

## Article

# AUV-Aided Optical—Acoustic Hybrid Data Collection Based on Deep Reinforcement Learning

Fanfeng Bu <sup>1</sup>, Hanjiang Luo <sup>1,\*</sup> , Saisai Ma <sup>1</sup>, Xiang Li <sup>1</sup>, Rukhsana Ruby <sup>2</sup>  and Guangjie Han <sup>3</sup> <sup>1</sup> College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China<sup>2</sup> College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China<sup>3</sup> Department of Information and Communication Systems, Hohai University, Changzhou 213022, China

\* Correspondence: hjliao@sdust.edu.cn

**Abstract:** Autonomous underwater vehicles (AUVs)-assisted mobile data collection in underwater wireless sensor networks (UWSNs) has received significant attention because of their mobility and flexibility. To satisfy the increasing demand of diverse application requirements for underwater data collection, such as time-sensitive data freshness, emergency event security as well as energy efficiency, in this paper, we propose a novel multi-modal AUV-assisted data collection scheme which integrates both acoustic and optical technologies and takes advantage of their complementary strengths in terms of communication distance and data rate. In this scheme, we consider the age of information (AoI) of the data packet, node transmission energy as well as energy consumption of the AUV movement, and we make a trade-off between them to retrieve data in a timely and reliable manner. To optimize these, we leverage a deep reinforcement learning (DRL) approach to find the optimal motion trajectory of AUV by selecting the suitable communication options. In addition to that, we also design an optimal angle steering algorithm for AUV navigation under different communication scenarios to reduce energy consumption further. We conduct extensive simulations to verify the effectiveness of the proposed scheme, and the results show that the proposed scheme can significantly reduce the weighted sum of AoI as well as energy consumption.



**Citation:** Bu, F.; Luo, H.; Ma, S.; Li, X.; Ruby, R.; Han, G. AUV-Aided Optical—Acoustic Hybrid Data Collection Based on Deep Reinforcement Learning. *Sensors* **2023**, *23*, 578. <https://doi.org/10.3390/s23020578>

Academic Editor: Jiachen Yang

Received: 27 November 2022

Revised: 20 December 2022

Accepted: 29 December 2022

Published: 4 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** autonomous underwater vehicles; optical–acoustic multi-modal communication; data collection; path planning; deep reinforcement learning

## 1. Introduction

Accompanied by the increasing demand for ocean exploration and protection, underwater sensor networks (UWSNs) have received more attention as these play an important role in diverse marine applications, such as coastal monitoring and protection, marine resource exploration, disaster warning and military operations [1–4]. However, due to the harsh hydrographic and geographical environment, it is difficult to collect data from underwater sensor devices via a long-range routing path. Even if the monitored data can be transmitted through multi-hop routing technologies, there may be heavy workload near the sink with extra energy consumption [5]. Furthermore, as the battery power of underwater sensor nodes is severely limited and difficult to be recharged underwater, it is not energy-efficient to upload large volume of ocean monitoring data to the sink directly. Moreover, with marine security operations, it is better to collect secret data nearby the monitoring sensors. To solve the aforementioned problems, autonomous underwater vehicles (AUVs) have been rapidly developed in recent years in terms of data storage and signal processing capabilities, which can better enable underwater mobile data collection. Moreover, the durability and mobility of AUVs alleviate the unbalanced energy consumption problem of underwater sensors [6,7].

To collect data in an efficient manner, various underwater communication technologies have been investigated, such as acoustic and optics [8]. Currently, although underwater acous-

tic communication (UAC) has become the most widely used technology due to its unique advantages (e.g., long-range communication), it is limited by its shortcomings (e.g., low bandwidth, slow speed, high bit error rate and large delay) [9]. To address these issues, underwater optical communication (UOC) has emerged as an alternative solution, as it has a higher propagation speed ( $2.255 \times 10^8$  m/s) and higher data rate (up to hundreds of Mbit/s) over short to medium-range transmissions [10,11]. As both acoustic and optical communication have their pros and cons, employing multi-modal underwater communication systems in UWSNs has become a potential approach to improve network performance [12,13].

To facilitate mobile data collection in such multi-modal networks, it is necessary to satisfy the varying requirements of marine applications by combining the potential advantages of AUVs [14]. Combined multi-modal data collection via AUV is divided into two categories, such as acoustic multi-modal and acoustic-optical multi-modal. In acoustic multi-modal data collection, the sensor node transmits control information using low-frequency acoustic waves and guides the AUV to the designated area, and then, it switches to a high-frequency UAC modem to transmit the data [15]. In this case, the high energy consumption of the UAC shortens the lifetime of the sensor node when transmitting large volume of network data. Whereas, in the acoustic-optical multi-modal data collection, the UAC provides the capability for the AUV to approach the sensor node through long-range guidance and assists with alignment for the optical communication. The subsequent proximity of the UOC data transmission not only improves data transmission efficiency but also saves energy for transmission [16]. However, limited by the UOC range, it is necessary for the AUV to move slowly close to the sensor node to build an optical link reliably, which increases traveling time. A promising solution to the above problem is to engage both the UAC and UOC in data collection, such as transmitting a small volume of data over long distances using the UAC and retrieving and offloading large amounts of data using the UOC [17,18].

Although the aforementioned pioneering studies have laid a solid foundation for multi-modal data collection, there are still some issues when applied to mobile data collection. Firstly, the AUV should adaptively select the best communication technology according to the specific marine operational data requirements (e.g., data importance and packet size). For example, high-quality data collection (e.g., high-resolution images with 4K size) with UOC can prolong the network lifetime of the UWSN by sacrificing the energy consumption of the AUV, and when the volume of data is relatively less, UAC can be used for remote collection to reduce the energy consumption and travel time of the AUV. Furthermore, the AUV should complete the data collection operation quickly to guarantee the freshness of data as the data value usually decays over time [19]. Generally, the age of information (AoI) can be used to measure data freshness in mobile data collection scenarios [5,20,21]. By optimizing AoI, the requirement of the network for timely data delivery can better be satisfied. Therefore, in a multi-modal AUV-enabled mobile data collection scheme, how to optimize the trajectories of AUV and select a communication option to minimize both AoI and energy consumption based on the size and the importance of the packets is a critical issue.

To solve the aforementioned issues, in this paper, we propose an acoustic-optical multi-modal mobile data collection scheme. Based on the type and the size of data, the AUV intelligently searches for the optimal trajectory and communication options using the deep reinforcement learning (DRL) approach, thereby minimizing the AoI and extending the lifetime of the sensor network. To the best of our knowledge, this is the first study which focuses on integrating an acoustic-optical multi-modal option with optimal AUV path planning for reliable and timely mobile data collection leveraging the DRL approach. The main contributions of this study are listed as follows.

- We investigate an AUV-assisted underwater trajectory planning problem for data collection by integrating the complementary advantages of both acoustic and optical communication with data diversity to perform reliable and timely mobile data collection.
- We propose a DRL-based AUV-assisted multi-modal mobile data collection scheme in which we consider several key factors, such as data importance, packet size and data collection option, to minimize AoI and reduce energy consumption.

- We propose an optimal angle steering algorithm for AUV navigation to reduce energy consumption, in which the steering angle of the AUV is determined based on the AUV and sensor positions as well as the data collection option.

The rest of the paper is structured as follows. We briefly review the related works in Section 2. In Section 3, we introduce the network model with necessary background. In Section 4, we analyze the problem of the multi-modal data collection. In Section 5, we describe the proposed scheme of DRL-based multi-modal data collection in detail. We evaluate the performance of the proposed scheme in Section 6. Finally, we conclude the paper in Section 7.

## 2. Related Works

In recent years, multi-modal communication has become a research topic to improve network performance and optimize data transmission in various marine application scenarios. Commonly adopted multi-modal technologies include acoustic multi-modal communication and acoustic–optical hybrid communication [13,22–24]. Among them, the acoustic multi-modal communication is constructed by a set of UAC modems working on different frequency bands [13]. In [22], the authors proposed a multi-modal underwater routing protocol based on the reinforcement learning technique. In this protocol, the reliability and delay of data transmission are optimized by UAC modems in multiple frequency bands. To explore the advantages of UAC and UOC during data transmission, Shen et al. proposed an acoustic–optical multi-modal routing scheme based on packet size and link adaptation, which reduces packet loss and end-to-end delay [23]. However, the challenge of unbalanced energy consumption still exists in the multi-hop underwater networks.

The AUV-assisted data collection can mitigate the energy consumption unbalance that occurs in multi-hop routing. Han et al. [18] explored the characteristics of underwater acoustic and optical communication in AUV-assisted data collection and showed that hybrid acoustic–optical data collection outperforms the one with a single acoustic modem in terms of both throughput and energy consumption. To cope with the impact of the harsh underwater environment, Luo et al. [25] maximize the network throughput by capturing the dynamic characteristics of the channel and the mobility of the AUV. Hu et al. [17] proposed a mobile data collection method for the heterogeneous sensor network using multi-hop acoustic communication to build an intra-cluster network where CHs collect large-scale data and upload them to a mobile receiver via optical communication. Although the aforementioned schemes improve the efficiency of data collection, these ignore the difference in the importance of data and the decay of freshness over time.

To handle the aforementioned issues, Gjanci et al. [16] proposed a greedy adaptive navigation algorithm to guide AUV for data collection, which considers the characteristics of data decay, but it is only applicable to sparse networks due to the unavoidable long paths. To deliver emergency data faster, Liu et al. [26] proposed a hybrid data collection scheme, in which the urgent data are routed using a multi-hop scheme, while the delay-insensitive data are collected by AUVs. Duan et al. [27] studied a hierarchical data collection problem using AUVs to optimize the information quality of the collected data while considering the importance and timeliness of the events.

