point-clouds	Controlled real-world testing
Collings et al. (2020)	LIDAR and SSS	LIDAR and SSS readings for observed seabeds were complementing each other, prompting further investigation into methods like SLAM	Real-world testing
Raaj et al. (2016a)	FLS and Camera	A particle filter is used for a computationally inexpensive and reliable fusion of an optical camera and FLS after calibration	Controlled Real-world testing
Joe et al. (2019a)	ALMS and MSIS (Sonar), Point Cloud Generation	MSIS at the bottom of the ROV in panning motion provides the best height information for ALMS to complement	Controlled Real-world testing
Gu et al. (2019)	Camera, IMU	Camera IMU integration done, such that calibration needs to happen only once in the air. Done to counter refraction effects	Controlled Real-world testing
Roznere and Li (2020)	Sonar, Camera, SLAM	Depth estimation is made possible using a combination of monocular cameras and Sonar (Single-beam Echosounder)	Controlled Real-world testing
Chemisky et al. (2021)	3D Acoustic Camera and Optical Camera	Development of a low-cost opto-acoustic navigation solution for calibration purposes. Spherical targets are most useful as they are visible in optical and acoustic images	Controlled Real-world testing

3.3. Close range object perception

The distance definition for a close-range operation varies based on different objectives. However, 'close range' objects usually refer to those at distances less than 1 meter from the AUV, where imaging sonars have the best resolution, and underwater optical sensors have the best visibility. Various types of underwater perception methodology within this range have been developed based on the sensor type. These close-range methodologies generally focus on either data quality enhancement, or maximization of extractable features and semantics from close-by objects for better perception results.

3.3.1. Underwater optical image enhancement

Underwater Image Enhancement (UIE) improves the quality of captured optical underwater images to improve visibility and effective range. Similar to underwater object detection, UIE can also be categorized into data-driven and non-data-driven methodologies. However, as of recent, data-driven UIE methods have achieved the state-of-the-art performance, usually leveraging synthetic data produced from Generative Adversarial Networks (GAN)s (Goodfellow et al., 2014) for their training, in addition to usage of existing real-world underwater image datasets (Li et al., 2020) combined with land-based datasets (Fabbri et al., 2018; Li et al., 2018; Wang et al., 2019; Guo et al., 2020; Hambarde et al., 2021).

Within the context of underwater object detection, recent underwater object detectors have integrated UIE methodologies into their model or architectures to enhance reliability in extreme or very turbid water environments (Li et al., 2021). While relatively new, this combination of fundamental object detectors with UIE architectures is generally seen as one of the more promising developments for optical-based object detection in the future (Shen et al., 2021b). Additionally, the proposals of alternative underwater image formation models such as Akkaynak and Treibitz (2018, 2019) to improve visibility could allow further improvement for UIE, which in turn should also improve optical-based underwater object detection in the future.

3.3.2. Methods using optical-based sensors

Optical-based RGB imaging sensors provide 2D colored images of objects that provide rich spatial and color features to be leveraged for object perception tasks, the most common of which is object detection. However, as explained in Section 3, optical signal attenuation cause reduced visibility, shape clarity, and noisy shape features, resulting in poor perception even at close ranges. The development of underwater optical-based perception has focused on addressing these visibility and clarity issues, leveraging the progress of land optical-based object perception, which utilizes similar RGB image inputs. The most common of these developed methods are underwater object detection methods which utilize specialized 2D image filters to extract key features of objects. Based on the type of filter used, these methods are categorized into two types: *non-data-driven methods*, and *data-driven methods*. Additionally, several types of underwater optical perception systems utilize external landmarks to highlight target objects' key features better.

Non-data-driven object detection. The feature detectors within this category are usually developed using more theoretical mathematical understandings of underwater environmental effects on light attenuation and image quality. Most notable example of these are detectors that extract region of interest (ROI) based on image analysis that incorporate these environmental equations (Chen et al., 2017; Shen et al., 2021a). Another method within this category focuses more on detecting specific objects based on a priori knowledge of the objects' model or shape that can be expected within a scene (Park and Kim, 2016). While proven to be accurate, reliable, and not require a lot of a priori data, this category of methods has the main limitation of being effective mainly within the apriori environment or objects that were used to model the detector. The effectiveness of these methods is slightly reduced when used for environments or objects that were not considered or modeled during the development of the detector (Shen et al., 2021b).

Data-driven object detection. In this category, the weights of the feature detectors are mainly developed from samples of data images collected within the expected underwater environment. This approach provides the advantage of usually producing detector models that are more generalizable and robust to noise, allowing more reliable usage in many areas. Following the marginal success on perception for land-based applications, most state-of-the-art detectors have utilized deep learning-based models to achieve reasonable success for several underwater applications, such as robot convoy (Shkurti et al., 2017) and close-range localization (Raaj et al., 2016b) with reasonable success. However, while more generalizable, these data-driven methods usually require many training samples to achieve models with adequate performance, which may only be practical and feasible for some applications and underwater environments. Additionally, it is still noticeably insufficient to address the visibility limitation of optical-based underwater perception, requiring further processing of data to achieve more reliable results in exceedingly turbid water areas (Shen et al., 2021b).

Usage of external landmarks. The visibility limitation of optical sensors, especially in turbid waters, means that more complex, high maneuverability applications like garage docking, direct object detection, and pose estimation remain a significant challenge, even with current UIE methods. Recent optical-based object detection alleviates this issue by attaching high-visibility landmarks on target objects, such as LEDs, specialized grid patterns, or beacons, to be trained as the highlighted features by detection and tracking methods (Yahya and Arshad, 2017; Nakamura et al., 2018; Singh et al., 2020; Ren et al., 2021). However, turbid water optical-based object detection independent of additional landmark features is still a preferable alternative in practice and thus remains an active area of research.

3.3.3. Methods using acoustic-based sensors

Acoustic-based sensors, especially multibeam FLS sonars, are the most commonly used sensor for underwater perception based on their proven reliability due to low signal attenuation. However, as described in Section 2, these acoustic sensors have low resolution and sparse spatial information. These limitations mean object shapes have low clarity or high ambiguity even at close range, reducing perception effectiveness. Research on acoustic-based perception methodologies has focused on addressing resolution and spatial information issues. The research problems include sonar image resolution enhancement, inferring critical spatial information from other features within the sonar images, or fusing data from multiple sonars to reconstruct or improve the quality of features.

Close range resolution enhancement of sonar images. Sonar images have lower resolution due to the acoustic signal bandwidth limitation and sonar array dimension. Hence, several sensing methodologies are focused on enhancing the image resolution to improve the clarity of object features to enable more efficient usage for perception-based processes such as object detection and classification. For close-range operations, more traditional methods were attempted, such as those based on back-projection, sparse representation (Kumudham R., 2019), and signal compressing (Andreas Gällström, 2019). However, most current state-of-the-art methods utilize deep learning to perform this resolution enhancement, leveraging their development for super-resolution on land-based images. Examples of these methods include using ResNet-based neural network (Sung et al., 2018a) and GAN (Sung et al., 2018c; Hua et al., 2021). The GAN models, in particular, show state-of-the-art results for this category of methodologies.

Object detection from single sonar view. While the image resolution of sonar is still relatively low, sonar hardware and computer vision method developments have allowed effective object detection with sonar images. Similar to the development of optical-based underwater object detection, sonar image-based object detection also follows the development object detection methods for land-based ones, from those dynamically creating filters from echo or background features within the sonar images (Galceran et al., 2012), to more recent methods that train neural networks to perform the object detection (Kim and Yu, 2016; Valdenegro-Toro, 2016; Fuchs et al., 2018b; Lee et al., 2019). The neural network-based methods show a recent trend of performing transfer-learning (Tan et al., 2018) of networks pre-trained on land-based images to fit the domain of underwater sonar images better. These methods enable faster development of models without requiring as much sonar image training data, which is expensive and time-consuming to obtain.

Reconstruction of 3D spatial information. Understanding the 3D spatial information of objects within a scene is most useful for environment mapping and navigation tasks, crucial for applications requiring high-risk spatial maneuvers in close ranges, such as garage docking (Sadjoli et al., 2021). Several 3D scanning sonar models with specialized receiver arrays, such as Echoscope (2021) have been developed, capable of preserving the complete 3D spatial information from reflected acoustic signals. However, these models utilize proprietary technology and are usually very expensive, limiting their usage in commercial and research settings. This has prompted research on methods using more readily available 2D imaging sonars – such as multibeam FLS – common on AUVs to infer 3D spatial data. These methods focus on recovering the depth information loss from sensor data, as explained in Section 2.1, to get the full 3D spatial information in the scene.

1. Reconstruction using a single sonar image: The first category of methodologies attempts to perform a 3D reconstruction of the scene using all the information available within just one sonar image. The most notable examples within this category are those based on space-carving methodology (Aykin and Negahdaripour, 2017), which infer the missing 'depth' information of objects from their respective acoustic shadows within the sonar images (Kim et al., 2019, 2020). While showing successful results, relying on acoustic shadow information means these methodologies heavily assume objects to be incident on relatively large surfaces, such as the seabed. This dependence prevents their practical usage for floating underwater objects or shallower areas where the quality of acoustic shadows may not be significantly affected by reflection signal noise.

2. Reconstruction with multiple sonars: The second category of acoustic-based underwater 3D reconstruction methods commonly researched is the fusion of data from multiple sonars at different orientations (Negahdaripour, 2020). Not dependent on specific landmarks or acoustic shadows within the sonar images, these methods have the advantage of being effective for objects in all states, including floating objects. However, determining the most optimized sonar configuration and multi-sonar calibration to perform 3D reconstruction are the main challenges of this category. Negahdaripour (2018) addressed these challenges by performing epipolar geometry analysis on the inputs from multiple sonars, which determined that vertically orienting the sonars is the optimized configuration for 3D reconstruction. Works done in John McConnell and Englot (2020), McConnell and Englot (2021) would later prove this analysis, achieving state-of-the-art 3D reconstruction using a realized implementation of orthogonally-oriented multi-sonar fusion (OMSF), effective for underwater mapping of relatively simple-shaped objects with repeating patterns. As summarized in Fig. 14, OMSF first uses a relatively simple constant false alarm (CFAR) based feature extraction (El-Darymli et al., 2013) to extract feature pixels from image inputs from orthogonally oriented pairs of FLS sonar. Then, norm-based pixel association is performed within a

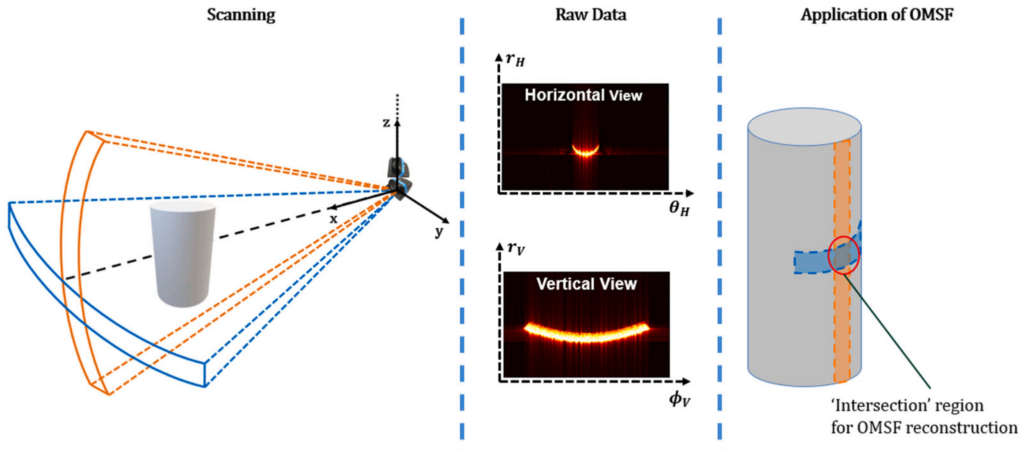


Fig. 14. Summary of OMSF methodology.

