USV-AUV Collaboration Framework for Underwater Tasks under Extreme Sea Conditions

Jingzehua Xu*,+, Guanwen Xie*,+, Xinqi Wang*, Yimian Ding*, Shuai Zhang†
*MicroMasters Program in Statistics and Data Science, Massachusetts Institute of Technology, USA
†Department of Data Science, New Jersey Institute of Technology, USA
Email: sz457@njit.edu

Abstract—Autonomous underwater vehicles (AUVs) are valuable for ocean exploration due to their flexibility and ability to carry communication and detection units. Nevertheless, AUVs alone often face challenges in harsh and extreme sea conditions. This study introduces a unmanned surface vehicle (USV)–AUV collaboration framework, which includes high-precision multi-AUV positioning using USV path planning via Fisher information matrix optimization and reinforcement learning for multi-AUV cooperative tasks. Applied to a multi-AUV underwater data collection task scenario, extensive simulations validate the framework's feasibility and superior performance, highlighting exceptional coordination and robustness under extreme sea conditions. The simulation code will be made available as open-source to foster future research in this area.

Index Terms—Autonomous underwater vehicle, Unmanned surface vehicle, Fisher information matrix, Reinforcement learning, Extreme sea conditions, Underwater tasks

I. INTRODUCTION

Autonomous underwater vehicles (AUVs), especially those multi-AUV systems, have attracted considerable interest for their versatility and mobility in applications such as environmental monitoring, seabed exploration, and biological research [1], [2]. The effectiveness of these vehicles in underwater tasks is largely dependent on precise and reliable positioning and highly efficient control strategies [3], particularly in challenging and extreme sea conditions [4]. Accurate positioning is crucial as it boosts operational efficiency and safety by preventing potential damage or accidents. Moreover, AUVs equipped with adaptive control policies [5] and autonomous decision-making capabilities are better suited to handle dynamic environments [6], providing enhanced resistance to interference and increased practical value [7].

Current positioning techniques for AUVs mainly consist of global positioning systems, inertial navigation systems, and ultrasonic positioning. However, these methods are vulnerable to problems like water refraction and scattering, signal attenuation, transmission delays [3], [8], and error accumulation [5]. In addition, traditional control strategies for multi-AUV systems depend heavily on mathematical models or model-based algorithms [9]. Model-based approaches face difficulties when controller parameters shift in dynamic environments [10], and both algorithms often struggle to predict the future maneuvers and behaviors of non-cooperative targets, thereby limiting their scalability and adaptability [9].

To address the aforementioned challenges, numerous researchers have concentrated on the development of USV-AUV co-localization and intelligent multi-AUV collaboration. Bahr *et al.* introduced a distributed algorithm that leverages multi-AUV to dynamically determine the locally optimal position of the beacon vehicle, utilizing data from the survey vehicle's broadcast communication [11]. Vasilijevi *et al.* constructed the Internet of underwater things using unmanned surface vehicles (USVs) to enhance the underwater positioning efficiency [12]. Jiang *et al.* applied a multi-agent proximal policy optimization reinforcement learning (RL) algorithm to direct efficient and energy-saving data collection for AUVs in unknown environments based on an objective uncertainty map [13]. Wang *et al.* proposed a collaborative data collection strategy for multi-AUV employing local-global deep Q-learning and data value, to facilitate hybrid data collection meeting various temporal demands [14]. Nevertheless, these methods encounter limitations in complex environments: as number of AUVs and sensors increases, the underwater acoustic channel becomes more intricate, computational complexity rises [15], and there are stringent requirements for battery and operational costs of the equipment. Furthermore, traditional AUV swarm control techniques such as heuristic algorithms [16], neural networks [7], and game theory [17], although effective for specific tasks, heavily rely on extensive prior information. In the absence of such information, performance of these methods significantly declines, particularly under extreme sea conditions.