More recently, researchers have proposed the concept of AoI to model the timeliness of data while considering the quality of experience (QoE) [28]. Khan et al. [29] provided an optimization algorithm to ensure the freshness of the collected data. Fang et al. [5] used a vocational queuing model to improve the data reliability and peak AoI of the data. Then, the communication link is established using the UAC when the AUV arrives near the node. Al-Habob et al. [30] proposed a framework to optimize the trajectories of AUV and minimize the normalized weighted sum of the average AoI. Wu et al. [31] studied the AUV transmission scheduling policy by considering both the age and the importance of the message.

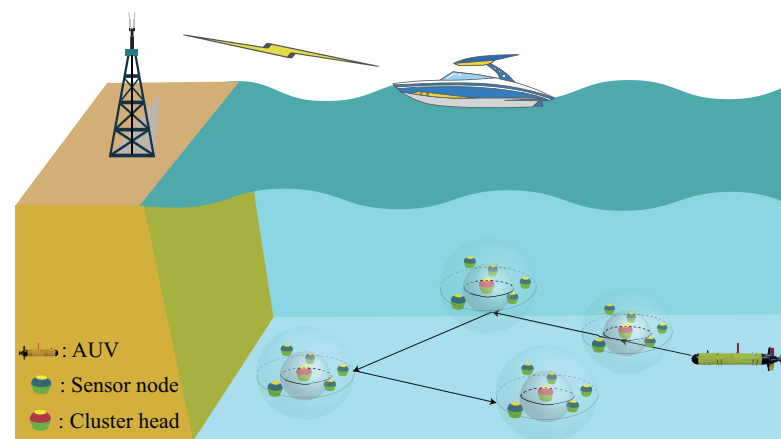
Although the aforementioned approaches have promoted the study of underwater data collection, only a single communication technology was considered for the data collection.

Moreover, none of them addressed the issue of leveraging multiple data types and multiple communication technologies to improve data freshness and energy efficiency. To address this issue, in this paper, we propose a AUV-assisted acoustic and optical multi-modal data collection scheme, in which we use the DRL method to optimize AUV trajectories, AoI and energy consumption by considering different communication options and packet size.

### 3. Network Model

#### 3.1. Network Architecture

As shown in Figure 1, we consider an AUV-based multi-modal data collection network where the deployed nodes are classified into ordinary nodes  $S = \{s_1, s_2, \dots, s_M\}$ , cluster heads  $CHs = \{c_1, c_2, \dots, c_N\}$  and sink node according to their different functions. The sensor nodes are statically deployed on the seabed using anchor chains, where the locations of the nodes are assumed to be known. The CHs perform intra-cluster data fusion and data compression [27] and then wait for the AUV to arrive and collect the data. In particular, during the network formation phase, all sensor nodes are divided into multiple clusters based on spatial distance, and only one node in each cluster is selected as the CH, while other nodes are used as ordinary nodes for data collection [32]. The AUV performs global data collection around all CH nodes and finally reports the data to the sink node.



**Figure 1.** Illustration of the underwater network model.

In the multi-modal network, each node is equipped with both UAC and UOC modems for multi-modal communication, and they have the same initial energy, sensing and communication capabilities. Specifically, it includes an acoustic modem for exchanging data at a low transmission rate over a long distance and an optical modem with a relatively short transmission distance and high data rate [33]. Meanwhile, the AUV has similar communication capability to ensure the data transmission [16]. Without loss of generality, we assume that the data arrival rate of sensor node obeys a Poisson random distribution with parameter  $\lambda$ . When the AUV visits  $c_i$ , the CHs package the sampled data block into the packets of length  $B_i$  with timestamp  $T_i$ .

#### 3.2. Node Clustering Phase

We assume that the nodes are randomly deployed in the target area to monitor the underwater environment, and the nodes are clustered. In the initial phase, the sink nodes know the location of each node and determine the number of clusters based on the network size, and then, the target area is divided equally into several square areas. The sink node broadcasts the subregion message to all nodes, and each node determines its own cluster identifier based on its position. Nodes with the same identifier belong to the same cluster [26].

The CHs should be selected for inter-cluster data collection and communication with the AUV. The selection of CH is carried out according to the procedure as follows. The number of optical and acoustic communication neighbors of each node in the sub-

region is first obtained, and then, the node with the highest number of optical communication neighbors and the remaining energy satisfying the energy threshold requirement is selected as the CH. Then, the above operation is repeated until all CHs in the target region are determined. Finally, a confirmation packet is sent by the sink to the designated CHs. At the end of the data collection process, all CHs are evaluated, and when the energy of the CH is less than the energy threshold, the network performs a new CH selection round.

### 3.3. Acoustic Data Collection Link

When the AUV traverses near the node, it is necessary to construct a communication link for data collection. As for the acoustic link, the acoustic wave is affected by the absorption of medium and the scattering of impurities in water. The path loss of underwater acoustic channels is related to frequency  $f$  and distance  $d_{ac}$ . To this end, the total attenuation is given as follows [34].

$$A(d_{ac}, f) = d_{ac}^k a(f)^{d_{ac}}, \quad (1)$$

where  $k = 1.5$  represents propagation loss, and  $a(f)$  is the absorption coefficient in dB/km given by the Thorp formula [35]

$$10 \log a(f) = 0.11 \frac{f^2}{1 + f^2} + 44 \frac{f^2}{4100 + f^2} + 2.75 \times 10^{-4} f^2 + 0.003. \quad (2)$$

Consequently, given the acoustic signal transmit power  $P_{trans}^{ac}$  and frequency  $f$ , the signal-to-noise ratio (SNR) can be expressed as [36]

$$SNR_{ac}(d_{ac}, f) = \frac{P_{trans}^{ac} / A(d_{ac}, f)}{N(f) \Delta f}, \quad (3)$$

where  $N(f)$  and  $\Delta(f)$  represent the total noise level including four kinds of interference noise and the bandwidth of the receiver, respectively. Therefore, the transmission power of acoustic communication satisfying the minimum  $SNR_{min}^{ac}$  is expressed as

$$P_{trans}^{ac} = SNR_{min}^{ac} A(d_{ac}, f) N(f) \Delta f. \quad (4)$$

### 3.4. Optical Data Collection Link

For the optical link, the path loss  $PL$  of the underwater wireless optical link can be expressed as [37]

$$PL \approx 10 \log \left( \left( \frac{D_r}{2\theta d_{op}} \right)^2 e^{-cd_{op}} \left( \frac{D_r}{\theta d_{op}} \right)^\zeta \right), \quad (5)$$

where  $D_r$  represents the aperture diameter of the receiver and  $\theta$  denotes half of the transmitter beamwidth,  $d_{op}$  represents the distance between transceivers, and  $c$  and  $\zeta$  represent the extinction coefficient and turbidity of water quality, respectively. Subject to the optical-to-electric conversion efficiency of the receiver, a minimum received power per bit is defined as  $P_{rec}^{op}$ . Then, the transmission power of the UOC is expressed as

$$P_{trans}^{op} = \frac{P_{rec}^{op}}{PL}. \quad (6)$$

In order to ensure robust optical communication, it is necessary to control the minimum  $SNR_{min}^{op}$  requirements [38].

$$SNR_{min}^{op} = \left[ \frac{P_{trans}^{op} e^{-cd_{op}} D_r^2 \cos \varphi}{(\tan^2 \theta) 4d_{op}^2 NEP} \right]^2, \quad (7)$$



where  $NEP$  represents the noise equivalent power and  $\varphi$  is the offset angle between transceivers. According to the Lambert W function [39], the maximum underwater optical communication distance while satisfying the communication  $SNR_{\min}^{op}$  can be obtained by

$$d_{op} = \frac{2W\left(\frac{c}{4}\left[\frac{(SNR_{\min}^{op})^{\frac{1}{2}}NEP\tan^2\theta}{P_{trans}^{op}D_r^2\cos\varphi}\right]^{-\frac{1}{2}}\right)}{c}. \quad (8)$$

Since optical modems are usually directional, in order to receive optical signals from any direction, we assume that an omni-directional optical modem can be achieved by using multiple LEDs [40].

#### 4. Multi-Modal Data Collection Analysis

When performing multi-modal data collection via AUV, the freshness of the collected data and the energy consumption of the network nodes need to be fully considered. The choice of communication options fundamentally affects the data collection efficiency. Among the various communication options, the UOC is capable of transmitting a large volume of data rapidly to reduce transmission latency but increases the navigation time and energy consumption of the AUV. Meanwhile, the UAC has a lower bandwidth but can collect small volume of data over long distances to reduce the travel time of the AUV. Consequently, to collect data in a timely and efficient manner, several key factors, such as the data collection option, data type, packet size and AoI requirement, should be fully analyzed and integrated into the optimal path-planning scheme for data collection.

