

单水听器强化学习无源自导方法

尹 涵 陈建峰[†] 张董哲

(西北工业大学 西安 710072)

2023 年 8 月 4 日收到

2023 年 11 月 25 日定稿

摘要 为降低水下无人自主航行器在搭载自导系统时对体积与成本的苛刻要求,提出了一种简便可行的无源自导方法,可仅使用单个声压水听器实现航行器对中心频率已知窄带声源目标的导引。该方法有效利用了航行器自身与目标间相对运动所产生的多普勒频移的变化,对航行器的追踪行为进行控制。基于强化学习 Actor-Critic 算法,利用实时获取的接收信号中心频率和信噪比参数构建价值函数,利用该函数输出的目标检测结果、目标方位角和中心频率变化率参数构建策略函数,结合历史动作,输出当前最优动作。通过仿真和湖试实验研究了该方法的训练效率以及目标跟踪性能。结果表明,所提无源自导方法可以方便地应用于水下无人平台,导引航行器对不同速度的声源目标进行较为精确的跟踪。

关键词 单水听器,无源自导,强化学习,多普勒频移

PACS: 43.30, 43.90

DOI: 10.12395/0371-0025.2023158

CSTR: 32049.14.11-2065.2023158

Passive homing method with reinforcement learning for a single hydrophone

YIN Han CHEN Jianfeng[†] ZHANG Dongzhe

(Northwestern Polytechnical University Xi'an 710072)

Received Aug. 4, 2023

Revised Nov. 25, 2023

Abstract In order to reduce the stringent volume and cost requirements for underwater unmanned autonomous vehicles (AUVs) when equipped with homing systems, this paper proposes a simple and feasible passive homing method. The proposed method can use only one sound pressure hydrophone to guide AUVs to approach the narrow-band sound source target with a known center frequency. The change in Doppler frequency shift caused by the relative motion between the AUV and target is used to control tracking behaviors. Based on the reinforcement learning Actor-Critic algorithm, the value function is constructed using the center frequency and signal-to-noise ratio parameters of the received signal obtained in real time. The target detection results, the target azimuth angle and the center frequency change rate parameters generated by the value function are utilized to construct the strategy function. The strategy function can combine historical actions to output the current optimal action. The training efficiency and the target tracking performance of the proposed method are studied through simulation experiments and lake trials. Results show that our passive homing method can be easily applied to underwater unmanned platforms to guide AUVs to more accurately track sound source targets with different speeds.

Keywords Single hydrophone, Passive homing, Reinforcement learning, Doppler frequency

引言

海洋中常用的各类有源声呐(如成像声呐、通信声呐)及水下声学信标等,其发射信号往往具有窄带

的特性^[1,2],在一些情况下,信号的中心频率可以提前获知。例如飞机黑匣子(飞行数据记录器和驾驶舱语音记录器)的发射信号中心频率通常是 37.5 kHz,这一特定频率已成为国际标准,以便搜寻人员在飞机失事后更容易对其定位和回收。如何有效利用水

[†] 通讯作者: 陈建峰, chenjf@nwpu.edu.cn

声信息引导自主水下航行器 (AUV) 对水下中心频率已知或可获得的窄带声源目标进行逼近, 是一个持续受到广泛关注的问题。

传统的无源自导方法主要基于多水听器阵列, 利用目标方向估计结果进行导引^[3], 这对 AUV 的体积与系统研发成本的要求较高。近年来, 小尺寸是 AUV 的发展趋势之一^[4]。与传统中大型 AUV 相比, 微小型 AUV 具有更灵敏和更隐蔽的优势, 适合批量生产和集群作业^[5]。相比多水听器阵列, 单水听器载荷对 AUV 体积要求更低, 同时可有效降低成本。