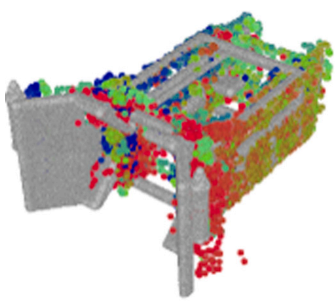


Fig. 15. Sample reconstruction result of a garage model by OMSF.

set intersection area, fusing information from the extracted features. Information from the associated features is then used to estimate the point and projected back onto 3D space. Preliminary results on OMSF testing in other works (Sadjoli et al., 2021) have also shown the method to work relatively well for more complex object shapes at varying distances (see Fig. 15). However, continued work in McConnell and Englot (2021) noted the limited effective reconstruction area of OMSF due to the restricted area during the feature association step. Additionally, OMSF has a biased tendency to reconstruct points onto solid surfaces due to the association method used. This means that with no modifications, reconstruction by OMSF would not work well on 'hollow' objects made up of thin frames, such as underwater trusses.

3.3.4. Methods using electromagnetic sensing

Extreme environmental conditions, such as shallow turbid waters, reduce the effectiveness of conventional underwater optical and acoustic sensors. Limited visibility and high environmental backscatter noise are tough to overcome due to the inherit limitation of modalities used. This limits these sensors' performance for tasks such as localization and object recognition for these extreme environments. Some research areas have attempted to overcome these issues via the development of underwater perception using more unconventional methods of electromagnetic sensing based on observations of some animals perceiving the environment using either electric or magnetic fields. This category of sensing methods can be divided into two closely related groups: electric sensing and magnetic sensing.

Electric-based sensing. The basic principle behind this sensing method is to utilize electric-sensitive receiver probes to detect and measure changes of electric fields within the probe vicinity, where measured changes can then be used to estimate information such as object location or shapes (Solberg et al., 2008; Bai et al., 2012, 2015; Boyer et al., 2015; Lebastard et al., 2016; Bazeille et al., 2020). Due to relatively recent development, sensors capable of facilitating these methods have yet to mature and are still in ongoing development. Thus far, two types of electric sensors and control schemes have been developed: active and passive, the main difference being that active electric sensing uses additional emitter probes that actively emit their electric fields (Bai et al., 2012).

The simplest type of active electric sensing for underwater perception can be seen in the work done by Solberg et al. (2008), developing a particle-filter-based active electric sensing capable of performing localization of a conductive target object. However, the reliance on prior electric field mapping of the environment means it is not practical for usage in new areas. Subsequent developments remove this dependency with sensor probes for detecting phase or potential differences in the electric fields, providing more distinctive information for different object geometry and conductivity (Bai et al., 2012). This development significantly improved the feasibility of active electric sensing, which can now create an input profile based on the potential reading of the target object as the sensor is aligned with the object (Bai et al., 2015). The distinct profiles can then be used as features for accurate shape recognition and localization of the spherical target objects tested, performance of which hypothetically can be improved using supervised learning methods.

In contrast, the development of passive, reactive electric sensing instead utilizes sensing probes that read to the lines of electric fields emitted by objects within nearby area (Boyer et al., 2015), without requiring any prior electric field mapping. However, this reactive sensing design initially only provided sparse localization information and tested in a particular role of passively guiding underwater robots to a docking station actively emitting electric fields. Subsequent work in Lebastard et al. (2016) then provided additional electrodes and memory controller to the reactive sensing system, enabling depth estimation of an object relative to the sensor without usage. This additional depth estimation enabled reactive sensing proved to give a significant advantage for object shape and pose estimation as shown in Bazeille et al. (2020), providing highly accurate results without requiring prior information.

While proven successful for close-range underwater perception, several limitations are inherent to all types of electric-based sensing developed thus far. First, as noted in Solberg et al. (2008), Bai et al. (2012), Boyer et al. (2015), electric fields emitted by objects or emitter probes are dependent conductivity of target objects and water environment. While electric field changes due to object conductivity can be a valuable feature for object identification, the variance of water conductivity can lead to inaccurate readings, negatively affecting perception performance for both active or reactive electric sensing (Bai et al., 2012; Boyer et al., 2015). Such limitation can be problematic for usage by underwater robots in operations where traversal through different types of underwater regions in one process is required.

Magnetic-based sensing. Magnetic sensing utilizes sensors sensitive to changes in magnetic signals – usually magnetometers or magnetic gradiometers – to estimate information of objects within the environment. Similar to electric sensing, magnetic sensing can be done actively via active induction of magnetic fields from target objects or passively detecting any magnetic signatures from the environment. However, most perception methods have been developed mainly with passive magnetic sensing due to more practical implementation (Xiang et al., 2016).

Perception methods developed for magnetic sensing depend highly on the type of task being solved. For relatively simple applications, basic processing of detected magnetic fields is sufficient, such as the detection of buried underwater mines using magnetic gradiometers developed with specialized signal processing and analytical inference (Kumar et al., 2004; Clem et al., 2004). This application of buried objects can also be enhanced via fusion with acoustic sensor data, as shown in work done by Pei et al. (2010). Similar analytical methods successfully track underwater pipes and guide AUVs as shown in Xiang et al. (2016). The heading of the target cable is inferred through analytic calculation of data from mounted magnetometers which are then used to guide the AUV through specialized magnetic-based feedback and control loops. More advanced magnetic-based perception methods have instead focused on using magnetic gradient tensors (MGTs) for more complex tasks, such as the work developed by Hu et al. (2019). The method successfully performs nonlinear localization of multiple objects simultaneously using MGT data. Development of such methodologies are also parallel with more advanced magnetic gradiometers that focus on higher quality MGT measurements (Keenan et al., 2010), or novel active atomic magnetic sensors that estimate magnetic fields from measurements of atom excitation (Kominis et al., 2003; Deans et al., 2018).

Current problems for industry use. From an industrial perspective, the additional electric or magnetic field from active electromagnetic sensing may introduce unintentional magnetic interference, creating complications to the navigational equipment – such as Inertial Measurement Units and compass – typically found on industrial-grade underwater robots. This interference could potentially be limited by equipping specialized electric insulator materials on the underwater robot or target objects, such as the case in the docking application proposed in Boyer et al. (2015). However, such a solution would not be practically feasible as only some target objects encountered can be expected to be equipped with similarly electric or magnetic insulating materials.

Overall, electric and magnetic sensing show promising early development for use in underwater perception. However, practical issues with dependence on water conductivity for electric sensing and potential electromagnetic complications with navigational sensors mean further development of these sensing methods to overcome these issues is still necessary.

3.3.5. Methods using flow sensing

Similar to the development of electromagnetic sensing, the development of flow sensing methods is mainly motivated as another alternative sensing system to address the limitations of conventional underwater perception sensors. Flow sensing refers to the measurement of flow from the surrounding liquid within an environment and is based on the lateral line system (LLS) usually found on fishes (Zheng et al., 2020).

For underwater robots, flow sensing is enabled by the development of artificial lateral lines (ALLs) sensor arrays, usually used to solve localization of the robot itself, which has been used to improve feedback control of an underwater vehicle, as shown by the works of DeVries et al. (2015). This self-state estimation has also enabled more accurate trajectory estimation of robots, as shown in the result of Zheng et al. (2020).

The development of flow sensing is also currently being expanded to enable the sensing of objects at further ranges. Work in Abdulsadda and Tan (2013) first showed successful tracking of a moving, vibrating dipole source using analytical-based nonlinear estimation approach on ALL flow readings. This ranged flow sensing ability has enabled more complex applications, such as obstacle avoidance, as done in Li and Zhang (2020). With integration of machine learning techniques, works in Pu et al. (2022), Chen et al. (2021) have also shown improved potential of flow sensing for object detection and localization.

Overall, the feasibility of close-range perception with flow sensing is well established for tasks such as robot state estimation, object detection, and localization. However, it has a notable drawback of not providing shape information of detected objects, meaning in-depth object classification is currently impossible (see Table 5).

4. From perception to navigation, localization, path planning, communication, collaborative operations, and machine learning

ROVs and AUVs are the motivations for the progress in underwater research and industries. The development of robot perception systems complements other related areas such as robot localization, navigation, path planning, and joint operation. In this section, we briefly introduce these research fields in relationship with the underwater perception field to provide insights into future developments.

Localization and navigation in an unstructured environment such as the deep ocean remain challenging for many underwater robotic platforms. Positioning systems such as LBL or USBL are costly and can only help the automated systems to cover areas that are reachable to humans to set up the network of transponders. However, automatic deployment at unknown terrain or unreachable to humans is still a problem. One primary reason is the need for an accurate map for the robot's navigation system. A key technology for solving this problem is Simultaneous Localization and Mapping (SLAM). Many SLAM techniques have been developed in the literature to help AUVs to navigate in an unconfined environment without a priori map. However, only some systems are tested in a practical scenario. A critical requirement of these SLAM techniques is the available features in the surrounding environment for pose discrimination which poses a considerable challenge for underwater navigation research. The feature extraction step can be achieved by processing data from an optical camera or multibeam-sonar. A robust perception system can improve the accuracy of SLAM techniques significantly. However, the underwater environment usually lacks informative features, and even the most advanced perception system will fail in extremely sparse underwater contexts. One possible solution is to plan the robot's trajectory through affluent feature areas to re-calibrate the localization system. This solution also relies on the perception system to classify the environment into different categories, such as free, occupied, or unknown areas, to improve its semantic environmental understanding.

Path planning for underwater vehicles involves generating an optimized path from an initial position to the goal position using particular evaluation metrics such as path length, navigation time, total energy consumption, or path feature availability. Various path-planning algorithms were developed, such as artificial potential field, geometric model search, random sampling, and intelligent bionic methods. Underwater path planning is a complex problem due to external environmental factors such as surrounding obstacles, ocean currents, limited sensing devices, and lack of undersea terrain information. Research on underwater path planning also has to consider the robot's physical and motion constraints. Many path planning algorithms have been proposed and applied to underwater robotics, such as Artificial Potential Field, Shortest Path Planning, A-star, D-star, and Level Set Method. While path planning for terrestrial robots is quite a mature field, underwater path planning still faces some problems, including the inability to take into account physical environment conditions such as ocean currents, the inability to be applied to 3D environments, and

Table 5
References for close range perception section.