##### 4.1. Problem Analysis

The primary goal of the mobile data collection in this paper is to minimize both the weighted average AoI and the energy consumption. The factors that influence the AoI include the AUV trajectory, the data transfer time and the importance of the data. Consequently, to minimize the weighted average AoI, the optimization problem can be expressed as

$$\min \frac{1}{N} \sum_{i=1}^N A_i + \beta \sum_{i=1}^N e_i, \quad (9)$$

$$\text{s.t. } x_{i,t}^{ac} \in \{0, 1\}, \quad (9a)$$

$$x_{i,t}^{op} \in \{0, 1\}, \quad (9b)$$

$$\sum_{t=1}^T x_{n,t}^{ac} + x_{n,t}^{op} = 1, \quad \forall n \in N \quad (9c)$$

$$E_{AUV} < energy_{AUV}, \quad (9d)$$

$$P_0 = (x_0, y_0), \quad (9e)$$

where  $A_i$  denotes the final result of AoI when the data from  $c_i$  reach the sink node,  $e_i$  denotes the energy consumption by  $c_i$  during data collection, and  $x_{i,t}^{ac} = 1$  indicates that the AUV has reached the acoustic communication range of CH  $c_i$  and receives data through UAC. Otherwise,  $x_{i,t}^{ac} = 0$  holds. Similarly,  $x_{i,t}^{op}$  represents an optical communication indicator.  $E_{AUV}$  is the energy consumption of the AUV in data collection, and  $energy_{AUV}$  indicates the initial total energy of the AUV. The constraint in (9c) is used to ensure that each node can select just one communication option during data collection. The constraint in (9d) is to guarantee that the AUV cannot consume all of its energy. Finally, the constraint in (9e) is to determine the initial position of AUV.

The optimization problem (9) is a non-linear integer programming problem, which is intractable due to the presence of binary variables and non-convex objective function. In the following section, we model this as a Markov decision process (MDP) to be solved by leveraging the DRL approach.

#### 4.2. Definition of AoI

The AoI is an important metric to portray the freshness of collected data and is defined as the time elapsed between the data collected by the AUV from the CHs until its delivery to the sink node [41]. We use  $\delta_{i,t}$  to denote the AoI collected from  $c_i$  in the navigation trajectory at time  $t$ . When  $t < T_i$ , the information of CH  $c_i$  is not sampled since it is not visited, and thus,  $\delta_{i,t} = 0$  holds. Otherwise,  $\delta_{i,t} = t - T_i$  holds. Then, the AoI of  $c_i$  at the start of time slot  $t$  is given by the following relation.

$$\delta_{i,t} = \begin{cases} 0, & \text{if } t < T_i \\ t - T_i, & \text{otherwise} \end{cases} \quad (10)$$

The primary factors affecting AoI during data collection include data transmission delay and AUV sailing time. We use  $\mathcal{T}_i^{ac}$  and  $\mathcal{T}_i^{op}$  to denote the data transmission time using UAC and UOC, respectively. The time to transmit  $B_i$  bits by the UAC can be written as

$$\mathcal{T}_i^{ac} = \frac{B_i}{R_{ac}} + \frac{d_i}{V_{ac}}, \quad (11)$$

where  $R_{ac}$  and  $V_{ac}$  indicate the data rate and transmission velocity of the underwater acoustic modem, respectively. Similarly, the data transmission time of the UOC at data rate  $R_{op}$  and transmission velocity  $V_{op}$  can be obtained as follows.

$$\mathcal{T}_i^{op} = \frac{B_i}{R_{op}} + \frac{d_i}{V_{op}}. \quad (12)$$

To collect the monitored data, the AUV travels from the sink  $p_0$ , collects data from each of the  $N$  CHs according to a pre-determined trajectory, and then returns to the sink node after completing the task. Assume that the travel trajectory of the AUV  $P = \{p_0, p_i, \dots, p_j, p_0\}$ , and thus the travel time of the AUV can be expressed as

$$\mathcal{T}_{travel} = \frac{D(P)}{V_{AUV}}, \quad (13)$$

where  $D(P)$  and  $V_{AUV}$  denote the total distance and velocity traveled by the AUV, respectively.

According to (11)–(13), from the moment  $T_i$  when AUV arrives at CH  $c_i$  to the moment  $T_{i+1}$  when it finishes collecting data and moves to the next data collection point, the AoI of CH  $c_i$  can be expressed as  $\mathcal{T}_i^m + t_{i,i+1}^{travel}$ . The optical communication has a much smaller transmission delay compared to the acoustic waves. However, the acoustic communication enables long-range transmission that significantly reduces the travel time of the AUV. The time delay caused by data transmission is mainly determined by the data size  $B_i$ , and so the data collection time and traveling time need to be considered jointly to reduce the decline of data freshness. Then, at moment  $t = T_{i+1}$ , the AoI collected from CH  $c_i$  refers to

$$\delta_{i,T_{i+1}} = \begin{cases} \mathcal{T}_i^{ac} + t_{i,i+1}^{travel}, & \text{if } b_i = 1 \text{ and } x_{i,t}^{ac} = 1 \\ \mathcal{T}_i^{op} + t_{i,i+1}^{travel}, & \text{if } b_i = 1 \text{ and } x_{i,t}^{op} = 1 \end{cases} \quad (14)$$

where  $b_i = 1$ ,  $i = \{1, 2, \dots, N\}$  indicates that the data of CH  $c_i$  has been collected; otherwise,  $b_i = 0$  holds. When the AUV arrives at the sink node, the AoI of  $c_i$  is

$$A_i = \sum_{k=i}^N \eta_i \delta_{k,T_{k+1}}, \quad (15)$$

where  $\eta_i$  denotes the importance weight of the data collected by CH  $c_i$ , and  $\sum_{i=1}^N \eta_i = 1$ . The higher its value, the greater the data importance.

#### 4.3. Energy Consumption Associated With Data Collection

To satisfy the energy constraint (9d) in the optimization problem (9), we analyze the AUV energy consumption and node energy consumption. In the data collection process, there are extra costs associated with the AUV if it runs out of energy before returning to the sink node. Therefore, the trajectory of the AUV should be scheduled to minimize energy consumption. The power of the AUV at each time slot mainly consists of the sum of propulsion power  $\Phi_{prop}$  and hotel load power  $\Phi_H$  [42]. The hotel load  $\Phi_H$  is the power consumed by all subsystems other than propulsion mechanism and is typically negligible in comparison with  $\Phi_{prop}$  [43]. Therefore, the power of the AUV trajectory can be expressed as

$$\Phi_{prop} = \frac{\rho}{2\eta_p} C_D A_s \|V_{AUV}\|^3, \quad (16)$$

where  $\|\cdot\|$  denotes the Euclidean vector norm and  $\rho$  is the density of water.  $\eta_p$ ,  $C_D$  and  $A_s$  indicate the efficiency of the AUV's propulsion system, the drag coefficient and the wetted surface area, respectively [7]. Consequently, with the relations in (4), (6), (11)–(13) and (16), the total energy consumption can be expressed as

$$\begin{aligned} E_{tot} &= \sum_{i=1}^N e_i + \omega \Phi_{prop} \mathcal{T}_{travel} \\ &= \sum_{i=1}^N x_i^{ac} \mathcal{T}_i^{ac} P_{trans}^{ac} + x_i^{op} \mathcal{T}_i^{op} P_{trans}^{op} + \omega \Phi_{prop} \mathcal{T}_{travel}, \end{aligned} \quad (17)$$

where  $\omega$  is a weighted parameter that measures the balance between the energy consumption of the sensor node and that of the AUV.

### 5. Proposed DRL-Based Multi-Modal Data Collection Scheme

In this section, we design the AUV multi-modal data collection scheme by leveraging the DRL approach. In this scheme, we first provide the MDP formulation and then present a multi-modal steering angle optimization (MSAO) algorithm for the AUV. Afterwards, we design the AUV path planning using the Deep Q Network (DQN) method for multi-modal mobile data collection.

#### 5.1. MDP Formulation

When the network nodes are clustered, the next goal is to find an optimal CHs data collection strategy. The AUV-assisted data collection problem can be formulated as an MDP to be solvable by the DRL approach, which is represented by  $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$  five tuples. Here,  $\mathcal{S}$  is the state of the environment,  $\mathcal{A}$  is the set of actions of the agent,  $\mathcal{P}$  is the state transition probability,  $\mathcal{R}$  is the reward function, and  $\gamma$  denotes the discount factor. In particular, at time slot  $t$ , the agent observes state  $s_t$  and chooses an action to be performed. Then, the environment state is transferred with probability  $p_{s_t, s_{t+1}}$  to  $s_{t+1}$  and the agent obtains a reward  $r_t$  from the environment. In this paper, the AUV is considered as the agent to collect data, and the details of each element are defined as follows.

- State space  $\mathcal{S}$ : The status of AUV mobile collection is defined as

$$s_t = \left\{ p_{a,t}, \psi_t, \Delta_t, \{x_{i,t}^{ac}, x_{i,t}^{op}, d_{i,t}, \delta_{i,t}\}_{i \in N} \right\}, \quad (18)$$

where  $p_{a,t}$  and  $\psi_t$  are the coordinates and sailing orientation of the AUV at time slot  $t$ , and its position can be obtained via ultra-short baseline (USBL) [44].  $\Delta_t$  is the difference between the remaining energy of the AUV and the AUV's arrival at its final destination from its current position.  $d_{i,t}$  records the Euclidean distance of the AUV to CH  $c_i$ .  $x_{i,t}^m$  is the data collection indicator related to the data collection option. When the AUV arrives at the data collection point  $c_i$ , the AoI of node  $c_i$  starts to be updated.