近年来, 研究人员先后提出了一些利用单个水听器实现无源定位的方法^[6-9], 从水听器的类别可分为矢量水听器和声压水听器两种实现方案。其中, 矢量水听器具有多个通道, 可提供声源方位信息, 但其体积相对较大。相比之下, 声压水听器通常为单通道, 体积较小, 结构更简单, 成本也更低廉。而在算法层面, 则需要在单一声压传感器条件下, 通过一定的信号处理手段来实现目标定向。

早在 1986 年, James 就提出基于多普勒频移信息来估计 AUV 与目标的相对运动趋势^[10]。当两者相互靠近时, AUV 接收信号的中心频率会增大, 反之亦然。2019 年, Bai 等介绍了单通道多普勒测向系统的组成、信号处理方法和测向原理^[11], 但在实际应用时, 单声压水听器会出现“左右模糊”的问题。本文综合了以上方法, 利用多普勒测向结果调节 AUV 的动作的大小, 而动作方向则由相对运动趋势进行控制: 当接收信号中心频率下降时, 表明 AUV 正在远离目标, 此时可以改变 AUV 的方向, 使其逐渐逼近目标。这种通过环境反馈不断“试错”从而对自身动作进行优化调整的策略被称为强化学习方法^[12]。

强化学习作为一种可根据环境反馈逐步提升控制策略的方法, 已经被广泛应用于 AUV 的控制问题研究, 主要分为两种范式: 基于价值的方法和基于策略的方法^[13-15]。在基于价值的方法中, 价值函数的估计方差较小, 但较难处理连续动作空间问题。而基于策略的方法直接对策略进行优化, 适用于连续动作空间, 但因为直接在策略空间中搜索, 所以更难以训练和收敛。除以上两种方法外, 目前更流行的是两者的结合, 即演员-评论家 (Actor-Critic) 方法^[16,17]。在该方法中, 价值函数通常位于 Critic 中, 而策略函数位于 Actor 中。这两个组件共同协作, Critic 提供了关于状态价值的反馈, Actor 则利用反馈来更新策略。这种协作使得 Actor-Critic 方法在处理连续动作空间和复杂环境中表现出色, 同时具有更好的样本效率和实时性, 有助于提高 AUV 在不确

定因素较多的水下环境中的鲁棒性。

本文主要研究如何利用单个声压水听器对中心频率已知的窄带声源目标进行无源式声自导, 可有效降低自导系统对 AUV 体积和研发成本的要求。针对单声压水听器定向“左右模糊”的问题, 提出利用航行器与目标相对运动产生的多普勒频移变化对 AUV 的运动方向进行控制。基于强化学习 Actor-Critic 方法, 利用接收信号参数构建价值函数, 利用价值函数的输出参数设计策略函数的决策过程。该方法可有效地结合 AUV 历史动作, 根据环境反馈, 输出当前最优动作, 对静止或运动声源目标进行逼近。在丹江口水库使用 TH-B010 型 AUV 进行了多次湖试实验, 验证了所提方法的有效性。

1 模型与方法

1.1 目标方位与多普勒频移

考虑水下无源跟踪场景, AUV 作为接收体, 声源目标作为信源。当目标与 AUV 之间存在相对运动时, AUV 接收到的来自目标发射信号的频率与发射信号的中心频率不相同, 此时接收频率与发射信号中心频率之差称为多普勒频移。

如果 AUV 和目标深度差异较小, 则可被视为在同一个二维平面中的两个不同质点。如图 1 所示, 其中 S 代表 AUV, K 代表声源目标。 v_s 与 v_k 分别代表 AUV 与目标的速度矢量。目标方位角 θ 为目标相对于 AUV 速度方向的夹角, φ 为 AUV 相对于目标速度方向的夹角。