Reference	Sensor type	Analysis summary	Validation	Applications
Sadjoli et al. (2021)	Multibeam Sonar, Reactive	Provided short justification for using acoustic-based sensors vs. optical, and performed initial simulator-based controllable testing on OMSF method. Initial results seem promising but require more experiments and data for validation.	Simulation experiments	ROV Perception, Underwater SLAM
Shen et al. (2021b)	Optical RGB, Single or Multiple views	Review paper on the image formation methods for underwater optical sensors, and the methods that have been used alongside them.	–	ROV Perception, Underwater Object Detection
Chen et al. (2017)	Optical RGB, Single View	Non-data driven underwater object detection method developed base on mathematical modeling of the light attenuation underwater, using single-view optical sensor. Results are shown to be satisfactory, albeit within a limited amount of tested environments.	Simulation and Field Experiments	ROV Perception, Underwater Object Detection
Shen et al. (2021a)	Optical RGB, Single View	Non-data driven object detection method using prior calculation. Satisfactory results, however, development of prior may not be generalizable to all environments and not tested within turbid environments.	Simulation and in-field experiments	ROV Perception, Object detection and underwater image segmentation
Park and Kim (2016)	Optical RGB, Single view	Non-data driven object detection that uses apriori object model data to perform specialized object detection on specific object models. Satisfactory results, however, only tested in clearer waters and notably would work objects within apriori data	In-Field experiments	ROV Perception, Underwater Object Detection
Shkurti et al. (2017)	Optical RGB, Single View	Example application of underwater object detection used for a multi-robot convoy, where several popular object detectors (e.g., YOLO) are transfer-learned to adjust to the underwater domain. However, good results are not fully tested for turbid environments.	In-field experiment	ROV Perception, ROV Convoy
Raaj et al. (2016b)	Optical RGB and FLS Sonar	Object Detection and localization using fusion between optical RGB and FLS sonar. The close-range perception section focuses more on the example of CNN-based object detection being used for the optical RGB portion. The method shows good results, with some testing done vs. some turbidity levels. However, the turbidity results shows significant performance decrease for optical portion once turbidity increases significantly	In-field experiment	ROV Perception, Underwater Object Detection
Goodfellow et al. (2014)	–	First paper on proposing the idea and implementation of General Adversarial Network (GAN)	–	Data synthesis and Deep learning-based training
Fabbri et al. (2018), Li et al. (2018), Wang et al. (2019), Guo et al. (2020), Hambarde et al. (2021)	Optical RGB	Examples of using GAN for data synthesis and development of a UIE model. Good results with more recent papers showing higher levels of accuracy and performance.	In-field experiment	Underwater Object Detection, Image Enhancement
Li et al. (2021)	Optical RGB	Example of combining UIE into core DL-based underwater detection methods that have been attempted before	In-field experiment	ROV Perception, Underwater Object Detection
Akkaynak and Treibitz (2018, 2019)	Optical RGB	Example of alternative image formation model proposed that provides more accurate image representation for underwater environments. Current experiments show promising results. However, effective integration would require more specialized low-level hardware to be developed.	In-field experiment	ROV Perception, Underwater Image Enhancement
Yahya and Arshad (2017), Nakamura et al. (2018), Singh et al. (2020), Ren et al. (2021)	Optical RGB	Example papers of using external beacons to alleviate the visibility issue for underwater optical RGB-s, but more specialized for garage docking. Satisfactory results are shown in all the papers. However, the necessity of external hardware limits full practicality in scenarios where the garage is remote or turbidity removes beacons' effectiveness.	In-field experiment	ROV Perception, ROV/AUV Garage Docking

(continued on next page)

the lack of path planning for the multi-AUV system. The path planning problem is coupled with underwater perception and SLAM problems. These problems should be solved simultaneously.

Underwater communication is also an important technical aspect to be enhanced. The field is significant and always is the center of research concentration. The technology allows exchanging of data between various underwater components, including network nodes, ROVs, AUVs, offshore stations, and surface stations. There are three common types of underwater communication including acoustic, optical, and radio frequency (RF). While optical and RF can provide high datarate, they are

Table 5 (*continued*).

Kumudham R. (2019), Andreas Gällström (2019), Sung et al. (2018a,c), Hua et al. (2021)	Multibeam Sonar, Active	Example of GAN used for enhancing the images from the multibeam FLS sonars. All papers show good results and highlight the lack of good quality sonar image data.	In-field experiment	ROV Perception, Sonar Image Enhancement
Galceran et al. (2012)	Multibeam Sonar, Active	Non-data driven object detector for FLS images that performs detection based on the background and echo data of the FLS image. It shows good results but no longer gets state-of-the-art results.	In-field experiment	ROV Perception, Object Detection
Kim and Yu (2016), Valdenegro-Toro (2016), Fuchs et al. (2018b), Lee et al. (2019)	Multibeam Sonar, Active	Examples of usage of CNN for object detection using FLS images. Good results but not yet performed on many object types require more data.	In-field experiment	ROV Perception, Object Detection
Tan et al. (2018)	–	Paper on Transfer Learning for domain transfer of Deep-Learning based models	–	–
Echoscope (2021)	Multibeam Sonar, Active	Article for an example of the active 3D sonar profiler models. However too expensive and not as practical for integration with many AUV systems.	–	Active 3D Underwater Perception
Instruments (2000)	Multibeam Sonar, Active	Basic guide from SeaBeam on how multibeam FLS works.	–	Underwater FLS sonar imaging.
Aykin and Negahdaripour (2017)	Multibeam Sonar, Active	Introduced Space-Carving concept for Underwater Object Reconstruction from Multibeam sonar images. Fast and relatively accurate reconstruction methodology, however heavily dependent on the presence of background surface as methodology utilizes acoustic shadow for the main feature. Hence, may not be suitable for floating underwater objects (where optical cameras may not be usable)	Simulation and in-field experiments	ROV Perception, Underwater Mapping
Kim et al. (2019)	Multibeam Sonar, Active	Fast and satisfactorily accurate reconstruction results based on Space-Carving methodology with testing that expanded upon Negahdaripour's testing. However, still faces similar problems of background surface dependence.	In-field Testing	ROV Perception, Underwater mapping
Kim et al. (2020)				
Negahdaripour (2020)	Multibeam Sonar, Active	In-field tests to expand upon the epipolar geometry analysis done from OCEANS 2018 paper. Accurate localization results were shown. However, the method is only tested on simple smaller objects, and not full 3D reconstruction of a more complex object	In-field testing	ROV Perception, underwater mapping
Negahdaripour (2018)	Multibeam Sonar, Active	Assessment on feasibility and best configuration of multiple multi-beam sonar configurations for underwater reconstruction by analyzing the epipolar geometry of the different acoustic images produced by the different sonar configurations tested. Test results suggest the orthogonal sonar configuration to produce the best optimized epipolar geometry calculations for best reconstruction results	Simulation experiments	ROV Perception, Underwater SLAM
John McConnell and Englot (2020), McConnell and Englot (2021)	Pair of Multibeam Sonar, Active	Fast and accurate reconstruction result that is not dependent on the presence of a background surface, thus suitable for floating underwater objects as well. However, require multiple sonars and reconstruction only occurs within a limited intersection area between the multiple sonars.	Simulation and In-field Experiment	Underwater ROV perception, Underwater SLAM
El-Darymli et al. (2013)	Multibeam Sonar, Active	Constant False-Alarm (CFAR) method that is commonly used for feature extraction among sonar images.		Underwater Object Detection
Solberg et al. (2008)	Electric, Active	Developed active electric sensing to perform localization. Results are highly accurate, however method is noted requiring a prior electric field mapping of the environment, reducing feasibility for practical usage.	Simulation and lab experiments	Underwater localization

(*continued on next page*)

only suitable for short-range due to transmission attenuation. Acoustic can support a more extended communication range. However, it also suffers from limited signal bandwidth, strong environment attenuation for considerable distances, multi-path propagation, and channel time variations. An underwater network provides interactive communication between underwater components and allows data to be uploaded or extracted in real time. The continuous data flow can keep the system updated and minimize data losses. Robust communication is

Table 5 (*continued*).

Bai et al. (2012)	Electric, Active	Improved electric sensing from Solberg et al. (2008) by sensing for phase differences of electric fields, providing more distinct information on objects of different shape and material. However, water conductivity could still potentially effect result accuracy	Simulation and lab experiments	Underwater localization, shape identification
Bai et al. (2015)	Electric, Active	Application of Solberg et al. (2008) and Bai et al. (2012) for object localization and shape identification via introduction of active alignment algorithm, with good results	In-field experiments	Underwater localization, shape identification
Boyer et al. (2015)	Electric, Reactive	Proposed reactive electric sensing for the case of guiding robot to target underwater dock. Reactive nature means no prior electric mapping is needed. However, requires specialized insulator materials to be used on docking and robot, however proven to be successful	Lab and In-field experiments	Guidance for underwater docking, underwater localization
Lebastard et al. (2016)	Electric, Reactive	Improved reactive sensing from Boyer et al. (2015) by introduction of additional electrode to enable depth measurement capability. Potentially allows imaging of object shape for object identification, without requiring prior electric mapping	Lab experiment	Object shape and pose estimation, Underwater object identification
Bazeille et al. (2020)	Electric, Reactive	Continuation of Lebastard et al. (2016) to explore potential of proposed reactive electric sensing for object shape and pose estimation. Results showed high accuracy on proposed process.	Lab experiment	Object shape and pose estimation
Clem et al. (2004)	Magnetic, Passive	Details the signal processing used for a developed magnetic gradiometer, and the object detection schematic used when mounted onto an AUV to detect buried underwater mines.	In-field tests	Buried object detection and localization
Kumar et al. (2004)	Magnetic, Passive	Sensor details for Clem et al. (2004).	In-field tests	Buried object detection and localization
Pei et al. (2010)	Magnetic, Passive	Developed a fusion-based scheme to combine magnetometer and acoustic sensor readings to perform object detection and localization of buried objects.	In-field tests	Buried object detection and localization
Xiang et al. (2016)	Magnetic, Passive	Example application of tracking of underwater cables and automated AUV guidance using magnetometers	In-field tests	Object Tracking
Hu et al. (2019)	Magnetic, Passive	Developed method to use MGT to perform simultaneous nonlinear multi-object localization	Simulation and In-field tests	Multiple Object Localization
Keenan et al. (2010)	Magnetic, Passive	Development of magnetic gradiometer to provide higher resolution MGTs	In-field tests	Object detection and localization
Deans et al. (2018)	Magnetic, Active	Development of active gradiometer using atomic magnetometers	Lab tests	Object detection and localization
Zheng et al. (2020)	Flow Sensing, ALL	Developed methodology of using flow sensing to self-localize and estimate robot state in real-time	Lab tests	Robot self-localization
DeVries et al. (2015)	Flow Sensing, ALL	Developed and demonstrated use of flow sensing with ALL to improve closed loop feedback control of an underwater vehicle	Simulation and Lab tests	Vehicle control loop
Abdulsadda and Tan (2013)	Flow Sensing, ALL	Ranged object tracking of vibrating moving dipoles using ALLs	Lab Tests	Object tracking
Li and Zhang (2020)	Flow Sensing, ALL	Simulation study of using flow sensing for AUV obstacle avoidance	Simulation	Obstacle avoidance
Pu et al. (2022)	Flow Sensing, ALL and Pressure sensor	Integration of Neural Network to perform localization of objects at range using combination of ALL and pressure sensor readings	Lab tests	Object localization
Chen et al. (2021)	Flow Sensing, Pressure sensor	Simulation study to incorporate Neural Network to perform object detection on flow data obtained from pressure sensors	Simulation	Object detection