- Action space  $\mathcal{A}$ : In state  $s_t$ , the action selection of the AUV is characterized by the target point  $c_{i,t} \in N_r$  with the transmission option  $m_{i,t}$ , and the next target point  $c_{j,t} \in N_r \setminus c_{i,t}$ , where  $N_r$  is the set of CHs that have not been collected. Then, the action performed by AUV at state  $s_t$  can be expressed as

$$a = \{c_{i,t}, m_{i,t}, c_{j,t} | s_t\}. \quad (19)$$

- State transition probability  $\mathcal{P}$ :  $\mathcal{P}(s_{t+1} | s_t, a_t)$  defines the transition probability from state  $s_t$  to the next state  $s_{t+1}$  under the action  $a_t$ , and  $\mathcal{P}(s_{t+1} | s_t, a_t) = 1$  holds.
- Reward  $\mathcal{R}$ : Applying action  $a_t$  in state  $s_t$ , the AUV enters state  $s_{t+1}$  and obtains an immediate reward  $r(s_{t+1} | s_t, a_t)$ . In the AUV-assisted multi-modal data collection scenario, the immediate reward  $r_t$  can be expressed as

$$r_t = \begin{cases} x_{i,t}^{ac} k_1 e_i + x_{i,t}^{op} k_2 e_i, & \text{if } b_i = 1 \\ J, & \text{if done} \\ \eta_i (dis_{p_a, p_i} + 1), & \text{otherwise} \end{cases}, \quad (20)$$

where  $k_1, k_2$  are constants and  $k_1 < k_2$  holds, and when the AUV has collected the data of CH  $c_i$ , the relevant reward is obtained according to the selected modem.  $dis_{p_a, p_i}$  is the Euclidean distance from the current position of the AUV to the target point.  $J$  denotes the reward at the end of the data collection process, including rewards for successful data collection and penalties for failure (e.g., exceeding maximum energy consumption and crossing boundaries).

$$J = \begin{cases} r_{out}, & \text{if } \Delta t < 0 \text{ or } p_a \notin \Omega \\ k_3 - \frac{1}{N} \sum_{i=1}^N A_i, & \text{if } p_a = p_0 \text{ and } N_r = \emptyset \end{cases}, \quad (21)$$

where  $k_3$  is a constant and  $\Omega$  is the region in which the AUV can move within.

- Discount factor  $\gamma$ :  $\gamma \in [0, 1]$  is the future reward discount factor.

### 5.2. Multi-Modal Steering Angle Optimization Algorithm

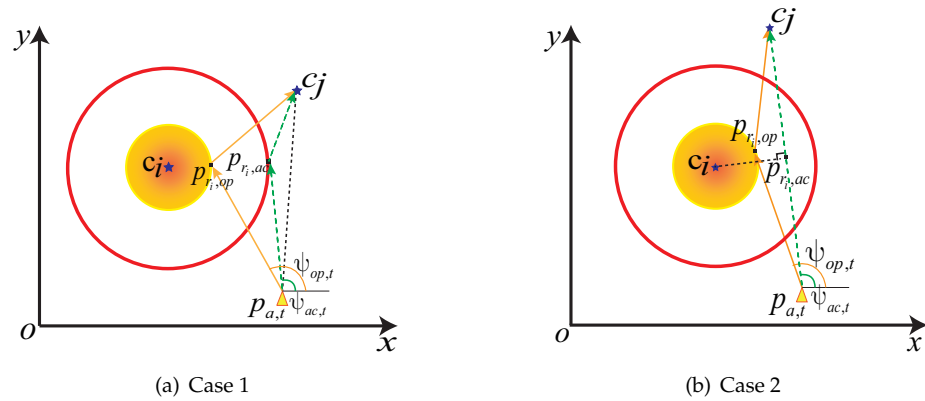
In the multi-modal data collection network, since the communication radius can reduce the navigation time of AUVs, we propose an MSAO algorithm to adjust the AUV heading under the maximum steering angle constraint. In MSAO, the steering angle of the AUV is calculated based on its position, the navigation target and the communication options. As shown in Figure 2, the yellow triangle indicates the position  $p_{a,t}$  of the AUV at time slot  $t$ , the blue pentagram indicates the CHs that need to perform the data collection operation, the outer circle  $C_{ac}$  and inner circle  $C_{op}$  indicate the communication range of UAC and UOC, respectively. Let  $c_i$  be the AUV's current target CH and  $c_j$  be the next target CH, the  $p_{r_i,m}$  indicates the target hover point when the AUV selects communication option  $m = \{ac, op\}$ , the  $\psi_{m,t}$  indicates the angle of the AUV toward the target hover point  $p_{r_i,m}$  at time slot  $t$ . The goal is to obtain the point  $p_{r_i,m}$  such that the  $\|p_{a,t} - p_{r_i,m}\| + \|c_j - p_{r_i,m}\|$  distance is shortest within the communication range  $C_m$  of the communication options  $m$ . This problem is a classical pilgrimage problem in ancient castles, and hence, an approximate solution of  $p_{r_i,m} = (x_{r_i}, y_{r_i})$  can be obtained by the following equation [45].

$$-\varsigma_1 \sqrt{1 - y^2} + \varsigma_2 y = 0, \quad (22)$$

where  $\varsigma_1 = \frac{x_i - x_j}{d_m d_{ij}} - \frac{x_i - x_{a,t}}{d_m d_{ai,t}}$ ,  $\varsigma_2 = \frac{y_i - y_{a,t}}{d_m d_{ai,t}} - \frac{y_i - y_j}{d_m d_{ij}}$ ,  $y = \frac{y_i - y_{r_i}}{d_m}$ ,  $d_{ai,t}$  and  $d_{ij}$  denote the distance of the AUV from the target at time slot  $t$  and the distance of the current target CH  $c_i$  from the next target CH  $c_j$ , respectively. Then, the steering angle of the AUV at time slot  $t$  can be expressed as

$$\Psi_{m,t} = \begin{cases} \min(\psi_{m,t} - \psi_t, \psi_{\max}), \psi_{m,t} \geq \psi_t \\ \max(\psi_{m,t} - \psi_t, -\psi_{\max}), \psi_{m,t} < \psi_t \end{cases} \quad (23)$$

where  $\psi_{\max}$  is the maximum steering angle allowed by the AUV. Then, depending on the target location and the communication option, the steering angle of the AUV can be adjusted in the following two cases.



**Figure 2.** The AUV multi-modal steering angle diagram. (a) The trajectory of the AUV from the current coordinates  $p_{a,t}$  towards the next target  $c_j$  will not sail through the communication region  $C_m$  of  $c_i$ . (b) The trajectory of the AUV from the current coordinates  $p_{a,t}$  towards the next target  $c_j$  will sail through the communication region  $C_m$ .

- Case 1: The AUV is not through the region  $C_m$  from the current position  $p_{a,t}$  to the next target collection point  $c_j$ ; i.e., the distance  $d_{seg_i}$  from point  $c_i$  to the segment  $\overline{p_{a,t}c_j}$  is greater than the UAC radius. As shown in Figure 2a, after determining the communication option, the points  $p_{r_i,ac}$  (or  $p_{r_i,op}$ ) are obtained in circle  $C_{ac}$  (or  $C_{op}$ ) to minimize the length of the AUV trajectory. For example, when the CH  $c_i$ ,  $c_j$  and acoustic modem are selected, the AUV hover position  $p_{r_i,ac} = (x_{r_i}, y_{r_i})$  for data collection and the steering angle  $\Psi_{ac,t}$  can be calculated by (22) and (23), respectively. Similarly, when  $m = op$  holds, the data collection hover point  $p_{r_i,op}$  and the steering angle  $\Psi_{op,t}$  can be obtained using the same approach.
- Case 2: The trajectory of the AUV from the current coordinate  $p_{a,t}$  to the next target CH  $c_j$  sails through the communication region  $C_m$  of  $c_i$ . If the AUV crosses the UAC area  $C_{ac}$  without crossing the communication area  $C_{op}$ ,  $d_{seg_i}$  becomes shorter than  $d_{ac}$  but greater than  $d_{op}$ . As shown in Figure 2b, the data collection hover point of the AUV is the vertical foot  $p_{r_i,ac}$  from  $c_i$  to segment  $\overline{p_{a,t}c_j}$  if UAC is selected as the communication option. Then, the steering angle of the AUV can be obtained by (23). If the selected communication option is UOC, the data collection point and steering angle are calculated following the method in Case 1. Furthermore, if  $d_{seg_i}$  is less than  $d_{op}$ , i.e., the AUV crosses the UOC range of  $c_i$ , then UOC is selected directly as the communication option. This is due to the superiority of UOC over UAC in terms of energy consumption and transmission time for the same AUV trajectory. The data collection hover point and steering angle of the AUV are similar to the method in Case 2.

Based on the above discussion, we obtain the MSAO algorithm that is shown in Algorithm 1.

**Algorithm 1** Proposed MSAO Algorithm

**Require:** Coordinate of the AUV  $p_{a,t}$ , coordinates of the current target CH  $c_i$ , coordinates of the next target CH  $c_j$ , UAC communication radius  $d_{ac}$  and UOC communication radius  $d_{op}$ .