Based on the above analysis, in this study we propose a USV-AUV collaboration framework for underwater tasks to improve the performance of the AUV completing tasks under extreme sea conditions. To make a conclusion, the contributions of this paper include the following:

- We realize accurate positioning of AUVs via USV path planning through minimizing the determinant of the Fisher information matrix (FIM) of the system. Based on this, through integrating environment-aware ability into state space, and USV-AUV collaboration into reward function in the Markov decision process (MDP), we further use RL to empower multi-AUV with intelligence and the adaptability to extreme sea conditions.
- We innovatively leverage two-dimensional tidal wave equations and ocean turbulence model to simulate the extreme sea conditions, which has impact on the positioning

---
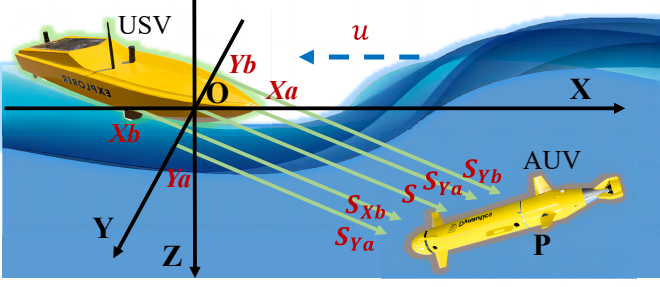
Fig. 1. Illustration of the AUV positioning via USV. The USV on the sea surface uses USBL to locate the underwater AUV.
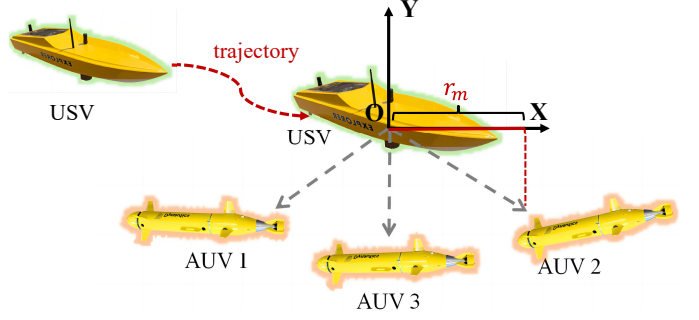

Fig. 2. The diagram of USV path planning. We can realize accurate positioning of AUVs via USV path planning through minimizing the determinant of the system's FIM.

accuracy and working efficiency of multi-AUV.

- Through comprehensive experiments in the underwater data collection task, our framework showcases superior feasibility and excellent performance in balancing multi-objective optimization under extreme sea conditions.

II. METHODOLOGY

In this section, we introduce the proposed USV-AUV collaboration framework, which comprises two main components: high-precision location of multi-AUV using USV path planning via FIM optimization, and RL enabled multi-AUV cooperative work. Besides, we also present simulation of extreme sea conditions using two-dimensional shallow water equation and ocean turbulence model.

A. USV Path Planning Based on Fisher Information Matrix Optimization

Our framework realizes accurate positioning of AUVs via USV path planning through minimizing the determinant of the system's FIM. Central to this consensus is that FIM's determinant is negatively correlated to the system's uncertainty.

Assume the coordinate of the USV is denoted as (x, y, η), and the coordinates of k-th AUV is (x_k, y_k, z_k). As illustrated in Fig. 1, the USV, which is positioned on the sea surface, employs an ultra-short baseline (USBL) system to determine the location of the underwater AUV. The arrays are uniformly spaced on the USV such that the distances $OXa = OXb = OYb = OYa = d/2$, where d represents the array spacing. Consequently, the probability density function of the measured data and the measurement equation can be expressed as follows:

$$p(\boldsymbol{Z}, \boldsymbol{X}) = \prod_{k=1}^{m} \frac{\exp\left\{-\frac{1}{2}[\boldsymbol{Z}_k - \boldsymbol{h}_k(\boldsymbol{X})]^T \boldsymbol{R}^{-1}[\boldsymbol{Z}_k - \boldsymbol{h}_k(\boldsymbol{X})]\right\}}{\sqrt{2\pi \det(\boldsymbol{R})}}, \quad (1a)$$