(*continued on next page*)

Table 5 (*continued*).

also essential for ROVs and AUVs. Effective communication greatly supports the interactions between the robot–robot and robot–human via information exchange, and it is a complement to the perception systems. While a perception system can independently help the robot to detect, recognize, predict and react to the operation of other objects in the surrounding environment, communication can provide the robot another channel to send commands, double confirm, or clarify any misunderstandings in joint operations. The underwater communication development focuses on topology design and implementation of Underwater Acoustic Networks (UAN) that can minimize energy consumption, reduce latency, enhance system security, and improve channel bandwidths. Acoustic and optical modems are also attracting enormous interest for improving their operation range and connectivity and reducing their cost, mainly from the transducer.

Underwater collaborative operations receive substantial interest from both academia and industry. The motivation originates from the demand to quickly collect massive data over a sizeable area. A group of underwater vehicles can also swiftly and collaboratively perform underwater missions such as seafloor surveying, territory inspection, search, rescue, or infrastructure maintenance and repair. Additionally, underwater cooperation could happen between divers and ROVs or AUVs, where robots become human partners in collaborative tasks. Joint operations require studying collaborative planning, navigation and control, cooperative world modeling, and collaborative skills. This research problem faces many difficulties comparing the terrestrial version due to the dynamic of the underwater environment and communication limitations. One such example is multiple AUVs cooperating to inspect an underwater infrastructure seamlessly. The task could be extended with the participant of a human partner. Both scenarios require the robots to be equipped with advanced cognitive systems, situation awareness for coordination, and cooperative control. Another exciting topic is robot control in communication losses. Such a scenario requires each robot in the group to rely on its advanced perception and control system to recognize and predict the intentions of the neighboring robot or human partners to safely and accurately perform the tasks.

Machine learning is the new motive for the development of underwater perception systems. Deep learning methodologies provide researchers with new opportunities to tackle classical problems in underwater perception, such as reverberation and speckle noise removal, underwater image enhancement, image resolution, object detection, object tracking, object classification, or multiple-view object reconstruction from a data-driven perspective. The main advantage of machine learning is the ability to automatically extract high-level features most appropriate for particular perceptions using a high-quality dataset. Machine learning can eliminate the ineffective and time-consuming processes of designing traditional handcrafted features. Machine learning also allows transferring of knowledge and models from other domains to be reused in underwater contexts. The feature benefits underwater robotics due to the need for decent datasets for underwater research. Other aspects of underwater robotics, such as underwater system modeling and control, benefit from learning-based methodologies. The main challenge in system modeling is to compute the optimal robot model with accurate motion states. Machine learning methodologies can estimate model parameters without using a prior mathematical form. The learning-based robot model parameter is directly extracted from the collected data. Different learning-based methods such as reinforcement learning, fuzzy logic, and optimal control are actively developed to deal with the dynamic and complex environment and vehicle hydrodynamic constraints. Machine learning is leading underwater robotics systems to achieve a higher autonomy level than ever.

For optical-based perception, further development and improvement of non-data driven perception methods first should be of higher priority, as these methods do not necessitate pre-training on extensive amount of data which may be expensive to get. This could be potentially very effective with increasing development of improved UIE using newer optical image formation models (Akkaynak and Treibitz, 2018, 2019) that can provide more accurate image results. However, newer hardware or processes supporting such new image models would also need to be developed in parallel. Once the non-data driven methods and newer image formation models are better established, the data-driven perception methods could then leverage data collected from these developments to then create more powerful data-driven models which can extract more detailed features and semantics from the newer data.

5. Conclusions and future developments

In the past half-century, developments in underwater robotics have allowed the technology to emerge from a purely military role to one that now finds widespread use in the offshore oil and gas industries, maritime search and rescue, oceanographic research, and environmental monitoring. The emerging field of deep-sea mining will push ROVs' capabilities to unprecedented levels. As with their terrestrial counterparts,families of autonomous underwater robots have been developed and arose from power sources, sensors, navigation technology, and communications advances. While initially aimed at research, AUVs are now being employed by the military, offshore industries, and other related fields. Underwater robots are the topic of an ongoing research effort that, combined with technological developments, will undoubtedly impart them with greatly enhanced capabilities in the future.

Although significant progress has been made in the underwater robotic field over the past decade, many challenges remain unsolved. First, in terms of sensing modalities, there are still limitations in sensing devices. Although new sensing modalities are being developed, such as electromagnetic sensing, bionic sensing, or underwater laser systems, optical cameras and acoustic sonar remain primary sensing modalities for underwater perception. As both sensors are constrained in resolution or perception range, an optimal sensing modality for underwater perception still needs to be addressed. Second, in terms of sensing algorithms, while machine learning is dominated in many research areas, their contribution to underwater perception is still unmatched by its immense potential. The reason is that underwater datasets are always much more difficult, expensive, and time-consuming to build. Other factors, such as safety, security, and environmental sustainability, limit the amount of data being collected. Multimodal sensing approaches, on the other hand, suffer from different data acquisitions and different data sources. Combining data from different types of underwater sensing modalities is still a challenge. Optical camera and acoustic sonar data have different operating ranges, resolutions, and Spatio-temporal alignments. The problem requires robust calibration algorithms to be applied before performing any data fusion processes. The requirement is a challenging problem for acoustic sonars as their accuracy is environmental-dependant.

Underwater perception involves many domains, such as computer vision, machine learning, sensor fusion, SLAM, 3D reconstruction, cooperative control, and communication. There are still many problems to be solved. The future development trend of underwater perception can be summarized as follows.

1. Improve current weaknesses of primary sensing devices. The current trend for acoustic sensing devices focuses on improving their captured resolution. Some FLSs enhance their data resolution by using higher operating frequencies to generate images with millimeter resolution with an operating range of less than 10 m. Another trend is to improve sonar resolution by combining data from multiple frequencies to provide high-resolution data. Multi-beam sonar with the capability to produce very high-resolution 3D data over a long distance is another possibility for many underwater projects. While these devices show promising abilities for underwater perception, their expensive cost is another problem to be addressed. On the other hand, optical-based devices will persist in improving their capability to operate in low-light conditions and minimize their energy consumption by using low-power imaging sensors and harvesting energy underwater. The artificial illumination system is another improved aspect to boost image quality. Besides, there is some interest in using the hyperspectral camera for underwater applications. While the device also suffers from seawater attenuation, it can capture underwater objects using hundreds of spectral bands, allowing potentially better robot perception.

2. Underwater sensing algorithms will focus on two main directions: learning-based perception and fusion-based perception. The quick progress of artificial intelligence and machine learning will boost the performance of underwater perception systems. More underwater datasets will be available thanks to the increasing availability of underwater sensors. The more data collected on the aquatic environment, the better we can train the robot to understand underwater environments. Intersensory learning is a promising trend for machine learning in underwater perception research. The methods encourage knowledge transfer between different types of sensing modalities. For example, by presenting low-resolution images from a FLS as input to a neural network and high-resolution photos from a camera as the output, we can learn a model that can render camera-like images using sonar data. However, the approach still requires a decent dataset to be collected for the model to adapt to various underwater scenarios and an algorithm for correctly mapping objects between the camera and FLS. Many successful machine learning models are being adopted for underwater technology, such as Attention Mechanisms or Transformers, to solve classical object perception problems. Fusion-based perception is also a primary trend in underwater robotics. More data fusion methodology will be implemented with new sensing modalities. For example, some research projects are experimenting with pairing the optical camera with underwater Lidar for high-resolution mapping and pairing multiple FLSs with different configurations for better object perception. The biggest challenge for data fusion is correctly calibrating various information sources. For example, the accuracy of sonar calibration is affected by water temperature and salinity.

3. 3D underwater reconstruction will be the focused research direction for underwater perception. 3D reconstruction of objects and environments is always highly interesting for underwater research. The technique allows underwater vehicles to build a 3D model of underwater objects or full coverage of underwater structures. The information enables underwater robots to perform path planning and obstacle avoidance or to navigate near the facilities effectively. For example, there is a recent work on 3D reconstruction using orthogonal stereo FLSs (OMSF). Its reconstruction result can also be integrated as PCD inputs into the PCD-based classifier and pose-estimator models, as shown in the diagram in Fig. 16. Other work are exploring similar research directions, such as those found in Sung et al. (2020). Preliminary development on OMSF integration for 3D classification are showing promising results with great accuracy and efficiency, even on sparse PCDs given by OMSF.

4. Cooperative perception is the new emerging technology for underwater research. The research direction receives tremendous interest in developing a systematic approach for controlling a group of robots. Such a system allows underwater perception tasks such as seafloor surveying, mapping, or 3D site reconstruction to perform effectively, collectively, and quickly. Cooperative perception allows a new allocation strategy for underwater sensing devices. All the sensors used for fusion need not be on just one underwater vehicle like the works of Liu et al. (2020) or Li et al. (2022). A step forward could be towards cooperative localization and imaging, whereby several robots can quickly scan a given target while sharing vital navigation information using acoustic modems (Djapic et al., 2013; Du et al., 2022), or optical modems like the Hydromea LUMA. The main bottleneck of such systems would be the communication protocol, which depends on factors like distance and bandwidth. However, large amounts of data can be transferred easily using such devices due to their speed. Another method of using these multi-robot systems could be in a master–slave fashion, whereby robots with more extended range sensors could help localize robots closer to targets of interest to produce imaging data totally independent of the characteristics of the scanned objects, leading to more robust implementations of SLAM or approaches like one-shot localization. Fig. 17 shows an example of the setup.