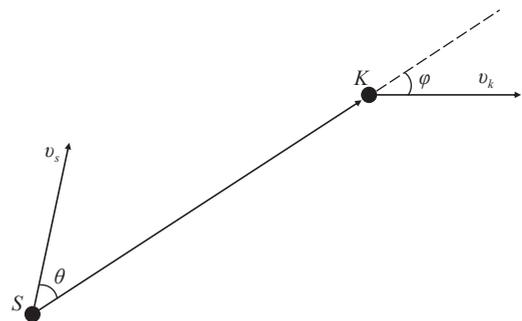


图 1 AUV 与目标相对运动示意图

由多普勒频移原理可得目标方位角的计算表达式为:

$$\theta = \arccos \left[\frac{f}{|v_s| f_0} (c + |v_k| \cos \varphi) - \frac{c}{|v_s|} \right], \quad (1)$$

其中, f_0 为信源目标发射信号的中心频率, f 为 AUV 端接收信号的中心频率, c 为水中声速。可以看到, 通过多普勒频移仅能确定目标方位角大小, 而

无法分清左右。同时,目标速度未知,所以不能直接通过式(1)对目标方位角进行估计。本文利用 sigmoid 函数^[18],得到目标方位角的估计表达式:

$$\hat{\theta} = \arccos[\text{est}(w \cdot f + b)], \quad (2)$$

$$\text{est}(x) = \begin{cases} x, & -1 < x < 1, \\ 2 \cdot \text{sigmoid}(x) - 1, & \text{其他}, \end{cases} \quad (3)$$

其中,权值 $w = c \cdot (|v_s| f_0)^{-1}$, 偏置 $b = -c \cdot |v_s|^{-1}$, $\text{sigmoid}(x) = (1 + e^{-x})^{-1}$ 。当目标方位角减小时,意味着 AUV 在逐渐靠近目标。为了探究目标方位角大小与接收信号中心频率的关系,对式(1)求导,可得

$$\frac{\partial \theta}{\partial f} = \frac{-1}{\sqrt{1-k^2}} \left[\frac{c + |v_k| \cos \varphi - f |v_k| \sin \varphi (\partial \varphi / \partial f)}{|v_s| f_0} \right], \quad (4)$$

其中, $k = f(c + |v_k| \cos \varphi) (|v_s| f_0)^{-1} - c |v_s|^{-1}$ 。

1.2 单水听器无源自导方法

由式(4)可知,在 AUV 对目标进行跟踪的过程中,当目标速度满足 $c + |v_k| \cos \varphi - f |v_k| \sin \varphi (\partial \varphi / \partial f) > 0$ 时, θ 对 f 的导数为负,表示接收信号中心频率会随着目标方位角的减小而不断增大。当 AUV 偏离目标航行时,目标方位角增大,接收信号中心频率会逐渐减小;当 AUV 靠近目标航行时,目标方位角减小,中心频率会逐渐增大;当 AUV 正对目标航行时,目标方位角为零,中心频率最大。

因此,如能保证 AUV 在跟踪目标的过程中接收信号中心频率最大化,当中心频率下降时进行纠正,便可使 AUV 不断靠近目标。基于强化学习 Actor-Critic 算法,利用声多普勒信息,设计合理的价值函数和策略函数,可以有效地实现这一自动纠错机制。图 2(a) 展示了一般的 Actor-Critic 算法框图,本文进行了简单改动,不再将状态变量输入到策略函数,并将历史动作输入策略函数。本文中 AUV 的动作定义为舵角(以 AUV 轴向为 0° ,取值范围为 -30° 至 30°)。奖励与导引误差成反比,导引误差越

小,奖励越大。

如图 3 所示,自导系统主要由两部分组成:信号处理单元和主控单元。其中,信号处理单元以固定的快拍时长不断地对接收到的水声信号进行频谱分析,计算中心频率,并计算接收信号在中心频率点处的信噪比。需要注意的是,所使用的频谱分析算法应当尽可能提高频谱的频率分辨率。相比于常规的快速傅里叶变换 (FFT) 算法,细化快速傅里叶变换 (Zoom-FFT) 算法可在不改变采样率和采样点数的情况下,获得比常规 FFT 算法更高的频率分辨率^[19,20]。