$$\boldsymbol{Z}_k = \boldsymbol{h}_k(\boldsymbol{X}) + \boldsymbol{u}_k, \quad (1b)$$

$$S_k = \sqrt{(x_k - x)^2 + (y_k - y)^2 + z_k^2}, \quad (1c)$$

where the target state vector is denoted by $\boldsymbol{X} = [x, y]^T$, while $\boldsymbol{h}_k(\boldsymbol{X}) = [\Delta\varphi_{x,k}, \Delta\varphi_{y,k}]^T$ stands for the phase difference vector between receiving units, including two elements $\Delta\varphi_{x,k} = \frac{2\pi f d}{c S_k}(x_k - x)$ and $\Delta\varphi_{y,k} = \frac{2\pi f d}{c S_k}(y_k - y)$, while

c represents the speed of sound and f indicates the signal frequency. Additionally, $\boldsymbol{u}_k$ is characterized as zero-mean Gaussian white noise, and the measurement noise covariance matrix is denoted by $\boldsymbol{R} = \sigma^2 \boldsymbol{I}$.

Then FIM of the system is subsequently obtained by determining the second derivative of the log-likelihood function

$$\boldsymbol{J}_m = \begin{bmatrix} -E\left[\frac{\partial^2 \ln p(\boldsymbol{Z}, \boldsymbol{X})}{\partial x^2}\right] & -E\left[\frac{\partial^2 \ln p(\boldsymbol{Z}, \boldsymbol{X})}{\partial x \partial y}\right] \\ -E\left[\frac{\partial^2 \ln p(\boldsymbol{Z}, \boldsymbol{X})}{\partial y \partial x}\right] & -E\left[\frac{\partial^2 \ln p(\boldsymbol{Z}, \boldsymbol{X})}{\partial y^2}\right] \end{bmatrix}. \quad (2)$$

Assume there are totally m AUVs, the determinant of the FIM can be simplified to the final expression after derivation, which can be expressed as

$$\det(\boldsymbol{J}_m) = \left(\frac{4\pi^2 f^2 d^2}{\sigma^2 c^2}\right)^2 \left[3m\frac{\sin^2\gamma_0}{S_0^4} + \frac{(\sin^4\gamma_0 + 1)^2}{S_0^4}\chi\right], \quad (3a)$$

$$r_m = \arg\max\{\det(\boldsymbol{J}_m)\}, \quad (3b)$$

where $\sin\gamma_0 = \frac{z_k}{S_0}$, while $\chi = \sum_{1 \le i < j \le m}^{m} \sin^2\alpha_{ij}$, with $\alpha_{ij} = \varphi_i - \varphi_j$ representing the angle between the projections of two AUVs and USV. By maximizing the determinant $\det(\boldsymbol{J}_m)$, we can determine the optimal horizontal distance r_m between the USV and multiple AUVs, as illustrated in Fig. 2.

B. Reinforcement Learning Enabled Multi-AUV Collaboration

Our framework leverages RL to train multi-AUV for collaborative operations. Rather than solely relying on standard MDP based RL algorithms, we introduce modifications to design of the state space and reward functions in standard MDP. Specifically, we incorporate the ocean current velocity perceived by AUV k into its state space, represented as

$$s_k = \|\boldsymbol{V}_c(\boldsymbol{P}_k(t))\|. \quad (4)$$

Furthermore, we integrate original reward functions with the distance differential between each AUV and the USV, which can be denoted as

$$r_k(t) = \left(l_{\max}^{k \leftrightarrow U} / l^{k \leftrightarrow U}(t)\right), \quad (5)$$

Algorithm 1: USV-AUV Collaboration Framework

---

1 Initialize USV-AUV's position, the replay buffer $\mathcal{D}$, critic and actor network parameters of each AUV.

2 **for** *each epoch k* **do**

3 | Reset the training environment and parameters.