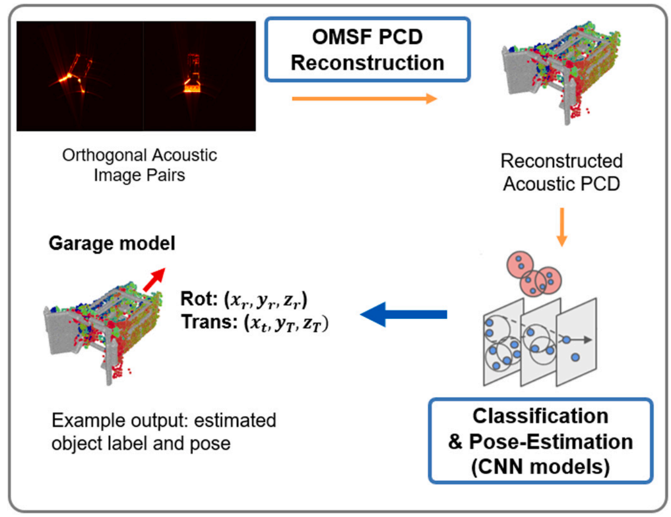


Fig. 16. Rough diagram showing example of a potential end-to-end PCD-based classification and pose-estimation integrating OMSF reconstruction.

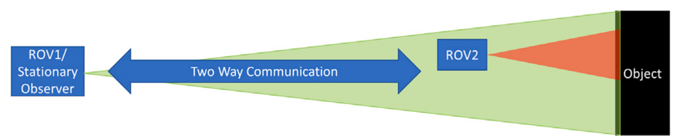


Fig. 17. A possible setup for collaborative robot control.

In this paper, we contribute a survey on marine robotic sensing modalities and perception algorithms. Our underwater perception framework classifies the perception ranges into long-range, medium-range, and close-range. Each perception range is characterized by the availability of the sensor and the sensing methods. Long-range object perception suffers from the low-contrast, low resolution, and noisy background of sonar images. Current sensing algorithms overcome this perception problem by using the object's unique features, such as motion and frequency, instead of relying on only traditional geometrical features, such as size and shape. Fusion-based approaches are appropriate for enhancing the robot's perception capabilities in the medium range. Finally, robots can perform advanced perception tasks such as object reconstructions or pose estimation with high-resolution data from optical and acoustic modalities in close range. With the recent fast-moving progress of related fields such as machine learning and underwater communication, a promising future awaits the development of the marine robotic area.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

This paper is done as part of the work conducted under the SaabNTU Joint Lab with support from Saab Singapore Pte. Ltd, Saab AB and the NTU Robotics Research Centre (RRC). The authors are grateful to Dr. Jovice Ng of Saab Singapore for her valuable advice and comments, to Burhan and Chia Hao Gen of RRC for the technical support and development of pool experiments, and to Dr. Andreas Gällström of Saab AB for his assistance in proofreading this article (SAAB - NTU Project Agreement - NTU Ref: 2018-0356).

References

Abdulsadda, A.T., Tan, X., 2013. Underwater tracking of a moving dipole source using an artificial lateral line: algorithm and experimental validation with ionic polymer–metal composite flow sensors. Smart Mater. Struct. 22 (4), 045010. http://dx.doi.org/10.1088/0964-1726/22/4/045010.

Abu, A., 2020. CFAR detection algorithm for objects in sonar images. IET Radar Sonar Navig. 14, 1757–1766. http://dx.doi.org/10.1049/iet-rsn.2020.0230.

Acosta, G.G., Villar, S.A., 2015. Accumulated CA–CFAR process in 2-D for online object detection from sidescan sonar data. IEEE J. Ocean. Eng. 40 (3), 558–569. http://dx.doi.org/10.1109/JOE.2014.2356951.

Akkaynak, D., Treibitz, T., 2018. A revised underwater image formation model. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6723–6732. http://dx.doi.org/10.1109/CVPR.2018.00703.

Akkaynak, D., Treibitz, T., 2019. Sea-thru: A method for removing water from underwater images. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 1682–1691. http://dx.doi.org/10.1109/CVPR.2019.00178.

Andreas Gällström, C.L., 2019. Enhanced sonar image resolution using compressive sending modelling. In: 5th Underwater Acoustics Conference and Exhibition. (5), pp. 995–999.

Anwer, A., Ali, S.S.A., Khan, A., Meriaudeau, F., 2017. Underwater 3-D scene reconstruction using kinect v2 based on physical models for refraction and time of flight correction. IEEE Access 5, http://dx.doi.org/10.1109/ACCESS.2017.2733003.

Aykin, M.D., Negahdaripour, S., 2017. Three-dimensional target reconstruction from multiple 2-D forward-scan sonar views by space carving. IEEE J. Ocean. Eng. 42 (3), 574–589. http://dx.doi.org/10.1109/JOE.2016.2591738.

Bai, Y., Snyder, J.B., Peshkin, M., MacIver, M.A., 2015. Finding and identifying simple objects underwater with active electrosense. Int. J. Robot. Res. 34 (10), 1255–1277. http://dx.doi.org/10.1177/0278364915569813.

Bai, Y., Snyder, J., Silverman, Y., Peshkin, M., MacIver, M.A., 2012. Sensing capacitance of underwater objects in bio-inspired electrosense. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 1467–1472. http://dx.doi.org/10.1109/IROS.2012.6386174.

Bazeille, S., Lebastard, V., Boyer, F., 2020. A purely model-based approach to object pose and size estimation with electric sense. IEEE Trans. Robot. 36 (5), 1611–1618. http://dx.doi.org/10.1109/TRO.2020.3000285.

Bongiorno, D.L., Bryson, M., Bridge, T.C., Dansereau, D.G., Williams, S.B., 2018. Coregistered hyperspectral and stereo image seafloor mapping from an autonomous underwater vehicle. J. Field Robotics 35, 312–329. http://dx.doi.org/10.1002/ROB.21713, https://onlinelibrary.wiley.com/doi/full/10.1002/rob.21713, https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21713, https://onlinelibrary.wiley.com/doi/10.1002/rob.21713.

Boyer, F., Lebastard, V., Chevallereau, C., Mintchev, S., Stefanini, C., 2015. Underwater navigation based on passive electric sense: New perspectives for underwater docking. Int. J. Robot. Res. 34 (9), 1228–1250. http://dx.doi.org/10.1177/0278364915572071.

Bruno, F., Lagudi, A., Ritacco, G., Muzzupappa, M., Guida, R., 2015. Opto-acoustic data fusion for supporting the guidance of remotely operated underwater vehicles (rovs). Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XL-5-W5, 47–53. http://dx.doi.org/10.5194/ISPRSARCHIVES-XL-5-W5-47-2015.

Chantler, M., Stoner, J., 1997. Automatic interpretation of sonar image sequences using temporal feature measures. IEEE J. Ocean. Eng. 22 (1), 47–56. http://dx.doi.org/10.1109/48.557539.

Charalambides, J., 2016. Getting to the bottom of subsea repairs. Control Eng. URL https://www.controleng.com/articles/getting-to-the-bottom-of-subsea-repairs/.

Chemisky, B., Nocerino, E., Menna, F., Nawaf, M.M., Drap, P., 2021. A portable opto-acoustic survey solution for mapping of underwater targets. 43, http://dx.doi.org/10.5194/isprs-archives-XLIII-B2-2021-651-2021,

Chen, H., Agarwal, S., Yang, J., Guan, J., Tang, X., Li, A., Jodin, G., Fan, D., 2021. Fish-lateral-inspired pressure sensing neural networks for underwater object identification. In: OCEANS 2021: San Diego – Porto. pp. 1–6. http://dx.doi.org/10.23919/OCEANS44145.2021.9706095.

Chen, Z., Zhang, Z., Dai, F., Bu, Y., Wang, H., 2017. Monocular vision-based underwater object detection. Sensors 17 (8), http://dx.doi.org/10.3390/s17081784, URL https://www.mdpi.com/1424-8220/17/8/1784.

Cheng, C., Sha, Q., He, B., Li, G., 2021. Path planning and obstacle avoidance for AUV: A review. Ocean Eng. 235, 109355. http://dx.doi.org/10.1016/j.oceaneng.2021.109355, URL https://www.sciencedirect.com/science/article/pii/S002980182100771X.

Cheng, C., Wang, C., Yang, D., Liu, W., Zhang, F., 2022. Underwater localization and mapping based on multi-beam forward looking sonar. Front. Neurorobot. 15, 189. http://dx.doi.org/10.3389/FNBOT.2021.801956/BIBTEX.

Christensen, L., de Gea Fernández, J., Hildebrandt, M., Koch, C.E.S., Wehbe, B., 2022. Recent advances in AI for navigation and control of underwater robots. Curr. Robot. Rep. http://dx.doi.org/10.1007/s43154-022-00088-3.

Christensen, J.H., Mogensen, L.V., Ravn, O., 2021. Side-scan sonar imaging: Automatic boulder identification. In: OCEANS 2021: San Diego – Porto. pp. 1–6. http://dx.doi.org/10.23919/OCEANS44145.2021.9705713.

Clem, T., Allen, G., Bono, J., McDonald, R., Overway, D., Sulzberger, G., Kumar, S., King, D., 2004. Magnetic sensors for buried minehunting from small unmanned underwater vehicles. In: Oceans '04 MTS/IEEE Techno-Ocean '04 (IEEE Cat. No.04CH37600). Vol. 2, pp. 902–910. http://dx.doi.org/10.1109/OCEANS.2004.1405594.

Collings, S., Martin, T.J., Hernandez, E., Edwards, S., Filisetti, A., Catt, G., Marouchos, A., Boyd, M., Embry, C., 2020. Findings from a combined subsea LiDAR and multibeam survey at kingston reef, western Australia. Remote Sens. 12 (15), http://dx.doi.org/10.3390/rs12152443, URL https://www.mdpi.com/2072-4292/12/15/2443.

Cui, J., Hu, C., Xu, H., 2018. Detection of the moving targets of forward-looking sonar based on background elimination method. In: 2018 IEEE International Conference on Signal Processing, Communications and Computing. ICSPCC, pp. 1–4. http://dx.doi.org/10.1109/ICSPCC.2018.8567815.

Dai, D., Chantler, M., Lane, D., Williams, N., 1995. A spatial-temporal approach for segmentation of moving and static objects in sector scan sonar image sequences. In: Fifth International Conference on Image Processing and Its Applications, 1995. pp. 163–167. http://dx.doi.org/10.1049/cp:19950641.

Deans, C., Marmugi, L., Renzoni, F., 2018. Active underwater detection with an array of atomic magnetometers. Appl. Opt. 57 (10), 2346–2351. http://dx.doi.org/10.1364/AO.57.002346, URL https://opg.optica.org/ao/abstract.cfm?URI=ao-57-10-2346.

Demirel, H., Anbarjafari, G., 2011. IMAGE resolution enhancement by using discrete and stationary wavelet decomposition. IEEE Trans. Image Process. 20 (5), 1458–1460. http://dx.doi.org/10.1109/TIP.2010.2087767.

Denos, K., Ravaut, M., Fagette, A., Lim, H.-S., 2017. Deep learning applied to underwater mine warfare. In: OCEANS 2017 - Aberdeen. pp. 1–7. http://dx.doi.org/10.1109/OCEANSE.2017.8084910.

DeVries, L., Lagor, F.D., Lei, H., Tan, X., Paley, D.A., 2015. Distributed flow estimation and closed-loop control of an underwater vehicle with a multi-modal artificial lateral line. Bioinspiration Biomim. 10 (2), 025002. http://dx.doi.org/10.1088/1748-3190/10/2/025002.