主控单元以固定的周期和信号处理单元建立一次通信,获取接收信号中心频率和信噪比,此通信周期在文中被统称为“自导采样周期”。主控单元中的策略函数连续输出两次最优动作的时间间隔在文中被统称为“自导控制周期”。自导采样周期总是小于等于自导控制周期。在每个自导控制周期内,信号处理单元会和主控单元建立多次通信,传输多次数据并进行累积,所累积的数据序列作为传输给价值函数的状态变量。

1.2.1 环境

强化学习算法需要提前设定智能体所处的环境。在该方法中,设定单水听器 AUV 与窄带声源目标在水中同一深度运动。水中声速为 c ,目标信号中心频率为 f_0 。

1.2.2 状态变量

在强化学习算法中,需要定义传递给价值函数或策略函数的一系列状态变量。同样,在此给出状态变量的计算方式。如图 3 所示,主控单元会以自导采样周期不断获取信号处理单元输出的接收信号中心频率 f_s 及信噪比 r_{sn} 。同时,会每隔一个自导控制周期进行一次动作修正。自导控制周期大于自导采样周期。传递给价值函数的状态变量为当前自导控制周期内累积的信噪比数据序列 R_{sn} 和中心频率数据序列 F_c :

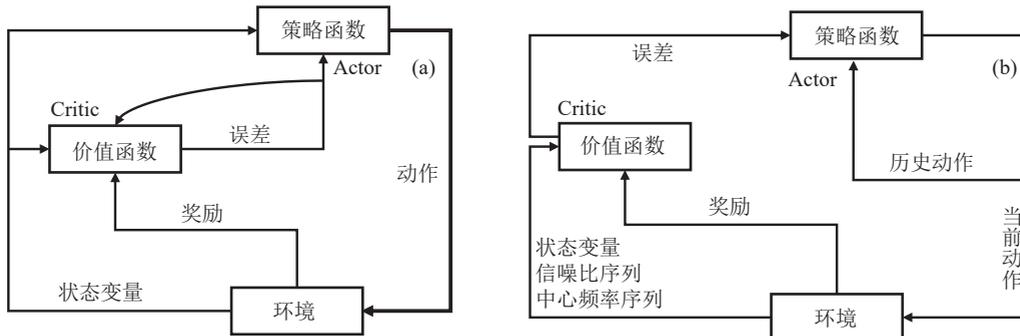


图 2 算法框图 (a) 强化学习 Actor-Critic 算法; (b) 所提单水听器无源自导方法

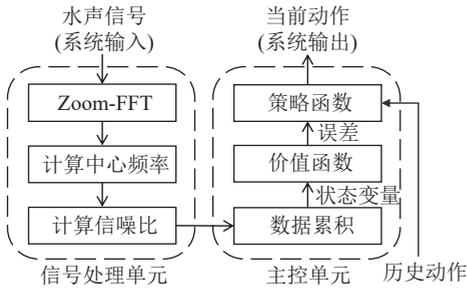


图3 所提基于强化学习的单水听器无源自导方法的信息处理流程框图

$$\mathbf{R}_{\text{sn}} = [r_{\text{sn}}(1), r_{\text{sn}}(2), \dots, r_{\text{sn}}(N)], \quad (5)$$

$$\mathbf{F}_c = [f_c(1), f_c(2), \dots, f_c(N)]. \quad (6)$$

1.2.3 价值函数

价值函数是 Actor-Critic 算法中 Critic 的核心, 该函数反映了历史动作的影响。价值函数的输入为当前环境所反馈的状态变量, 输出的误差决定了 Actor 所产生的动作。在 1.2.2 节已对状态变量进行了定义, 在此只需要对价值函数的输出进行定义。在本文方法中, 误差由三部分组成, 分别记为 E_1 、 E_2 和 E_3 。