4 | **for** *each time step i* **do**

5 | | USV path planning via FIM optimization.

6 | | Measure each AUV's position via Eq. (1).

7 | | **for** *each AUV k* **do**

8 | | | Obtain current state information s_i.

9 | | | Sample and execute action a_i while receiving reward r_i.

10 | | | Transfer to the next state s_{i+1}.

11 | | | **while** *In updating period* **do**

12 | | | | Extract N tuples of data $(s_n, a_n, r_n, s_{n+1})_{n=1,\cdots,N}$ from $\mathcal{D}$.

13 | | | | Update the Critic Nework.

14 | | | | Update the Actor Network.

15 | | | **end**

16 | | | Store (s_i, a_i, r_i, s_{i+1}) in $\mathcal{D}$.

17 | | **end**

18 | **end**

19 **end**

---

where $l_{\max}^{k \leftrightarrow U}$ and $l^{k \leftrightarrow U}$ stand for the maximum distance, and indicates the current distance between AUV k and USV. Through extensive epochs of RL training, the collaborative behavior and decision-making capabilities of multi-AUV, enhanced with environment-awareness, will progressively converge to an expert level.

Combining Sections II-A and II-B, we finally propose the USV-AUV collaboration framework, whose algorithm pseudocode has been listed in Algorithm 1.

C. Simulation of Extreme Sea Conditions

Given that the USV operates on the surface and the AUVs navigate underwater, the USV-AUV collaboration framework is highly susceptible to disturbances from severe waves and ocean turbulence.

Based on this intuition, our study employs two-dimensional shallow water equations to simulate the sea surface with wave dynamics. If we denote the wave velocity as $\boldsymbol{V}_w = [u, v]$, and the gravitational acceleration as g, the water level η at coordinate point (x', y') can be calculated by

$$\frac{\partial u}{\partial t} + g \frac{\partial \eta}{\partial x'} = 0, \tag{6a}$$

$$\frac{\partial v}{\partial t} + g \frac{\partial \eta}{\partial y'} = 0, \tag{6b}$$

$$\frac{\partial \eta}{\partial t} + \frac{\partial (u \cdot h)}{\partial x'} + \frac{\partial (v \cdot h)}{\partial y'} = 0, \tag{6c}$$

where h represents the water depth, and we can further derive the expression

$$\eta = R_L \frac{\cos kx'}{\cos kL} e^{-i\omega t}, \tag{7}$$

where R_L denotes the variable associated with the offshore length L, while $k = \frac{2\pi}{\lambda}$, with wave length $\lambda = \frac{2\pi}{\omega}\sqrt{gh}$. Consequently, we observe that $\cos kL$ becomes zero when $L = \frac{\lambda}{4}$, resulting in a significant rise in the water level.

Additionally, this study employs the superposition of multiple viscous vortex functions, derived from the simplified Navier-Stokes equations, to simulate ocean turbulence. The functions are presented as follows:

$$V_x(\boldsymbol{P}_k(t)) = -\Gamma \cdot \frac{y' - y_0}{2\pi \|\boldsymbol{P}_k(t) - \boldsymbol{P}_0\|_2^2} \cdot \left(1 - e^{-\frac{\|\boldsymbol{P}_k(t) - \boldsymbol{P}_0\|_2^2}{\delta^2}}\right), \tag{8a}$$

$$V_y(\boldsymbol{P}_k(t)) = -\Gamma \cdot \frac{x' - x_0}{2\pi \|\boldsymbol{P}_k(t) - \boldsymbol{P}_0\|_2^2} \cdot \left(1 - e^{-\frac{\|\boldsymbol{P}_k(t) - \boldsymbol{P}_0\|_2^2}{\delta^2}}\right), \tag{8b}$$

$$\varpi(\boldsymbol{P}_k(t)) = \frac{\Gamma}{\pi \delta^2} \cdot e^{-\frac{\|\boldsymbol{P}_k(t) - \boldsymbol{P}_0\|_2^2}{\delta^2}}, \tag{8c}$$