Djapic, V., Nad, D., Ferri, G., Omerdic, E., Dooly, G., Toal, D., Vukic, Z., 2013. Novel method for underwater navigation aiding using a companion underwater robot as a guiding platforms. In: OCEANS 2013 MTS/IEEE Bergen: The Challenges of the Northern Dimension. http://dx.doi.org/10.1109/OCEANS-BERGEN.2013.6608153.

Dos Santos, M., Ribeiro, P.O., Núñez, P., Drews-Jr, P., Botelho, S., 2017. Object classification in semi structured enviroment using forward-looking sonar. Sensors 17 (10), http://dx.doi.org/10.3390/s17102235, URL https://www.mdpi.com/1424-8220/17/10/2235.

Doucet, A., Godsill, S., Andrieu, C., 2003. On sequential Monte Carlo sampling methods for Bayesian filtering. Stat. Comput. 10, http://dx.doi.org/10.1023/A:1008935410038.

Du, Z., Chai, H., Li, Z., Xiang, M.Z., Zhang, F., Hui, J., 2022. Acoustic ray-trace correction for UUVs cooperative localization in deep ocean applications. Mar. Geod. http://dx.doi.org/10.1080/01490419.2022.2059601.

Echoscope, 2021. Echoscope 4G. URL https://www.codaoctopus.com/products/3d/echoscope4G.

El-Darymli, K., McGuire, P., Power, D., Moloney, C.R., 2013. Target detection in synthetic aperture radar imagery: a state-of-the-art survey. J. Appl. Remote Sens. 7 (1), 1–35. http://dx.doi.org/10.1117/1.JRS.7.071598.

Fabbri, C., Islam, M.J., Sattar, J., 2018. Enhancing underwater imagery using generative adversarial networks. In: 2018 IEEE International Conference on Robotics and Automation. ICRA, pp. 7159–7165. http://dx.doi.org/10.1109/ICRA.2018.8460552.

Fuchs, L.R., Gällström, A., Folkesson, J., 2018a. Object recognition in forward looking sonar images using transfer learning. In: 2018 IEEE/OES Autonomous Underwater Vehicle Workshop. AUV, pp. 1–6. http://dx.doi.org/10.1109/AUV.2018.8729686.

Fuchs, L.R., Gällström, A., Folkesson, J., 2018b. Object recognition in forward looking sonar images using transfer learning. In: 2018 IEEE/OES Autonomous Underwater Vehicle Workshop. AUV, pp. 1–6. http://dx.doi.org/10.1109/AUV.2018.8729686.

Galceran, E., Djapic, V., Carreras, M., Williams, D.P., 2012. A real-time underwater object detection algorithm for multi-beam forward looking sonar. IFAC Proc. Vol. 45 (5), 306–311. http://dx.doi.org/10.3182/20120410-3-PT-4028.00051, URL https://www.sciencedirect.com/science/article/pii/S1474667016306206, 3rd IFAC Workshop on Navigation, Guidance and Control of Underwater Vehicles.

Gandhi, P., Kassam, S., 1988. Analysis of CFAR processors in nonhomogeneous background. IEEE Trans. Aerosp. Electron. Syst. 24 (4), 427–445. http://dx.doi.org/10.1109/7.7185.

Gao, G., Liu, L., Zhao, L., Shi, G., Kuang, G., 2009. An adaptive and fast CFAR algorithm based on automatic censoring for target detection in high-resolution SAR images. IEEE Trans. Geosci. Remote Sens. 47, 1685–1697. http://dx.doi.org/10.1109/TGRS.2008.2006504.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems. Vol. 27, Curran Associates, Inc., URL https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf.

Granstrom, K., Lundquist, C., Orguner, O., 2012. Extended target tracking using a Gaussian-mixture PHD filter. IEEE Trans. Aerosp. Electron. Syst. 48 (4), 3268–3286. http://dx.doi.org/10.1109/TAES.2012.6324703.

Gu, C., Cong, Y., Sun, G., 2019. Environment driven underwater camera-IMU calibration for monocular visual-inertial SLAM. In: Proceedings - IEEE International Conference on Robotics and Automation. Institute of Electrical and Electronics Engineers Inc., pp. 2405–2411. http://dx.doi.org/10.1109/ICRA.2019.8793577.

Guo, Y., Li, H., Zhuang, P., 2020. Underwater image enhancement using a multiscale dense generative adversarial network. IEEE J. Ocean. Eng. 45 (3), 862–870. http://dx.doi.org/10.1109/JOE.2019.2911447.

Guo, Y., Liu, H., Fan, X., Lyu, W., 2021. Research progress of path planning methods for autonomous underwater vehicle. Math. Probl. Eng. 2021, 8847863. http://dx.doi.org/10.1155/2021/8847863.

Hambarde, P., Murala, S., Dhall, A., 2021. UW-GAN: Single-image depth estimation and image enhancement for underwater images. IEEE Trans. Instrum. Meas. 70, 1–12. http://dx.doi.org/10.1109/TIM.2021.3120130.

Hu, S., Tang, J., Ren, Z., Chen, C., Zhou, C., Xiao, X., Zhao, T., 2019. Multiple underwater objects localization with magnetic gradiometry. IEEE Geosci. Remote Sens. Lett. 16 (2), 296–300. http://dx.doi.org/10.1109/LGRS.2018.2870839.

Hua, J., Liu, M., Wang, S., 2021. A Super-Resolution Reconstruction Method of Underwater Target Detection Image by Side Scan Sonar. Association for Computing Machinery, New York, NY, USA, pp. 135–140. http://dx.doi.org/10.1145/3483845.3483869.

Instruments, S., 2000. Multibeam sonar theory of operation.

Jakuba, M.V., Pizarro, O., Williams, S.B., 2010. High resolution, consistent navigation and 3D optical reconstructions from AUVs using magnetic compasses and pressure-based depth sensors. In: OCEANS'10 IEEE Sydney, OCEANSSYD 2010. http://dx.doi.org/10.1109/OCEANSSYD.2010.5603895.

Jalal, F., Nasir, F., 2021. Underwater navigation, localization and path planning for autonomous vehicles: A review. In: 2021 International Bhurban Conference on Applied Sciences and Technologies. IBCAST, pp. 817–828. http://dx.doi.org/10.1109/IBCAST51254.2021.9393315.

Ji, Y., Xie, L., Shi, Y., 2021. Adaptive underwater sonar image target detection. In: 2021 40th Chinese Control Conference. CCC, pp. 7355–7360. http://dx.doi.org/10.23919/CCC52363.2021.9549369.

Jie, F., Pingbo, W., 2020. Target detection in sonar image based on faster RCNN. In: 2020 International Conference on Information Science, Parallel and Distributed Systems. ISPDS, pp. 25–30. http://dx.doi.org/10.1109/ISPDS51347.2020.00013.

Joe, H., Cho, H., Sung, M., Kim, J., cheol Yu, S., 2021. Sensor fusion of two sonar devices for underwater 3D mapping with an AUV. Auton. Robots 45, http://dx.doi.org/10.1007/s10514-021-09986-5.

Joe, H., Cho, H., Yu, S.C., 2019a. Implementation of point cloud algorithm on three different installations of mechanically scanning imaging sonar for seabed mapping. In: IEEE Region 10 Annual International Conference, Proceedings/TENCON. 2018-October, http://dx.doi.org/10.1109/TENCON.2018.8650281.

Joe, H., Kim, J., Yu, S.C., 2019b. Sensor fusion-based 3D reconstruction by two sonar devices for seabed mapping. 52, http://dx.doi.org/10.1016/j.ifacol.2019.12.302,

Joe, H., Kim, J., Yu, S.C., 2020. 3D reconstruction using two sonar devices in a Monte-Carlo approach for AUV application. Int. J. Control Autom. Syst. 18, http://dx.doi.org/10.1007/s12555-019-0692-2.

Joe, H., Yu, S.C., 2018. Submap based normal distribution transform scan matching by using heterogeneous sonars for AUV navigation. In: AUV 2018 - 2018 IEEE/OES Autonomous Underwater Vehicle Workshop, Proceedings. http://dx.doi.org/10.1109/AUV.2018.8729734.

John McConnell, J.D.M., Englot, B., 2020. Fusing concurrent orthogonal wide-aperture sonar images for dense underwater 3D reconstruction. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS, IEEE.

Karoui, I., Quidu, I., Legris, M., 2015. Automatic sea-surface obstacle detection and tracking in forward-looking sonar image sequences. IEEE Trans. Geosci. Remote Sens. 53, 1–10. http://dx.doi.org/10.1109/TGRS.2015.2405672.

Keenan, S.T., Young, J.A., Foley, C.P., Du, J., 2010. A high-it/isubc/subflip-chip SQUID gradiometer for mobile underwater magnetic sensing. Supercond. Sci. Technol. 23 (2), 025029. http://dx.doi.org/10.1088/0953-2048/23/2/025029.

Kim, B., Kim, J., Cho, H., Kim, J., Yu, S.-C., 2020. AUV-based multi-view scanning method for 3-D reconstruction of underwater object using forward scan sonar. IEEE Sens. J. 20 (3), 1592–1606. http://dx.doi.org/10.1109/JSEN.2019.2946587.

Kim, J., Lee, M., Song, S., Kim, B., Yu, S.-C., 2019. 3-d reconstruction of underwater objects using image sequences from optical camera and imaging sonar. In: OCEANS 2019 MTS/IEEE SEATTLE. pp. 1–6. http://dx.doi.org/10.23919/OCEANS40490.2019.8962558.

Kim, J., Yu, S.-C., 2016. Convolutional neural network-based real-time ROV detection using forward-looking sonar image. In: 2016 IEEE/OES Autonomous Underwater Vehicles. AUV, pp. 396–400. http://dx.doi.org/10.1109/AUV.2016.7778702.

Kominis, I., Kornack, T., Allred, J., Romalis, M., 2003. A subfemtotesla multichannel atomic magnetometer. Nature 422, 596–599. http://dx.doi.org/10.1038/nature01484.

Kronauge, M., Rohling, H., 2013. Fast two-dimensional CFAR procedure. IEEE Trans. Aerosp. Electron. Syst. 49 (3), 1817–1823. http://dx.doi.org/10.1109/TAES.2013.6558022.

Kumar, M., Mondal, S., 2021. Recent developments on target tracking problems: A review. Ocean Eng. 236, 109558. http://dx.doi.org/10.1016/j.oceaneng.2021.109558, URL https://www.sciencedirect.com/science/article/pii/S0029801821009471.

Kumar, S., Perry, A., Moeller, C., Skvoretz, D., Ebbert, M., Ostrom, R., Bennett, S., Czipott, P., 2004. Real-time tracking magnetic gradiometer for underwater mine detection. In: Oceans '04 MTS/IEEE Techno-Ocean '04 (IEEE Cat. No.04CH37600). Vol. 2, pp. 874–878. http://dx.doi.org/10.1109/OCEANS.2004.1405583.