(1) 目标检测结果 E_1

E_1 主要是用来进行目标检测, 可由以下公式计算:

$$E_1 = \begin{cases} 1, & N_1/N > \sigma, \\ 0, & N_1/N \leq \sigma, \end{cases} \quad (7)$$

其中, N_1 为信噪比数据序列 \mathbf{R}_{sn} 大于等于 ϖ 的频数; ϖ 为信噪比阈值, 当且仅当接收信号的信噪比大于等于 ϖ 时认为接收到有效的目标信号; σ 为检测率阈值, 当且仅当有效信号占比 N_1/N 大于 σ 时认为目标存在。

(2) 目标方位角 E_2

E_2 是当前的目标方位角估计值, 影响策略函数输出的舵角大小。需要将 \mathbf{F}_c 中对应信噪比大于 ϖ 的频点取出构成新的中心频率序列 $\mathbf{F}'_c = [f'_c(1), f'_c(2), \dots, f'_c(N_1)]$, 并计算该序列的平均值:

$$\bar{f} = \frac{1}{N_1} \sum_{n=1}^{N_1} f'_c(n). \quad (8)$$

将 \bar{f} 代入式 (2) 可得 E_2 。

(3) 中心频率平均变化率 E_3

E_3 主要用来估计 AUV 的当前航向是否正确, 影响策略函数输出的舵角方向。利用移动平均法计算 \mathbf{F}'_c 的平均变化率作为 E_3 , 计算过程为

$$f''(i) = \frac{1}{M} \sum_{i=j}^{j+M-1} f'_c(i), \quad (9)$$

$$\nabla f''_c(l) = \frac{f''_c(l+1) - f''_c(l)}{T_c}, \quad (10)$$

$$E_3 = \frac{1}{N_1 - M} \sum_{m=1}^{N_1 - M} \nabla f''_c(m), \quad (11)$$

其中, $j = 1, 2, \dots, N_1 + 1 - M$, $l = 1, 2, \dots, N_1 - M$, M 为窗函数长度, T_c 为自导采样周期。

1.2.4 策略函数

策略函数是强化学习中最重要的一部分, 它直接决定了智能体的当前动作, 同时会影响后续的状态变量。在本文方法中, 策略函数的输入是历史动作值以及来自价值函数的误差, 输出是当前最优动作。假设 AUV 上一时刻的舵角为 φ_{last} , 当前舵角为 φ_{now} , 将策略函数记为 $g(\cdot)$, 则上述过程可表示为

$$\varphi_{\text{now}} = g(\varphi_{\text{last}}, E_1, E_2, E_3). \quad (12)$$

在自导过程中, 当未检测到有效目标 ($E_1 = 0$) 时, 应根据具体情况设定特殊的再搜索策略, 以确保 AUV 能够再次航行到信噪比较高的区域。常规情况下, 可以采取蛇形或环形搜索的方法以扩大搜索扇面。在知道目标大致方位的情况下, 也可以采取定向直航的方式使得 AUV 再次航行到高信噪比区域。因此, 初始信噪比应当大于信噪比阈值。否则, 应按照提前设定的再搜索策略使得 AUV 航行到较高信噪比区域, 才能进行有效的导引。

当检测到有效目标 ($E_1 = 1$) 时, 需要基于多普勒频移对 AUV 运动趋势进行判断: 当 $E_3 \leq \varepsilon$ 时, 表示 AUV 正在远离目标, 此时应改变舵角方向, 反之亦然。而舵角的大小则由误差 E_2 控制。该决策过程为

$$\varphi_{\text{now}} = \begin{cases} \text{sign}(\varphi_{\text{last}}) \cdot E_2, & E_3 > \varepsilon, \\ (-1) \cdot \text{sign}(\varphi_{\text{last}}) \cdot E_2, & E_3 \leq \varepsilon, \end{cases} \quad (13)$$