- 1: **if**  $d_{seg_i} > d_{ac}$  and  $m = ac$  **then**
- 2:   Calculate the data collection hover position  $p_{c_i,ac}$  by (22).
- 3:   Calculate steering angle  $\Psi_{ac,t}$  by (23).
- 4: **else if**  $d_{seg_i} \leq d_{ac}$  and  $d_{seg_i} > d_{op}$  and  $m = ac$  **then**
- 5:   The data collection hover position is the vertical foot  $p_{r_i,ac}$  from  $c_i$  to the segment  $\overline{p_{a,t}c_j}$ .
- 6:   Calculate steering angle  $\Psi_{ac,t}$  by (23).
- 7: **else if**  $d_{seg_i} > d_{op}$  and  $m = op$  **then**
- 8:   Calculate the data collection hover position  $p_{r_i,op}$  by (22).
- 9:   Calculate steering angle  $\Psi_{op,t}$  by (23).
- 10: **else if**  $d_{seg_i} \leq d_{op}$  **then**
- 11:   The data collection hover position is the vertical foot  $p_{r_i,op}$  from  $c_i$  to the segment  $\overline{p_{a,t}c_j}$ .
- 12:   Calculate steering angle  $\Psi_{op,t}$  by (23).
- 13: **end if**

**Ensure:** The steering angle of the AUV:  $\Psi_{m,t}$ .

**5.3. DRL-Based Multi-Modal Path Planning Scheme**

Due to the uncertainty of reference access points and node data arrivals, the locations of AUV and the AoI of collected data are inherently random, which leads to a proliferation of state space dimensions. In comparison, DRL can handle extremely large state space by estimating the  $Q$  values of states  $s$  and actions  $a$  through neural networks [45,46]. The training framework of DQN includes a current  $Q$ -network and a target  $Q$ -network. In order to balance experience and exploration of the unknown, the agent at state  $s_t$  selects the action  $a_t$  to be performed by the  $\epsilon$ -greedy algorithm [47].

$$a_t = \begin{cases} \text{random } a \in \mathcal{A}, & \text{with probability } \epsilon \\ \arg \max_{a_t} Q(s_t, a_t; \theta), & \text{with probability } 1 - \epsilon \end{cases} \quad (24)$$

Immediately after adjusting the navigation angle and execution of action  $a_t$ , the AUV receives reward  $r_t$  and the data acquisition network moves to the next state  $s_{t+1}$ . Aiming to reduce the correlation between the online  $Q$ -network samples, an experience replay  $\mathcal{B}$  is used to store historical experience samples  $(s_t, a_t, r_t, s_{t+1})$ . At each training step, a small batch of randomly selected empirical samples  $\Phi_b$  from the experience replay is used to update the parameters of the online  $Q$ -network. In addition, we denote the parameters of DQN as  $\theta$ , and the parameters of the online  $Q$ -network are determined by minimizing the loss function.

$$L(\theta) = E_{\Phi_b} [\delta(s, a)^2], \quad (25)$$

where  $\delta(s, a) = y_t - Q(s_t, a_t | \theta)$  is the temporal difference, and  $y_t$  is the target  $Q$ -value, which can be calculated by

$$y_t = r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1} | \theta^-), \quad (26)$$

where  $\theta^-$  denoted the parameter of the target  $Q$ -network. Then, the weight of the current network  $\theta$  is updated by the following formula.

$$\theta = \theta + \frac{\alpha}{\Phi_b} \sum_{i=1}^{\Phi_b} \delta(s, a) \nabla_{\theta} Q(s_t, a_t, \theta). \quad (27)$$

The proposed AUV-assisted data collection algorithm is shown in Algorithm 2. The algorithm starts by initializing all neural networks as well as the replay buffer  $\mathcal{B}$ . The training iterates over  $E$  episodes, and the environment is initialized in each episode by observing the distribution of CHs. The action is first obtained according to the  $\epsilon$ -greedy policy, which is followed by inputting the action to Algorithm 1 to obtain the steering angle. Then, the AUV moves to the next state  $s_{t+1}$  and receives an immediate reward  $r_t$ . After storing the transition tuple  $(s_t, a_t, r_t, s_{t+1})$  in experience replay  $\mathcal{B}$ , a randomly selected sample of  $\Phi_b$  is utilized to learn the current network  $Q$ , and it updates the weights of the current network  $\theta$  and that of the target network  $\theta^-$ . Then,  $c_i$  is removed from  $N_r$  if the current state is able to collect the data of  $c_i$ , and the current loop is terminated when  $N_r = \emptyset$  holds.

---

**Algorithm 2** DRL-Based Multi-Modal Data Collection Algorithm
 

---

```

1: Input: Initialize the constants  $k_1, k_2$  and  $k_3$ , maximum number of training sets  $E$ , reward
   discount factor  $\gamma$ , learning rate  $l_r$ , experience replay  $\mathcal{B}$ , minimum batch  $\Phi_b$ , exploration
   probability  $\epsilon$ , and update step  $\chi$ ;
2: Initialize the current network  $Q(s_t, a_t, \theta)$  with weights  $\theta$  and the target network
    $Q(s_t, a_t, \theta^-)$  with weights  $\theta^-$ .
3: for  $episode = 1, \dots, E$  do
4:   for  $t = 1, \dots, \mathcal{T}$  do
5:     Initialize the data collection network environment and observe the initial
     state  $s_t$ .
6:     Select a random action  $a_t$  according to the  $\epsilon$ -greedy algorithm.
7:     Determine the AUV steering angle with Algorithm 1.
8:     Execute action  $a_t$  and observe the reward  $r_t$  and the next state  $s_{t+1}$ .
9:     Store experience  $(s_t, a_t, r_t, s_{t+1})$  in experience replay  $\mathcal{B}$ .
10:    Sample a random mini-batch of  $\Phi_b$  experiences from  $\mathcal{B}$ .
11:    Calculate the target value  $y_t$  by (26).
12:    Update the current network weights  $\theta$  by (27).
13:    Update the weights of the target network  $\theta^- = \theta$  every  $\chi$  steps.
14:    if  $s_{t+1}$  is the collection stop  $n_i$  then
15:      Remove the CH  $c_i$  from  $N_r$ .
16:    end if
17:    Terminate the episode if  $N_r = \emptyset$  holds.
18:  end for
19: end for
20: Output: The AUV trajectory  $p_{a,t}$  and the AoI  $A_i$ .
  
```

---

## 6. Results and Discussion

In this section, we conduct extensive simulations to verify the effectiveness of the proposed scheme. The simulation setup and numerical performance results are given as follows.

### 6.1. Simulation Setup

To evaluate the proposed scheme, we assume that there are 50 sensor nodes uniformly distributed in an  $800 \text{ m} \times 800 \text{ m}$  square target area. After CHs are designated, data fusion and data compression are performed by CHs. It is assumed that the data types collected and transmitted by the normal sensor nodes are text, records and images, and the amount of data pooled by the CHs is set to be between 10 and 300 packets, with the size of each packet 1024 bits. The AUV starts from the start point  $p_0 = (50, 120)$  with an initial orientation angle  $\psi_{m,0} = 0^\circ$  and returns to  $p_0$  after collecting data from all the CHs.

To evaluate the performance of the algorithm, a python 3.8 simulation environment was chosen. The target Q-network and the current Q-network are two-layer fully connected networks with 256 neurons per layer, and we use the ReLU function as the activation

function to train both networks using the Adam optimizer. Other simulation parameters and their specific values are provided in Table 1.

**Table 1.** Simulation parameters.

Parameters	Description	Value (Unit)
$f$	Carrier frequency	35 (kHz)
$\Delta f$	Bandwidth	2 (kHz)
$k$	Propagation loss	1.5
$SNR_{\min}^{ac}$	UAC minimum SNR	3 (dB)
$R_{ac}$	UAC data rate	3.16 (kbps)
$R_{op}$	UOC data rate	0.5 (Gbps)
$\theta$	Half of the transmitter beamwidth	3 ( $^{\circ}$ )
$c$	Extinction coefficient	0.18 ( $\text{m}^{-1}$ )
$\zeta$	Turbidity of water quality	0.05
$D_r$	Aperture diameter	0.25 (m)
$SNR_{\min}^{op}$	UOC minimum SNR	3 (dB)
$NEP$	Noise equivalent power	1 (mW)
$P_{rec}^{op}$	Average transmitted power	0.01 (mW)
$\rho$	Density of water	997 ( $\text{Kg}/\text{m}^3$ )
$\eta_p$	Efficiency of the AUV propulsion system	100%
$C_D$	Drag coefficient	0.0064
$A_s$	Wetted surface area	0.8856 ( $\text{m}^2$ )
$B$	Experience replay buffer size	500,000
$\Phi_b$	Mini-batch size	256
$\gamma$	Reward discount factor	0.95
$\chi$	Update step	1000

For the sake of performance comparison, the benchmark algorithms are provided as follows.

- **Single Acoustic:** The AUV exchanges data utilizing acoustic waves during data collection, and the hovering positions are determined by the UAC radius during the selection process of the steering angle. The AUV trajectories are learned using the DQN algorithm.
- **Single Optical:** The AUV can exchange data only by selecting optical waves and calculating the AUV hovering locations by means of the UOC radius. The DQN algorithm is used to learn the AUV trajectory.
- **Energy Greedy:** The AUV performs steering Algorithm 1 and then greedily selects the nodes with the shortest path length in the data collection sequence.

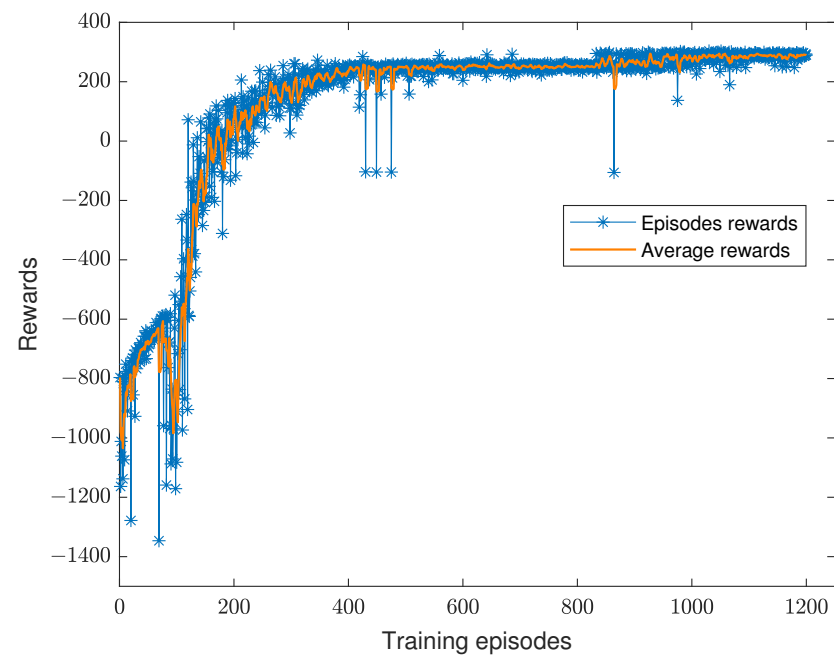
### 6.2. The Convergence Performance

To demonstrate the convergence of the AoI optimization algorithm for AUV data collection, in Figure 3, we show the variation of the cumulative reward, where the X-axis represents the number of iterations trained and the Y-axis presents the cumulative reward. It can be seen that in the early stages of training, the cumulative reward values are very low due to the high chance of  $\epsilon$ -greedy random exploration. As the training period continues to increase, the reward value gradually increases and stabilizes.