Kumudham R., R.V., 2019. Super resolution enhancement of underwater sonar images. SN Appl. Sci. (852), http://dx.doi.org/10.1007/s42452-019-0886-5.

Lagudi, A., Bianco, G., Muzzupappa, M., Bruno, F., 2016. An alignment method for the integration of underwater 3D data captured by a stereovision system and an acoustic camera. Sensors (Basel, Switzerland) 16, 536. http://dx.doi.org/10.3390/s16040536, https://pubmed.ncbi.nlm.nih.gov/27089344, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4851050/.

Lama, R.K., Shin, S., Kang, M., Kwon, G.-R., Choi, M.-R., 2016. Interpolation using wavelet transform and discrete cosine transform for high resolution display. In: 2016 IEEE International Conference on Consumer Electronics. ICCE, pp. 184–186. http://dx.doi.org/10.1109/ICCE.2016.7430573.

Lane, D., Chantler, M., Dai, D., 1998a. Robust tracking of multiple objects in sector-scan sonar image sequences using optical flow motion estimation. IEEE J. Ocean. Eng. 23 (1), 31–46. http://dx.doi.org/10.1109/48.659448.

Lane, D., Chantler, M., Dai, D.Y., Tena Ruiz, I., 1998b. Tracking and classification of multiple objects in multibeam sector scan sonar image sequences. In: Proceedings of 1998 International Symposium on Underwater Technology. pp. 269–273. http://dx.doi.org/10.1109/UT.1998.670107.

Lane, D., Stoner, J., 1994. Automatic interpretation of sonar imagery using qualitative feature matching. IEEE J. Ocean. Eng. 19 (3), 391–405. http://dx.doi.org/10.1109/48.312915.

Lebastard, V., Boyer, F., Lanneau, S., 2016. Reactive underwater object inspection based on artificial electric sense. Bioinspiration Biomim. 11 (4), 045003. http://dx.doi.org/10.1088/1748-3190/11/4/045003.

Lee, S., Park, B., Kim, A., 2019. Deep learning based object detection via style-transferred underwater sonar images. IFAC-PapersOnLine 52 (21), 152–155. http://dx.doi.org/10.1016/j.ifacol.2019.12.299, URL https://www.sciencedirect.com/science/article/pii/S2405896319321846, 12th IFAC Conference on Control Applications in Marine Systems, Robotics, and Vehicles CAMS 2019.

Li, H., Gao, J., Du, W., Zhou, T., xu, C., Baowei, C., 2017. Object representation for multi-beam sonar image using local higher-order statistics. EURASIP J. Adv. Signal Process. 2017, http://dx.doi.org/10.1186/s13634-016-0439-7.

Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., Tao, D., 2020. An underwater image enhancement benchmark dataset and beyond. IEEE Trans. Image Process. 29, 4376–4389. http://dx.doi.org/10.1109/TIP.2019.2955241.

Li, M., Mathai, A., Lau, S.L.H., Yam, J.W., Xu, X., Wang, X., 2021. Underwater object detection and reconstruction based on active single-pixel imaging and super-resolution convolutional neural network. Sensors 21 (1), http://dx.doi.org/10.3390/s21010313, URL https://www.mdpi.com/1424-8220/21/1/313.

Li, J., Skinner, K.A., Eustice, R.M., Johnson-Roberson, M., 2018. Watergan: Unsupervised generative network to enable real-time color correction of monocular underwater images. IEEE Robot. Autom. Lett. 3 (1), 387–394. http://dx.doi.org/10.1109/LRA.2017.2730363.

Li, X., Wang, H., Shen, X., 2022. Multi-UAV active target search based on perceptual adaptation in marine environment. Cogn. Comput. Syst. 4, 218–228. http://dx.doi.org/10.1049/CCS2.12050, https://onlinelibrary.wiley.com/doi/full/10.1049/ccs2.12050, https://onlinelibrary.wiley.com/doi/abs/10.1049/ccs2.12050, https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ccs2.12050.

Li, Y., Zhang, M., 2020. Simulation study for obstacle avoidance of autonomous underwater vehicles. J. Coast. Res. 104–108, URL https://www.jstor.org/stable/48639958.

Liu, Z.-S., Siu, W.-C., Huang, J.-J., 2015. Image super-resolution via hybrid NEDI and wavelet-based scheme. In: 2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. APSIPA, pp. 1131–1136. http://dx.doi.org/10.1109/APSIPA.2015.7415447.

Liu, Y., Wang, M., Su, Z., Luo, J., Xie, S., Peng, Y., Pu, H., Xie, J., Zhou, R., 2020. Multi-AUVs cooperative target search based on autonomous cooperative search learning algorithm. J. Mar. Sci. Eng. 8, 843. http://dx.doi.org/10.3390/JMSE8110843, https://www.mdpi.com/2077-1312/8/11/843/htm, https://www.mdpi.com/2077-1312/8/11/843.

Ma, L.Y., He, X.L., Huang, Z.S., 2011. Sidescan sonar image super resolution based on sparse representation. In: Measuring Technology and Mechatronics Automation. In: Applied Mechanics and Materials, Vol. 48, Trans Tech Publications Ltd, pp. 1049–1052. http://dx.doi.org/10.4028/www.scientific.net/AMM.48-49.1049.

Mallios, A., Ridao, P., Romagós, D., Hernández, E., 2014. Scan matching SLAM in underwater environments. Auton. Robots 36, http://dx.doi.org/10.1007/s10514-013-9345-0.

Maussang, F., Chanussot, J., Hetet, A., Amate, M., 2007. Mean–standard deviation representation of sonar images for echo detection: Application to SAS images. IEEE J. Ocean. Eng. 32 (4), 956–970. http://dx.doi.org/10.1109/JOE.2007.907936.

McConnell, J., Englot, B., 2021. Predictive 3D sonar mapping of underwater environments via object-specific Bayesian inference. In: IEEE/RSJ International Conference on Robotics and Automation. ICRA, IEEE.

Melzi, M., Ouldali, A., 2011. Joint multiple target tracking and classification using the unscented Kalman particle PHD filter. In: 2011 IEEE 9th International New Circuits and Systems Conference. pp. 534–537. http://dx.doi.org/10.1109/NEWCAS.2011.5981202.

Modalavalasa, N., 2012. A novel approach for segmentation of sector scan sonar images using adaptive thresholding. Int. J. Inf. Sci. Tech. 2, 113–119. http://dx.doi.org/10.5121/ijist.2012.2411.

Nakamura, S., Yamada, D., Mukada, N., Myint, M., Lwin, K.N., Matsuno, T., Toda, Y., Minami, M., 2018. Development of dual-eyes docking system for AUV with lighting 3D marker. In: OCEANS 2018 MTS/IEEE Charleston. pp. 1–9. http://dx.doi.org/10.1109/OCEANS.2018.8604527.

Negahdaripour, S., 2018. Analyzing epipolar geometry of 2-D forward-scan sonar stereo for matching and 3-D reconstruction. In: OCEANS 2018 MTS/IEEE Charleston. pp. 1–10. http://dx.doi.org/10.1109/OCEANS.2018.8604540.

Negahdaripour, S., 2020. Application of forward-scan sonar stereo for 3-D scene reconstruction. IEEE J. Ocean. Eng. 45 (2), 547–562. http://dx.doi.org/10.1109/JOE.2018.2875574.

Park, J., Kim, J., 2016. High-precision underwater navigation using model-referenced pose estimation with monocular vision. In: 2016 IEEE/OES Autonomous Underwater Vehicles. AUV, pp. 138–143. http://dx.doi.org/10.1109/AUV.2016.7778661.

Park, J., Ku, B., Jin, Y., Ko, H., 2019. Side scan sonar image super resolution via region-selective sparse coding. IEICE Trans. Inf. Syst. 102-D, 210–213.

Pei, Y.H., Yeo, H.G., Kang, X.Y., Pua, S.L., Tan, J., 2010. Magnetic gradiometer on an AUV for buried object detection. In: OCEANS 2010 MTS/IEEE SEATTLE. pp. 1–8. http://dx.doi.org/10.1109/OCEANS.2010.5664272.

Perry, S., Guan, L., 2004. A recurrent neural network for detecting objects in sequences of sector-scan sonar images. IEEE J. Ocean. Eng. 29 (3), 857–871. http://dx.doi.org/10.1109/JOE.2004.831616.

Petillot, Y., Tena Ruiz, I., Lane, D., 2001. Underwater vehicle obstacle avoidance and path planning using a multi-beam forward looking sonar. IEEE J. Ocean. Eng. 26 (2), 240–251. http://dx.doi.org/10.1109/48.922790.

Phung, S.L., Nguyen, T.N.A., Le, H.T., Chapple, P.B., Ritz, C.H., Bouzerdoum, A., Tran, L.C., 2019. Mine-like object sensing in sonar imagery with a compact deep learning architecture for scarce data. In: 2019 Digital Image Computing: Techniques and Applications. DICTA, pp. 1–7. http://dx.doi.org/10.1109/DICTA47822.2019.8945982.

Piao, Y., Shin, l.-h., Park, H., 2007. Image resolution enhancement using inter-subband correlation in wavelet domain. In: 2007 IEEE International Conference on Image Processing. Vol. 1, pp. I–445–I–448. http://dx.doi.org/10.1109/ICIP.2007.4378987.

Pike, J., 1999. Cable-controlled undersea recovery vehicle (curv). URL https://irp.fas.org/program/collect/curv.htm.

Priyadharsini, R., Sharmila, T.S., Rajendran, V., 2015. Underwater acoustic image enhancement using wavelet and K-L transform. In: 2015 International Conference on Applied and Theoretical Computing and Communication Technology (ICATccT). pp. 563–567. http://dx.doi.org/10.1109/ICATCCT.2015.7456948.

Pu, Y., Hang, Z., Wang, G., Hu, H., 2022. Bionic artificial lateral line underwater localization based on the neural network method. Appl. Sci. 12 (14), http://dx.doi.org/10.3390/app12147241, URL https://www.mdpi.com/2076-3417/12/14/7241.

Quidu, I., Bertholom, A., Dupas, Y., 2010. Ground obstacle tracking on forward looking sonar images. In: Proceedings of ECUA'10.

Quidu, I., Jaulin, L., Bertholom, A., Dupas, Y., 2012. Robust multitarget tracking in forward-looking sonar image sequences using navigational data. IEEE J. Ocean. Eng. 37 (3), 417–430. http://dx.doi.org/10.1109/JOE.2012.2190572.

Raaj, Y., John, A., Jin, T., 2016a. 3D object localization using forward looking sonar (FLS) and optical camera via particle filter based calibration and fusion. In: OCEANS 2016 MTS/IEEE Monterey, OCE 2016. http://dx.doi.org/10.1109/OCEANS.2016.7761077.

Raaj, Y., John, A., Jin, T., 2016b. 3D object localization using forward looking sonar (FLS) and optical camera via particle filter based calibration and fusion. In: OCEANS 2016 MTS/IEEE Monterey. pp. 1–10. http://dx.doi.org/10.1109/OCEANS.2016.7761077.