where $\boldsymbol{P}_k(t)$ and $\boldsymbol{P}_0$ represent the current location of the AUV k and the coordinate vector of Lamb vortex center, respectively. $V_x(\boldsymbol{P}_k(t))$ and $V_y(\boldsymbol{P}_k(t))$ are the velocities of the ocean turbulence on the X and Y axes perceived by AUV k at position $\boldsymbol{P}_k(t)$ at time t, respectively. Meanwhile, ϖ, δ, and Γ stand for the vorticity, radius, and intensity of the vortex, respectively.

Finally, we can employ the finite difference method to simulate sea waves and the ocean turbulence model, which together constitute the extreme sea conditions in this study.

III. EXPERIMENTS

In this section, we verify the effectiveness of the proposed USV-AUV collaboration framework under extreme sea conditions through comprehensive simulation experiments. Furthermore, we present the experimental results with further analysis and discussion.

A. Task Description and Settings

Since open-source underwater tasks are scarce, we have chosen the multi-AUV data collection task as a representative example to evaluate our framework. This task involves utilizing multi-AUV to gather data from underwater sensor nodes within the Internet of underwater things (IoUT), with multiple objectives such as maximizing the total data rate, avoiding collisions, and minimizing energy consumption, etc. The parameters used in this paper are detailed in Table 1. For additional detail and parameters related to the task, please refer to the previous work [4].

B. Experiment Results and Analysis

To evaluate the feasibility of our framework, we first employed two mainstream RL algorithms, DDPG and SAC, to train two AUVs positioned by the USV on the sea surface to collaboratively complete the underwater data collection task under both ideal and extreme sea conditions, respectively. As illustrated in Fig. 3, the training curves progressively converge
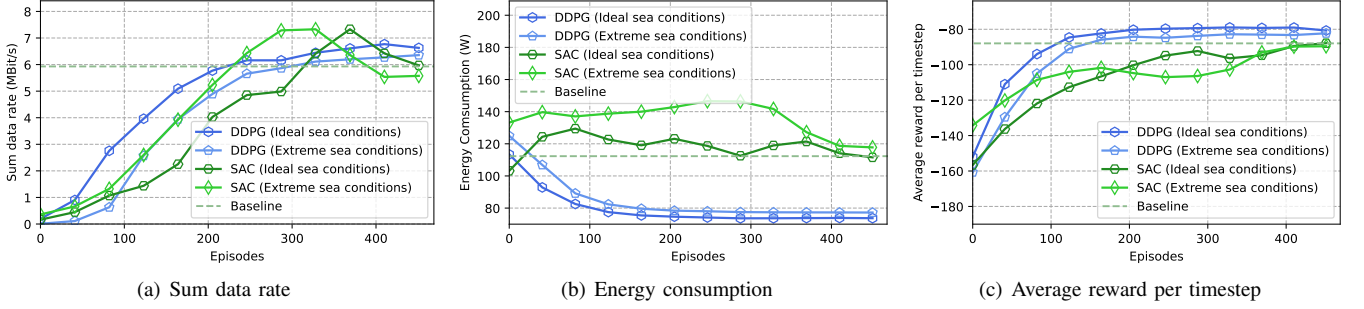
(a) Sum data rate   (b) Energy consumption   (c) Average reward per timestep

Fig. 3. The curves of sum data rate, energy consumption and average reward per timestep, using DDPG and SAC for RL training in ideal and extreme sea conditions, respectively. (a) Sum data rate. (b) Energy consumption. (c) Average reward per timestep.

TABLE I
PARAMETERS OF THE ENVIRONMENT AND PROPOSED
FRAMEWORK IN THIS STUDY.