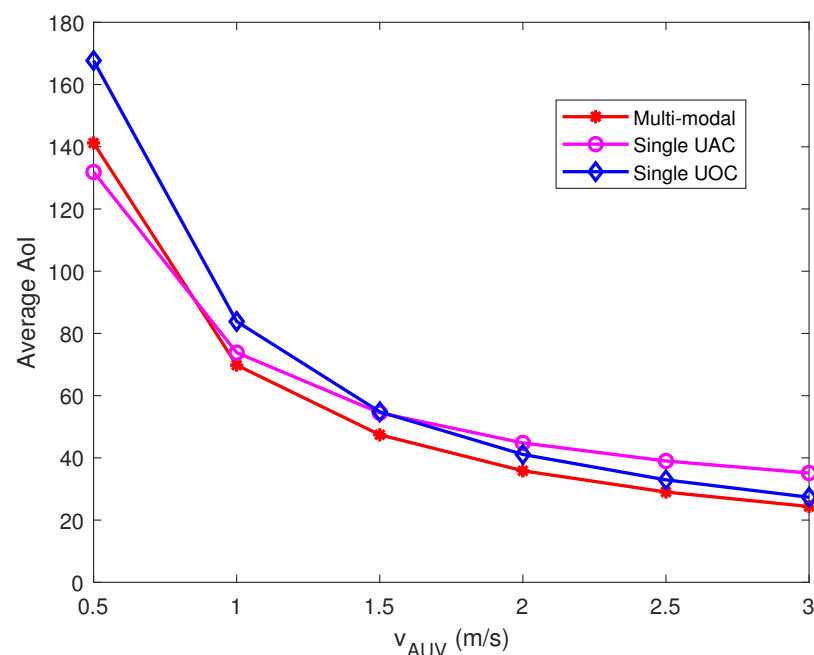
### 6.3. Impact of the AUV Velocity on Performance

To explore the effect of AUV velocity on the average AoI, we simulated the average AoI performance of the collected data with data arrival rate  $\lambda = 20$  and 300 Kbits. The experimental results are shown in Figure 4, where it can be observed that the average AoI of the data collected by AUV gradually decreases with increasing AUV velocity. The weighted average AoI of the single UAC is lower than that of the multi-modal and single UOC when the AUV speed is 0.5 m/s. With the increase of the AUV velocity,

the average AoI of the multi-modal data collection scheme is better than that of the single communication option.



**Figure 3.** Convergence performance of the DQN-based data collection algorithm.



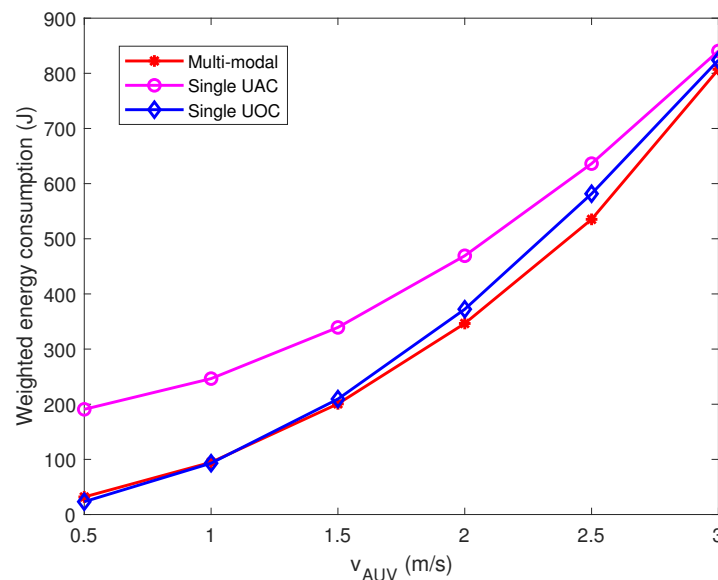
**Figure 4.** Effect of AUV velocity on AoI of collected data.

The primary reason for this performance is that the AUV travels slower and increases travel time, whereas the long-distance data collection via the single UAC is able to reduce the travel time of the AUV, which mitigates the increase in AoI. As the AUV velocity increases, the effect of AUV travel time on AoI is weakened, which makes the weight of data transmission time increase for AoI; thus, the UOC scheme outperforms the UAC scheme at higher AUV velocities. The multi-modal data collection scheme selects the best communication option according to the data characteristics, AUV navigation time and data transmission time so that the overall performance is better than the single-modal scheme.



Furthermore, the average AoI of our proposed multi-modal data collection scheme at the AUV velocity of 0.5 m/s is inferior to that of the single UAC scheme; this is because the multi-modal scheme not only considers AoI but also focuses on data collection energy consumption, so it sacrifices some AoI performance to reduce CHs energy consumption.

Under the parametric conditions of Figure 4, we analyzed the effect of AUV velocity on the energy consumption of data collection. Considering that the energy consumption of CHs is irreversible, we pay more attention to the energy consumption of data transmission, and therefore, we set  $\omega = 0.03$ . As shown in Figure 5, it can be observed that the weighted energy consumption of the single UAC scheme always remains at the highest level owing to the high weight of data transmission energy consumption. The single UOC scheme performs well in terms of CHs data transmission energy consumption, and hence, the weighted energy consumption is better than the single UAC scheme. The multi-modal scheme is able to reduce both AUV energy consumption and CHs data transmission energy consumption by jointly deciding on the best communication option based on packet size and path length. Furthermore, with the increasing AUV velocity, the weighted energy consumption of the three schemes will be convergent, since the AUV power increases geometrically with the velocity of travel.

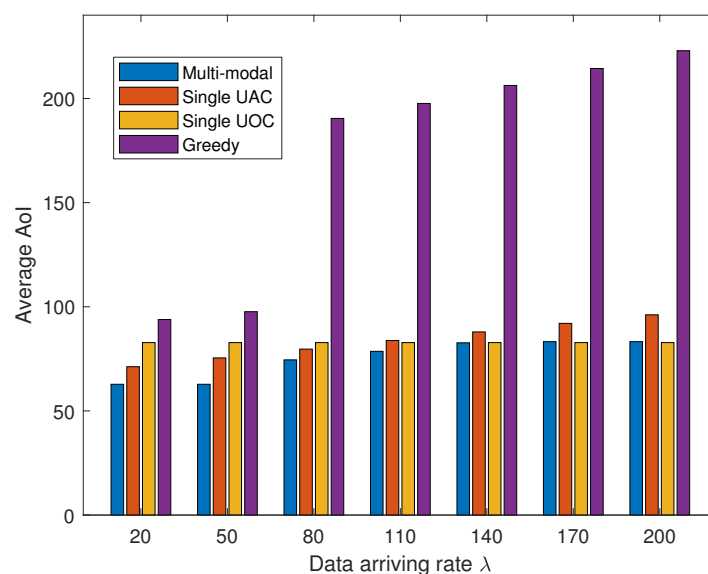


**Figure 5.** Effect of AUV velocity on the weighted energy consumption of the task.

#### 6.4. Impact of the Data Arrival Rate on Performance

Figure 6 shows the effect of different data arrival rates on the weighted average AoI. The velocity of the AUV is set to 1 m/s, and the length of each time slot is set to 6 s. It can be observed from the figure that at lower data arrival rates, the single UAC is superior to the single UOC since the acoustic waves can be deployed for long-range data transmission, which significantly saves the travel time of the AUV. As the data arrival rate increases, the weight of data transmission time on AoI improves, which results in the single UOC scheme being superior to the single UAC scheme in terms of AoI. In addition, the greedy algorithm performs poorly in weighted average AoI as it greedily selects the closest visit location ignoring data importance and AoI. Our proposed multi-modal data collection scheme outperforms the other three schemes for different data arrival rates and is near the single UOC performance when the data size is over 140 Kbits. The main reason for this phenomenon is that the multi-modal scheme selects acoustic communication to reduce the sailing time when the data size is small and optical communication for fast data transmission when the data size is large, and thus, it can adapt to different data conditions and achieve a relatively low weighted average AoI.

To verify the superiority of the proposed multi-modal data collection algorithm in terms of energy consumption for data collection, we compare the transmission energy consumption of CHs and AUV energy consumption for different data sizes. In this study, we set the AUV velocity to 1 m/s and the data size to 20–200 Kbits, and the experimental results are shown in Table 2 and Figure 7. It is observed that the average data transmission energy consumption of CHs under a single UOC approach is the smallest, and the AUV energy consumption is the highest. The energy consumption of CHs is the highest for the single UAC and greedy approaches, but the AUV energy consumption is kept at a low level. In the multi-modal scheme, the energy consumption of CH increases and then decreases with the increasing data size, and the AUV energy consumption gradually increases, but the overall energy consumption remains at a low level.