Rahman, S., Li, A.Q., Rekleitis, I., 2018. SVIn2: An underwater SLAM system using sonar, visual, inertial, and depth sensor. http://dx.doi.org/10.1109/IROS40897.2019.8967703, http://arxiv.org/abs/1810.03200.

Rajamohan, K., Rajendran, V., 2019. Super resolution enhancement of underwater sonar images. SN Appl. Sci. 1, http://dx.doi.org/10.1007/s42452-019-0886-5.

Rajamohan, K., Tarun, S., Avinash, R., Rajendran, V., 2018. Resolution enrichment of side scan sonar image using wavelet based interpolation methods. Int. J. Eng. Technol. 7, 375. http://dx.doi.org/10.14419/ijet.v7i2.21.12445.

Ravisankar, P., Sharmila, T., Rajendran, V., 2018. A wavelet transform based contrast enhancement method for underwater acoustic images. Multidimens. Syst. Signal Process. 29, http://dx.doi.org/10.1007/s11045-017-0533-5.

Ren, R., Zhang, L., Liu, L., Yuan, Y., 2021. Two AUVs guidance method for self-reconfiguration mission based on monocular vision. IEEE Sens. J. 21 (8), 10082–10090. http://dx.doi.org/10.1109/JSEN.2020.3042306.

Rixon Fuchs, L., Larsson, C., Gällström, A., 2019. Deep learning based technique for enhanced sonar imaging.

Roznere, M., Li, A.Q., 2020. Underwater monocular image depth estimation using single-beam echosounder. In: IEEE International Conference on Intelligent Robots and Systems. Institute of Electrical and Electronics Engineers Inc., pp. 1785–1790. http://dx.doi.org/10.1109/IROS45743.2020.9340919.

Ruiz, I., Petillot, Y., Lane, D., Bell, J., 1999. Tracking objects in underwater multibeam sonar images. IEE Colloquium (Digest) (103), 69–75, IEE Colloquium on Moation Aanalysis and Tracking ; Conference date: 01-05-1999.

Sadjoli, N., Gen, C.H., Huy, D.Q., Yiyu, C., Lee, G.S.G., Elhadidi, B., 2021. Acoustic PCD-based AUV perception for enhanced automated garage docking. In: OCEANS 2021: San Diego - Porto. pp. 1–9. http://dx.doi.org/10.23919/OCEANS44145.2021.9705817.

Sharmila, T., Kadarkarai, R., Thangaswamy, S.R.R., 2013. Developing an efficient technique for satellite image denoising and resolution enhancement for improving classification accuracy. J. Electron. Imaging 22, 3013. http://dx.doi.org/10.1117/1.JEI.22.1.013013.

Shen, J., Xu, Z., Chen, Z., Wang, H., Shi, X., 2021a. Optical prior-based underwater object detection with active imaging. In: Complexity. (6656166), http://dx.doi.org/10.1155/2021/6656166.

Shen, Y., Zhao, C., Liu, Y., Wang, S., Huang, F., 2021b. Underwater optical imaging: Key technologies and applications review. IEEE Access 9, 85500–85514. http://dx.doi.org/10.1109/ACCESS.2021.3086820.

Shkurti, F., Chang, W.-D., Henderson, P., Islam, M.J., Higuera, J.C.G., Li, J., Manderson, T., Xu, A., Dudek, G., Sattar, J., 2017. Underwater multi-robot convoying using visual tracking by detection. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS, pp. 4189–4196. http://dx.doi.org/10.1109/IROS.2017.8206280.

Singh, P., Gregson, E., Ross, J., Seto, M., Kaminski, C., Hopkin, D., 2020. Vision-based AUV docking to an underway dock using convolutional neural networks. In: 2020 IEEE/OES Autonomous Underwater Vehicles Symposium. AUV, pp. 1–6. http://dx.doi.org/10.1109/AUV50043.2020.9267926.

Solberg, J.R., Lynch, K.M., MacIver, M.A., 2008. Active electrolocation for underwater target localization. Int. J. Robot. Res. 27 (5), 529–548. http://dx.doi.org/10.1177/0278364908090538.

Sozer, E., Stojanovic, M., Proakis, J., 2000. Underwater acoustic networks. IEEE J. Ocean. Eng. 25 (1), 72–83. http://dx.doi.org/10.1109/48.820738.

Sung, M., Joe, H., Kim, J., Yu, S.-C., 2018a. Convolutional neural network based resolution enhancement of underwater sonar image without losing working range of sonar sensors. In: 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans. OTO, pp. 1–6. http://dx.doi.org/10.1109/OCEANSKOBE.2018.8559421.

Sung, M., Kim, J., Cho, H., Lee, M., Yu, S.-C., 2020. Underwater-sonar-image-based 3D point cloud reconstruction for high data utilization and object classification using a neural network. Electronics 9 (11), http://dx.doi.org/10.3390/electronics9111763, URL https://www.mdpi.com/2079-9292/9/11/1763.

Sung, M., Kim, J., Yu, S.-C., 2018b. Image-based super resolution of underwater sonar images using generative adversarial network. In: TENCON 2018 - 2018 IEEE Region 10 Conference. pp. 0457–0461. http://dx.doi.org/10.1109/TENCON.2018.8650176.

Sung, M., Kim, J., Yu, S.-C., 2018c. Image-based super resolution of underwater sonar images using generative adversarial network. In: TENCON 2018 - 2018 IEEE Region 10 Conference. pp. 0457–0461. http://dx.doi.org/10.1109/TENCON.2018.8650176.

Tamsett, D., McIlvenny, J., Baxter, J., Gois, P., Williamson, B., 2019. On the information advantage of sidescan sonar three-frequency colour over greyscale imagery. J. Mar. Sci. Eng. 7, http://dx.doi.org/10.3390/jmse7080276.

Tamsett, D., McIlvenny, J., Watts, A., 2016. Colour sonar: Multi-frequency sidescan sonar images of the seabed in the inner sound of the pentland firth, Scotland. J. Mar. Sci. Eng. 4, http://dx.doi.org/10.3390/jmse4010026.

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., Liu, C., 2018. A survey on deep transfer learning. In: Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I. (Eds.), Artificial Neural Networks and Machine Learning – ICANN 2018. Springer International Publishing, Cham, pp. 270–279. http://dx.doi.org/10.1007/978-3-030-01424-7_27.

Trucco, E., Pétillot, Y.R., Ruiz, I.T., Plakas, K., Lane, D.M., 2000. Feature tracking in video and sonar subsea sequences with applications. Comput. Vis. Image Underst. 79, 92–122.

Valdenegro-Toro, M., 2016. End-to-end object detection and recognition in forward-looking sonar images with convolutional neural networks. In: 2016 IEEE/OES Autonomous Underwater Vehicles. AUV, pp. 144–150. http://dx.doi.org/10.1109/AUV.2016.7778662.

Vo, B.-N., Ma, W.-K., 2006. The Gaussian mixture probability hypothesis density filter. IEEE Trans. Signal Process. 54 (11), 4091–4104. http://dx.doi.org/10.1109/TSP.2006.881190.

Vo, B.-N., Singh, S., Doucet, A., 2003. Sequential Monte Carlo implementation of the PHD filter for multi-target tracking. In: Proceedings of 6th International Conference Information Fusion. Vol. 2, pp. 792–799. http://dx.doi.org/10.1109/ICIF.2003.177320.

Wang, X., Wang, G., Wu, Y., 2018. An adaptive particle swarm optimization for underwater target tracking in forward looking sonar image sequences. IEEE Access 6, 46833–46843. http://dx.doi.org/10.1109/ACCESS.2018.2866381.

Wang, N., Zhou, Y., Han, F., Zhu, H., Zheng, Y., 2019. UWGAN: Underwater GAN for real-world underwater color restoration and dehazing. arXiv:1912.10269.

Weiss, M., 1982. Analysis of some modified cell-averaging CFAR processors in multiple-target situations. IEEE Trans. Aerosp. Electron. Syst. AES-18 (1), 102–114. http://dx.doi.org/10.1109/TAES.1982.309210.

Weng, M., Huang, G., Da, X., 2010. A new interframe difference algorithm for moving target detection. In: 2010 3rd International Congress on Image and Signal Processing. Vol. 1, pp. 285–289.

Williams, D.P., 2016. Underwater target classification in synthetic aperture sonar imagery using deep convolutional neural networks. In: 2016 23rd International Conference on Pattern Recognition. ICPR, pp. 2497–2502. http://dx.doi.org/10.1109/ICPR.2016.7900011.

Wu, D., Du, X., Wang, K., 2018. An effective approach for underwater sonar image denoising based on sparse representation. In: 2018 IEEE 3rd International Conference on Image, Vision and Computing. ICIVC, pp. 389–393. http://dx.doi.org/10.1109/ICIVC.2018.8492877.

Xiang, X., Lapierre, L., Jouvencel, B., Xu, G., Huang, X., 2009. Cooperative acoustic navigation scheme for heterogenous autonomous underwater vehicles. In: Inzartsev, A.V. (Ed.), Underwater Vehicles. IntechOpen, Rijeka, http://dx.doi.org/10.5772/6719.

Xiang, X., Yu, C., Niu, Z., Zhang, Q., 2016. Subsea cable tracking by autonomous underwater vehicle with magnetic sensing guidance. Sensors 16 (8), http://dx.doi.org/10.3390/s16081335, URL https://www.mdpi.com/1424-8220/16/8/1335.

Yahya, M.F., Arshad, M.R., 2017. Image-based visual servoing for docking of an autonomous underwater vehicle. In: 2017 IEEE 7th International Conference on Underwater System Technology: Theory and Applications. USYS, pp. 1–6. http://dx.doi.org/10.1109/USYS.2017.8309453.

Yan, J., Li, X., Shao, Z., 2015. Intelligent and fast two-dimensional CFAR procedure. In: 2015 IEEE International Conference on Communication Problem-Solving. ICCP, pp. 461–463. http://dx.doi.org/10.1109/ICCPS.2015.7454202.

Yuan, X., Martínez, J.-F., Eckert, M., López-Santidrián, L., 2016. An improved otsu threshold segmentation method for underwater simultaneous localization and mapping-based navigation. Sensors 16 (7), http://dx.doi.org/10.3390/s16071148, URL https://www.mdpi.com/1424-8220/16/7/1148.

Zhang, T., Liu, S., He, X., Huang, H., Hao, K., 2020. Underwater target tracking using forward-looking sonar for autonomous underwater vehicles. Sensors 20 (1), http://dx.doi.org/10.3390/s20010102, URL https://www.mdpi.com/1424-8220/20/1/102.

Zheng, X., Wang, W., Xiong, M., Xie, G., 2020. Online state estimation of a fin-actuated underwater robot using artificial lateral line system. IEEE Trans. Robot. 36 (2), 472–487. http://dx.doi.org/10.1109/TRO.2019.2956343.