Parameters	Values
Experimental area	200m × 200m
USV sound level	135dB
USV transmit frequency	15kHz(AUV 1)/18kHz(AUV 2)
Hydrophone array d	0.033m
Water Depth	120m
Time step Δt	10s
Space step Δx	4m
Angular frequency ω	$2\pi/43200$ rad/s
Initial amplitude η_0	5m

TABLE II
COMPARISON OF DIFFERENT RL ALGORITHMS UNDER IDEAL AND
EXTREME SEA CONDITIONS, RESPECTIVELY.

	SDR	EC	ARPS
DDPG (ISC)	6.63±0.97	73.72±1.11	-80.70±6.94
DDPG (ESC)	6.35±1.57	77.23±0.89	-82.28±7.41
SAC (ISC)	5.97±2.66	111.57±13.23	-87.86±13.10
SAC (ESC)	5.58±2.35	117.76±17.30	-89.58±15.22
Baseline [4]	5.93±3.51	112.30±10.56	-87.95±12.08

to expert-level performance, indicating that the AUVs have successfully acquired the expert policy through RL training. Additionally, we performed environmental generalization experiments to compare performance in both ideal and extreme sea conditions. The corresponding results (here ISC, ESC denote ideal and extreme sea conditions, while SDR, EC, ARPS indicate sum data rate, energy consumption, and average reward per timestep, respectively), presented in Table 2, show that despite the presence of ocean waves and turbulence, the performance remains comparable in both scenarios. This demonstrates the framework's high robustness under extreme sea conditions.

Moreover, we utilized the expert policy trained with the DDPG algorithm to guide the multi-AUV system in an underwater data collection task under extreme sea conditions. The trajectories of the AUVs and the USV during a single RL training episode are depicted in Fig. (4a). To further evaluate

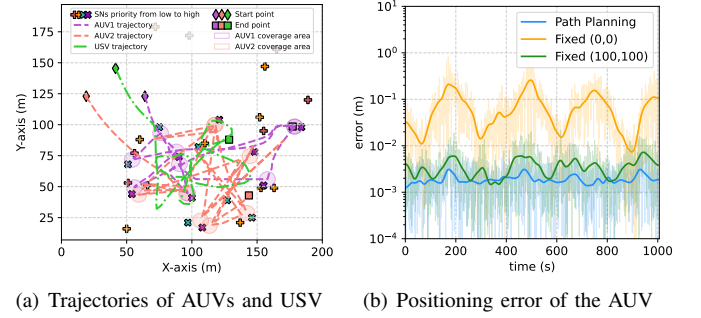(a) Trajectories of AUVs and USV   (b) Positioning error of the AUV

Fig. 4. Trajectories of AUVs and USV, and positioning error of the AUV with USV fixed at (0,0) and (100, 100), and path planning using FIM optimization, respectively. (a) Trajectories of AUVs and USV. (b) Positioning error of the AUV.

the advantages of the USV-AUV collaboration framework, we also presented the positioning error of the multi-AUV system. We examined three different scenarios: employing USV path planning based on FIM optimization, fixing the USV at coordinates (0,0), and fixing it at (100,100). As shown in Fig. (4b), the first approach achieves the lowest positioning error, illustrating the superior performance of USBL positioning via USV path planning using FIM optimization, even under extreme sea conditions.

IV. CONCLUSION

In this paper, we propose a USV–AUV collaboration framework. The two parts of our framework, including high-precision multi-AUV location using USV path planning by FIM optimization and RL training for multi-AUV cooperative tasks, collaboratively enhance the performance of multi-AUV underwater tasks in extreme sea conditions. Experiment results from underwater data collection task verify the superior feasibility and performance of the proposed framework, which effectively embodies high coordination between USV and AUV while showcasing excellent robustness in extreme sea conditions. To accelerate the relevant research in this field, the code for simulation will be released as open-source in the future.

References

[1] X. Hou, J. Wang, T. Bai, Y. Deng, Y. Ren, and L. Hanzo, "Environment-aware auv trajectory design and resource management for multi-tier underwater computing," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 2, pp. 474–490, 2023.