其中, $\text{sign}(\cdot)$ 为取符号函数; $\varepsilon \leq 0$ 为频率变化率阈值, 当 AUV 接收信号中心频率的平均变化率小于 ε 时表示 AUV 正在偏离目标航行。需注意的是, 实际工程应用中 AUV 的舵角不能过大。本文中 AUV 的最大舵角为 30° , 当 E_2 超过极限值时, 应对其限幅。

2 仿真实验及结果

2.1 仿真条件

仿真中训练的模型参数最终会在湖试实验中进行测试, 因此仿真实验环境的参数设置应与湖试实验相同。声源发射信号为窄带连续信号, 中心频率为 37.5 kHz, 带宽为 30 kHz, 声源级为 120 dB, 此频率

和声源级参数来自本文湖试使用的水下信标。噪声级设为 30 dB, 与实际湖试环境相近。AUV 和目标的初始距离不大于 1000 m, 航行深度均在水下 2 m。AUV 的速度为 8 kn, 旋转角速度由舵角大小决定。由 1.2.4 节可知, 舵角最大值为 30°, 因此 AUV 最大旋转角速度为 30 (°)/s。最后, 选择常规的 1 s 为自导控制周期, 自导采样周期为 T_c ($T_c < 1$)。

2.2 评价指标

最近会遇点 (Closest Point of Approach, CPA) 指 AUV 航行过程中与声源目标距离最近的点。脱靶量 (Miss Distance, MD) 指 AUV 在 CPA 点与目标的距离。采用平均脱靶量 (Mean Miss Distance, MMD) 作为评价指标, 表示 1000 次导引仿真实验脱靶量的平均值。

2.3 仿真结果

如 1.2 节所述, 本文所提导引方法中有多个重要的参数, 即信噪比阈值、检测率阈值和频率变化率阈值, 它们将直接影响导引任务的性能。为了满足实际导引过程稳定性的需求, 采取离线强化学习方法。即在实际部署或测试时将参数固定下来不再更新, 但会提前通过大量的任务模拟进行最优参数组合的选择。具体来讲, 首先进行仿真环境搭建, 然后让 AUV 按照所提方法在不同初始距离和不同目标运动速度下进行多次导引任务, 将 MMD 作为目标函数, 以目标函数最小化作为参数优化方向。

在实际自导任务中, 目标通常保持静止或直线运动, 偶尔会采取蛇形或环形的方式航行以扩大搜索扇面。在训练时, 目标只需保持静止或者直线运动, 最大航行速度为 8 kn。而在测试时, 增加了其他曲线运动情况 (蛇形运动和环形运动) 以验证训练出的模型对不同目标运动轨迹的鲁棒性, 目标的旋转角速度不大于 30 (°)/s。

2.3.1 训练结果

为了探究自导采样周期对算法性能和模型训练的影响, 在自导控制周期为 1 s 的情况下, 使用不同的自导采样周期训练模型, 结果如图 4 所示。自导采样周期越短, 训练总时长越长, 相应的 MMD 越小。当自导采样周期为 0.01 s 时, MMD 为 2.32 m, 但训练总时长达到了 6.95 h; 而当自导采样周期为 0.5 s 时, 训练总时长降低到了 1.17 h, 但 MMD 增加到 52.2 m。总的来说, 对于 1 s 的导引控制周期, 0.1 s 的自导采样周期可以同时达到较低的模型训练时长和较小的 MMD。