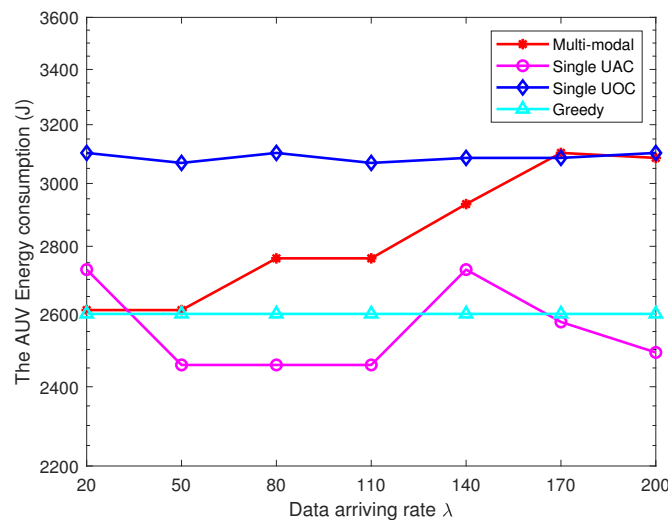
**Figure 6.** The average AoI of collected data with the increasing data arriving rate.

**Table 2.** Average data transmission energy consumption of CHs versus the data arriving rate.

Collected Data (Kbits)	Single UAC (mJ)	Single UOC (mJ)	Greedy (mJ)	Multi-Modal (mJ)
20	6659.29	0.02	5650.31	5112.19
50	15,403.83	0.05	12,780.47	3968.71
80	24,081.09	0.08	20,852.35	5583.11
110	32,623.83	0.11	28,789.69	7668.36
140	41,099.29	0.14	379,37.81	0.14
170	49,103.91	0.16	46,346.02	0.16
200	52,399.92	0.19	55,359.61	0.19

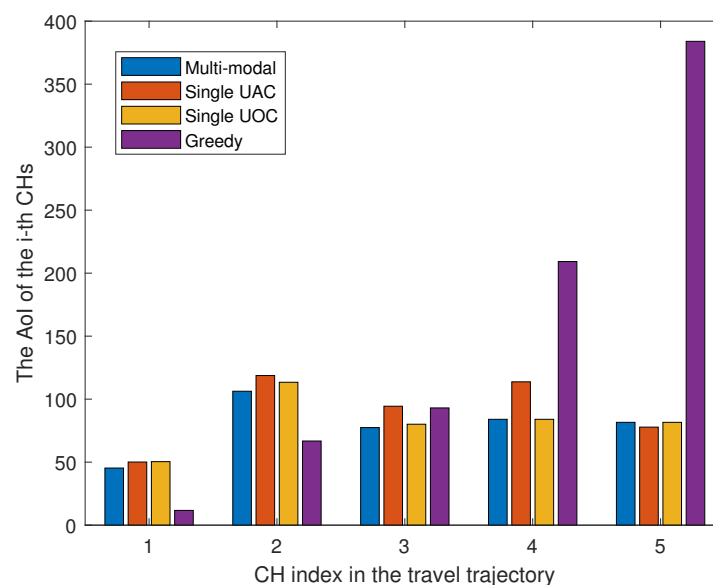
The reason for such a phenomenon can be explained as follows. The single UOC requires the AUV to travel to the immediate vicinity of the node, which increases the energy consumption of the AUV for navigation. Fortunately, due to the low energy consumption and high bandwidth of the optical modems, the energy consumption of the CHs owing to the data transmission is low. Similarly, the single UAC and greedy algorithm allow the AUV to collect data over longer distances using acoustic waves, which greatly saves AUV energy consumption. However, with the increasing data size, the low bandwidth and high energy consumption of the UAC make the energy consumption of the CHs significant. In the multi-modal scheme, the AUV selects the UAC to collect data when the data size is small, keeping the transmission energy consumption of the CHs low while reducing the mobile energy consumption. When the data size is larger than 140 Kbits, the multi-modal scheme switches to strictly optical communication mode to reduce the excessive energy consumption of the CHs in order to extend the lifetime of the UWSN. Note that our

proposed multi-modal data collection scheme has an excellent performance in the face of diversified data, and when the data size of each CHs is large (or small), the multi-modal scheme will become a strict UOC (or UAC) scheme.



**Figure 7.** Performance comparison in terms of AUV energy consumption.

In Figure 8, we show the weights of AoI for each CHs under different schemes. The results show that the greedy scheme has the maximum AoI value for CH index = 5 and the lowest AoI value for index = 1. This is because the greedy algorithm ignores the effect of data importance when selecting the nearest nodes to visit, resulting in a large data AoI for the first visited node. The other three schemes use reinforcement learning methods to select the best node access order based on the importance of the data, which avoids the extreme cases of AoI values. Furthermore, the multi-modal scheme flexibly selects the communication options based on the data size and importance of the nodes, and hence, its performance is better compared to the other two single-modal schemes. It is worth noting that we neglected the specific details of light alignment and the time consumed during data collection, which result in a seemingly promising AoI performance for the single UOC. In future work, we will consider more details of underwater optical communication.



**Figure 8.** AoI for different paths when the number of CHs is 5.

## 7. Conclusions

In this paper, we proposed an AUV-assisted multi-modal data collection scheme which provides timely and reliable data collection by utilizing underwater acoustic and optical communication technologies in an adaptive manner. The trajectory planning problem is formulated as a mixed integer nonlinear problem to minimize the weighted average AoI and energy consumption, and the data collection problem is formulated as an MDP considering data importance, packet size, and data collection options. We then developed a DQN-based learning algorithm to determine the optimal strategy. In addition, an AUV multi-modal corner optimization algorithm is proposed to reduce the energy consumption of AUV navigation. Through numerical simulations, we showed that our proposed algorithm has convergence capability as well as verified that the AUV path-planning algorithm has excellent performance which can effectively reduce the AoI and energy consumption of collected data.

**Author Contributions:** F.B. proposed the main ideas, wrote the paper, designed the description framework, and conducted the simulations. H.L. provided guidance for the work, discussed and provided ideas, wrote and modified the paper and acquired funding. X.L., R.R. and G.H. provided guidance for the work, and collaborated in discussion on the proposed system model and techniques. S.M. assisted in testing the code and checked the paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the Shandong Provincial Natural Science Foundation, China, under Grant ZR2020MF059; in part by the China NSFC under Grant 62072287.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank the anonymous reviewers for their time towards the manuscript.

**Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Jahanbakht, M.; Xiang, W.; Hanzo, L.; Azghadi, M.R. Internet of underwater things and big marine data analytics—A comprehensive survey. *IEEE Commun. Surveys Tuts.* **2021**, *23*, 904–956. [\[CrossRef\]](#)
2. Wei, X.; Guo, H.; Wang, X.; Wang, X.; Qiu, M. Reliable data collection techniques in underwater wireless sensor networks: A survey. *IEEE Commun. Surveys Tuts.* **2021**, *24*, 404–431. [\[CrossRef\]](#)
3. Luo, H.; Wu, K.; Ruby, R.; Liang, Y.; Guo, Z.; Ni, L.M. Software-defined architectures and technologies for underwater wireless sensor networks: A survey. *IEEE Commun. Surveys Tuts.* **2018**, *20*, 2855–2888. [\[CrossRef\]](#)
4. Luo, H.; Wang, J.; Bu, F.; Ruby, R.; Wu, K.; Guo, Z. Recent progress of air/water cross-boundary communications for underwater sensor networks: A review. *IEEE Sens. J.* **2022**, *22*, 8360–8382. [\[CrossRef\]](#)
5. Fang, Z.; Wang, J.; Jiang, C.; Zhang, Q.; Ren, Y. AoI-inspired collaborative information collection for AUV-assisted internet of underwater things. *IEEE Internet Things J.* **2021**, *8*, 14559–14571. [\[CrossRef\]](#)
6. Jawhar, I.; Mohamed, N.; Al-Jaroodi, J.; Zhang, S. An architecture for using autonomous underwater vehicles in wireless sensor networks for underwater pipeline monitoring. *IEEE Trans. Ind. Informat.* **2018**, *15*, 1329–1340. [\[CrossRef\]](#)
7. Mahmoodi, K.A.; Uysal, M. AUV Trajectory Optimization for an Optical Underwater Sensor Network in the Presence of Ocean Currents. In Proceedings of the 2021 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), Bucharest, Romania, 24–28 May 2021; pp. 1–6.
8. Wang, X.; Luo, H.; Yang, Y.; Ruby, R.; Wu, K. Underwater Real-time Video Transmission via Optical Channels with Swarms of AUVs. In Proceedings of the 2021 IEEE 27th International Conference on Parallel and Distributed Systems (ICPADS), Beijing, China, 14–16 December 2021; pp. 859–866.
9. Wang, J.; Li, J.; Yan, S.; Shi, W.; Yang, X.; Guo, Y.; Gulliver, T.A. A novel underwater acoustic signal denoising algorithm for Gaussian/non-Gaussian impulsive noise. *IEEE Trans. Veh. Technol.* **2020**, *70*, 429–445. [\[CrossRef\]](#)
10. Akhoundi, F.; Jamali, M.V.; Hassan, N.B.; Beyranvand, H.; Minoofar, A.; Salehi, J.A. Cellular underwater wireless optical CDMA network: Potentials and challenges. *IEEE Access* **2016**, *4*, 4254–4268. [\[CrossRef\]](#)