[2] H. Xing, Y. Liu, S. Guo, L. Shi, X. Hou, W. Liu, and Y. Zhao, "A multi-sensor fusion self-localization system of a miniature underwater robot in structured and gps-denied environments," *IEEE Sensors Journal*, vol. 21, no. 23, pp. 27 136–27 146, 2021.

[3] J. Du, B. Jiang, C. Jiang, Y. Shi, and Z. Han, "Gradient and channel aware dynamic scheduling for over-the-air computation in federated edge learning systems," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 4, pp. 1035–1050, 2023.

[4] Z. Zhang, J. Xu, G. Xie, J. Wang, Z. Han, and Y. Ren, "Environment- and energy-aware auv-assisted data collection for the internet of underwater things," *IEEE Internet of Things Journal*, vol. 11, no. 15, pp. 26 406–26 418, 2024.

[5] C. Hu, S. Zhu, Y. Liang, Z. Mu, and W. Song, "Visual-pressure fusion for underwater robot localization with online initialization," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 8426–8433, 2021.

[6] C. Lin, G. Han, M. Guizani, Y. Bi, J. Du, and L. Shu, "An sdn architecture for auv-based underwater wireless networks to enable cooperative underwater search," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 132–139, 2020.

[7] J. Xu, Z. Zhang, J. Wang, Z. Han, and Y. Ren, "Multi-auv pursuit-evasion game in the internet of underwater things: An efficient training framework via offline reinforcement learning," *IEEE Internet of Things Journal*, pp. 1–1, 2024.

[8] L. Zhang, C. Tang, P. Chen, and Y. Zhang, "Gaussian parameterized information aided distributed cooperative underwater positioning algorithm," *IEEE Access*, vol. 8, pp. 64 634–64 645, 2020.

[9] Z. Fang, J. Wang, J. Du, X. Hou, Y. Ren, and Z. Han, "Stochastic optimization-aided energy-efficient information collection in internet of underwater things networks," *IEEE Internet of Things Journal*, vol. 9, no. 3, pp. 1775–1789, 2022.

[10] Y. Wu, K. H. Low, and C. Lv, "Cooperative path planning for heterogeneous unmanned vehicles in a search-and-track mission aiming at an underwater target," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6782–6787, 2020.

[11] A. Bahr, J. J. Leonard, and A. Martinoli, "Dynamic positioning of beacon vehicles for cooperative underwater navigation," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 3760–3767.

[12] A. Vasilijevic, D. Nad, and N. Miskovic, "Autonomous surface vehicles as positioning and communications satellites for the marine operational environment — step toward internet of underwater things," in *2018 IEEE 8th International Conference on Underwater System Technology: Theory and Applications (USYS)*, 2018, pp. 1–5.

[13] B. Jiang, J. Du, C. Jiang, Z. Han, and M. Debbah, "Underwater searching and multiround data collection via auv swarms: An energy-efficient aoi-aware mappo approach," *IEEE Internet of Things Journal*, vol. 11, no. 7, pp. 12 768–12 782, 2024.

[14] J. Wang, S. Liu, W. Shi, G. Han, and S. Yan, "A multi-auv collaborative ocean data collection method based on lg-dqn and data value," *IEEE Internet of Things Journal*, vol. 11, no. 5, pp. 9086–9106, 2024.

[15] J. Du, C. Jiang, J. Wang, Y. Ren, and M. Debbah, "Machine learning for 6g wireless networks: Carrying forward enhanced bandwidth, massive access, and ultrareliable/low-latency service," *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 122–134, 2020.

[16] H. Zhao, J. Yan, X. Luo, and X. Guan, "Ubiquitous tracking for autonomous underwater vehicle with iout: A rigid-graph-based solution," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14 094–14 109, 2021.

[17] K. Zhang, H. Wang, H. Zhang, N. Luo, and J. Ren, "Target tracking of uuv based on maximum correntropy high-order ughf," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–16, 2023.