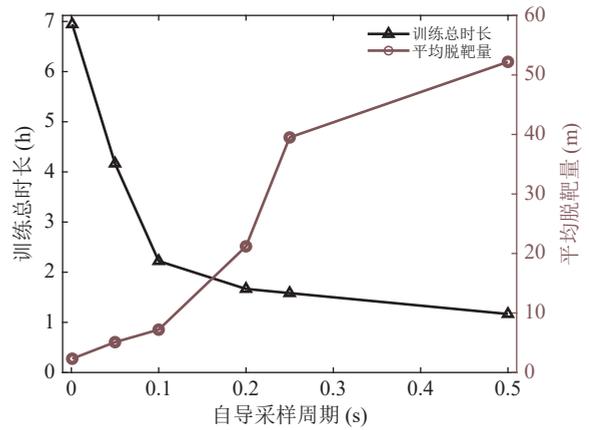


图 4 不同自导采样周期下的模型训练时长和平均脱靶量

出现以上现象的原因是, 当自导采样周期越短, 自导控制周期内累积的数据量越大, 解算的中心频率变化率误差越小, 因此导引误差越低。但是, 计算量会增加, 进而影响训练的效率, 同时也会影响实际导引的实时性。因此, 自导采样周期需要折衷选择。

2.3.2 测试结果

由图 4 可知, 0.1 s 的自导采样周期较为合理, 因此选择使用 0.1 s 的自导采样周期训练的模型参数进行测试。AUV 和目标的初始距离均为 500 m。

(1) 目标运动速度和轨迹的影响

从图 5 可以看到, 目标速度增大时, MMD 明显增大。若目标速度小于 3.2 kn, 则 MMD 可达 10 m 以下; 若目标速度达到 4.8 kn 及以上, 则 MMD 可达 100 m 以上, 算法性能下降。这是因为如 1.2 节所述, 本文方法对目标速度具有一定的限制, 过高运动速度的目标对多普勒频移的影响太大, 此时根据接收信号中心频率变化率对 AUV 相对目标运动趋势进行判断的错误概率会较大。此外, 对于训练时未曾考虑的目标蛇形和环形运动情况, MMD 并没有受太大影响。这表明, 本文方法对目标运动方向的鲁棒性较强, 可跟踪不同运动轨迹的目标。图 6 对某三