11. Wang, J.; Luo, H.; Ruby, R.; Liu, J.; Guo, K.; Wu, K. Reliable Water-Air Direct Wireless Communication: Kalman Filter-Assisted Deep Reinforcement Learning Approach. In Proceedings of the 2022 IEEE 47th Conference on Local Computer Networks (LCN), Edmonton, AB, Canada, 26–29 September 2022; pp. 233–238.
12. Luo, H.; Xie, X.; Han, G.; Ruby, R.; Hong, F.; Liang, Y. Multimodal acoustic-RF adaptive routing protocols for underwater wireless sensor networks. *IEEE Access* **2019**, *7*, 134954–134967. [[CrossRef](#)]
13. Zhao, Z.; Liu, C.; Qu, W.; Yu, T. An Energy Efficiency Multi-Level Transmission Strategy based on underwater multimodal communication in UWSNs. In Proceedings of the IEEE INFOCOM 2020—IEEE Conference on Computer Communications, Toronto, ON, Canada, 6–9 July 2020; pp. 1579–1587.
14. Gauni, S.; Manimegalai, C.; Krishnan, K.M.; Shreeram, V.; Arvind, V.; Srinivas, T. Design and analysis of co-operative acoustic and optical hybrid communication for underwater communication. *Wireless Pers. Commun.* **2021**, *117*, 561–575. [[CrossRef](#)]
15. Yu, T.; Liu, C.; Qu, W.; Zhao, Z. OD-PPS: An On-Demand Path Planning Scheme for Maximizing Data Completeness in Multimodal UWSNs. In Proceedings of the International Conference on Wireless Algorithms, Systems, and Applications, Nanjing, China, 25–27 June 2021; pp. 16–28.
16. Gjanci, P.; Petrioli, C.; Basagni, S.; Phillips, C.A.; Bölöni, L.; Turgut, D. Path finding for maximum value of information in multi-modal underwater wireless sensor networks. *IEEE Trans. Mobile Comput.* **2018**, *17*, 404–418. [[CrossRef](#)]
17. Hu, Y.; Zheng, Y.; Liu, H.; Wang, Z.; Mao, Y.; Han, H. Mobile sink path planning research for underwater heterogeneous sensor network. In Proceedings of the 2018 Chinese Control And Decision Conference (CCDC), Shenyang, China, 9–11 June 2018; pp. 4443–4448.
18. Han, S.; Noh, Y.; Liang, R.; Chen, R.; Cheng, Y.J.; Gerla, M. Evaluation of underwater optical-acoustic hybrid network. *China Commun.* **2014**, *11*, 49–59.
19. Sun, Y.; Uysal-Biyikoglu, E.; Yates, R.D.; Koksall, C.E.; Shroff, N.B. Update or wait: How to keep your data fresh. *IEEE Trans. Inf. Theory* **2017**, *63*, 7492–7508. [[CrossRef](#)]
20. Talak, R.; Karaman, S.; Modiano, E. Optimizing information freshness in wireless networks under general interference constraints. In Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing, Los Angeles, CA, USA, 26–29 June 2018; pp. 61–70.
21. Yi, M.; Wang, X.; Liu, J.; Zhang, Y.; Bai, B. Deep reinforcement learning for fresh data collection in UAV-assisted IoT networks. In Proceedings of the IEEE INFOCOM 2020—IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Toronto, ON, Canada, 6–9 July 2020; pp. 716–721.
22. Basagni, S.; Di Valerio, V.; Gjanci, P.; Petrioli, C. MARLIN-Q: Multi-modal communications for reliable and low-latency underwater data delivery. *Ad Hoc Netw.* **2019**, *82*, 134–145. [[CrossRef](#)]
23. Shen, Z.; Yin, H.; Jing, L.; Liang, Y.; Wang, J. A Cooperative Routing Protocol Based on Q-Learning for Underwater Optical-Acoustic Hybrid Wireless Sensor Networks. *IEEE Sens. J.* **2021**, *22*, 1041–1050. [[CrossRef](#)]
24. Júnior, E.P.C.; Vieira, L.F.; Vieira, M.A. CAPTAIN: A data collection algorithm for underwater optical-acoustic sensor networks. *Comput. Netw.* **2020**, *171*, 1–11.
25. Luo, H.; Xu, Z.; Wang, J.; Yang, Y.; Ruby, R.; Wu, K. Reinforcement Learning-Based Adaptive Switching Scheme for Hybrid Optical-Acoustic AUV Mobile Network. *Wirel. Commun. Mob. Com.* **2022**, *2022*, 1–14. [[CrossRef](#)]
26. Liu, Z.; Meng, X.; Liu, Y.; Yang, Y.; Wang, Y. AUV-Aided Hybrid Data Collection Scheme Based on Value of Information for Internet of Underwater Things. *IEEE Internet Things J.* **2021**, *9*, 6944–6955. [[CrossRef](#)]
27. Duan, R.; Du, J.; Jiang, C.; Ren, Y. Value-based hierarchical information collection for AUV-enabled internet of underwater things. *IEEE Internet Things J.* **2020**, *7*, 9870–9883. [[CrossRef](#)]
28. Fang, Z.; Wang, J.; Jiang, C.; Wang, X.; Ren, Y. Average Peak Age of Information in Underwater Information Collection with Sleep-scheduling. *IEEE Trans. Veh. Technol.* **2022**, *71*, 10132–10136. [[CrossRef](#)]
29. Khan, M.T.R.; Jembre, Y.Z.; Ahmed, S.H.; Seo, J.; Kim, D. Data freshness based AUV path planning for UWSN in the internet of underwater things. In Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 9–13 December 2019; pp. 1–6.
30. Al-Habob, A.A.; Dobre, O.A.; Poor, H.V. Age-optimal information gathering in linear underwater networks: A deep reinforcement learning approach. *IEEE Trans. Veh. Technol.* **2021**, *70*, 13129–13138. [[CrossRef](#)]
31. Wu, T.; Wen, P.; Tang, S. Optimal scheduling strategy of AUV based on importance and age of information. *Wirel. Netw.* **2022**, 1–9. [[CrossRef](#)]
32. Huang, M.; Zhang, K.; Zeng, Z.; Wang, T.; Liu, Y. An AUV-assisted data gathering scheme based on clustering and matrix completion for smart ocean. *IEEE Internet Things J.* **2020**, *7*, 9904–9918. [[CrossRef](#)]
33. Cheng, W.; Teymorian, A.Y.; Ma, L.; Cheng, X.; Lu, X.; Lu, Z. Underwater localization in sparse 3D acoustic sensor networks. In Proceedings of the IEEE INFOCOM 2008—The 27th Conference on Computer Communications, Phoenix, AZ, USA, 13–18 April 2008; pp. 236–240.
34. Stojanovic, M.; Preisig, J. Underwater acoustic communication channels: Propagation models and statistical characterization. *IEEE Commun. Mag.* **2009**, *47*, 84–89. [[CrossRef](#)]
35. Brekhovskikh, L.M.; Lysanov, Y.P.; Beyer, R.T. *Fundamentals of Ocean Acoustics*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2003.

36. Stojanovic, M. On the relationship between capacity and distance in an underwater acoustic communication channel. In Proceedings of the ACM SIGMOBILE Mobile Computing and Communications Review, Los Angeles, CA, USA, 25 September 2007; ACM: Los Angeles, CA, USA, 2007; Volume 11, pp. 34–43.
37. Elamassie, M.; Miramirkhani, F.; Uysal, M. Performance characterization of underwater visible light communication. *IEEE Trans. Commun.* **2018**, *67*, 543–552. [[CrossRef](#)]
38. Giles, J.W.; Bankman, I.N. Underwater optical communications systems. Part 2: Basic design considerations. In Proceedings of the MILCOM 2005–2005 IEEE Military Communications Conference, Atlantic City, NJ, USA, 17–20 October 2005; pp. 1700–1705.
39. Steinvall, O. Laser system range calculations and the Lambert W function. *Appl. Opt.* **2009**, *48*, B1–B7. [[CrossRef](#)]
40. Farr, N.; Chave, A.; Freitag, L.; Preisig, J.; White, S.; Yoerger, D.; Titterton, P. Optical modem technology for seafloor observatories. In Proceedings of the Proceedings of OCEANS 2005 MTS/IEEE, Washington, DC, USA, 17–23 September 2005; pp. 928–934.
41. Liu, J.; Wang, X.; Bai, B.; Dai, H. Age-optimal trajectory planning for UAV-assisted data collection. In Proceedings of the IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Honolulu, HI, USA, 15–19 April 2018; pp. 553–558.
42. Phillips, A.; Haroutunian, M.; Murphy, A.J.; Boyd, S.; Blake, J.; Griffiths, G. Understanding the power requirements of autonomous underwater systems, Part I: An analytical model for optimum swimming speeds and cost of transport. *Ocean Eng.* **2017**, *133*, 271–279. [[CrossRef](#)]
43. Furlong, M.E.; McPhail, S.D.; Stevenson, P. A concept design for an ultra-long-range survey class AUV. In Proceedings of the OCEANS 2007-Europe, Aberdeen, UK, 18–21 June 2007; pp. 1–6.
44. Rigby, P.; Pizarro, O.; Williams, S.B. Towards geo-referenced AUV navigation through fusion of USBL and DVL measurements. In Proceedings of the OCEANS 2006, Boston, MA, USA, 18–21 September 2006; pp. 1–6.
45. Su, N.; Wang, J.B.; Zeng, C.; Zhang, H.; Lin, M.; Li, G.Y. Unmanned-Surface-Vehicle-Aided Maritime Data Collection Using Deep Reinforcement Learning. *IEEE Internet Things J.* **2022**, *9*, 19773–19786. [[CrossRef](#)]
46. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G.; et al. Human-level control through deep reinforcement learning. *Nature* **2015**, *518*, 529–533. [[CrossRef](#)]
47. dos Santos Mignon, A.; da Rocha, R.L.d.A. An adaptive implementation of  $\epsilon$ -greedy in reinforcement learning. *Procedia Comput. Sci.* **2017**, *109*, 1146–1151. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.