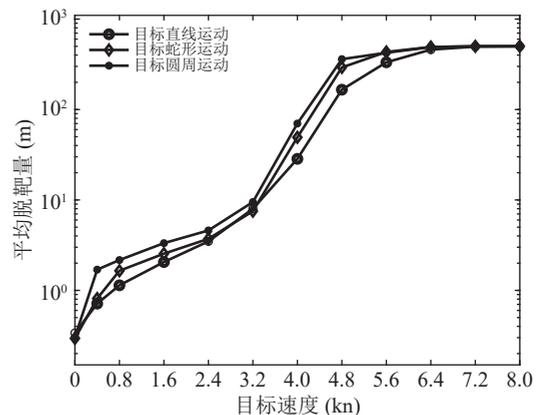


图 5 不同目标运动速度和轨迹下的平均脱靶量测试结果

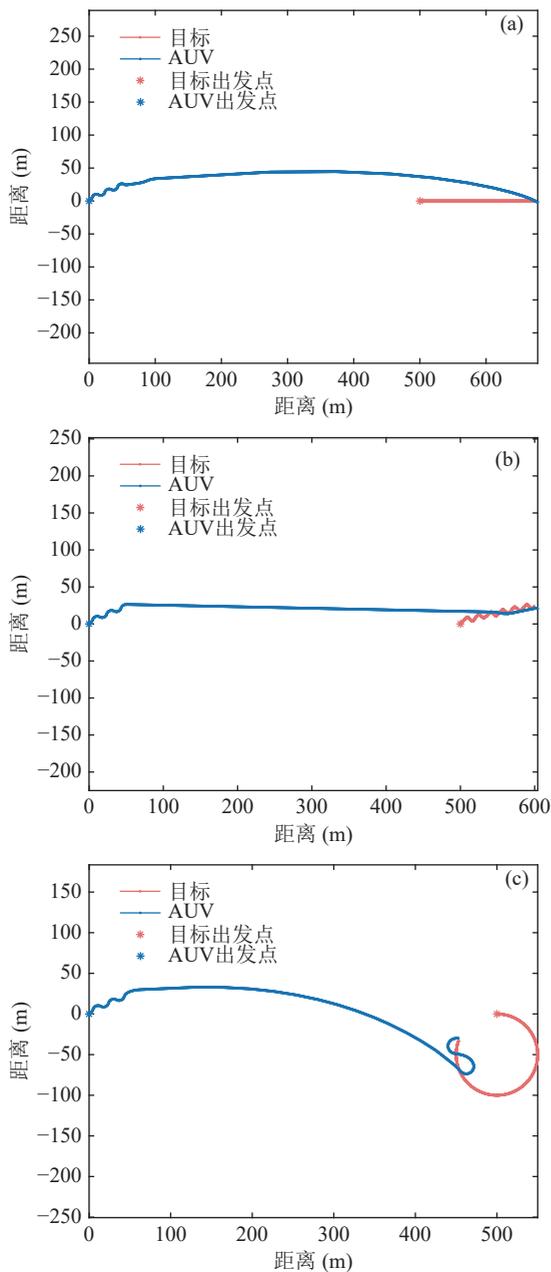


图6 不同目标运动情况下的测试导引轨迹 (a) 目标直线运动; (b) 目标蛇形运动; (c) 目标环形运动

次不同目标运动情况下的导引轨迹进行了可视化, 目标速度均为 2 kn。在图 6(b) 和图 6(c) 中, 目标的旋转角速度分别为 $30\text{ }(^{\circ})/\text{s}$ 和 $2.5\text{ }(^{\circ})/\text{s}$ 。

(2) 目标信号中心频率设置误差的影响

由式 (13) 可知, E_2 决定了策略函数所输出舵角的大小。根据 1.2.3 节, 在计算 E_2 的过程中, 需要使用目标信号中心频率。因此, 如果目标信号中心频率设置出现误差, 则会使得 AUV 导引过程中的舵角大小产生误差, 进而影响平均脱靶量。通过仿真实验分析了不同目标信号中心频率设置误差对平均脱靶量的影响, 此处的误差指相对误差, 表示为百分

比。实验中, 目标始终保持静止。

由图 7 可见, MMD 随着目标信号中心频率设置的相对误差增大而逐渐增加, 随后趋于平稳。出现此现象的原因时, 当中心频率设置误差过大时, 多数情况下所计算的多普勒频移过大, 进而导致所计算舵角过大。由 1.2.4 节可知, 此时 AUV 会对舵角进行限幅。因此, 当误差过大时, AUV 多数情况下都按照统一的限幅值进行打舵, 使得 MMD 相近。此外, AUV 的舵角方向仅由中心频率变化率和历史舵角方向进行控制, 不会受到此误差的影响。

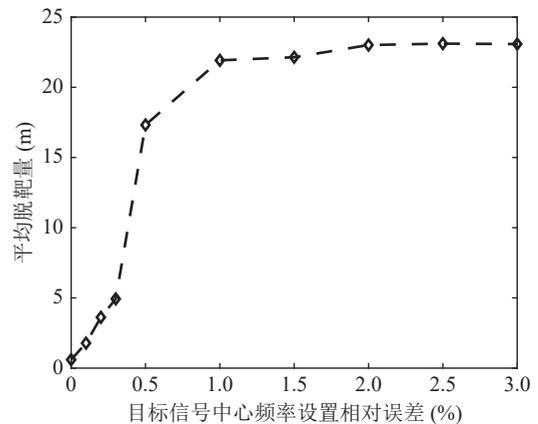


图7 目标中心频率设置相对误差对平均脱靶量的影响

3 湖试验证

在丹江口水库进行了湖试实验验证本文方法。湖面最大宽度约为 20 km, 湖水深度约为 100 m。所使用 TH-B010 型 AUV 的质量为 50 kg, 总长度为 2 m, 直径为 200 ms。声源目标固定在航船下方水下 2 m 处, 跟随航船一起运动或保持静止, 其位置由航船 GPS 代替。声源发射信号为窄带连续信号, 中心频率为 37.5 kHz, 声源级为 120 dB。经测量, 环境噪声级约为 30 dB。自导采样周期和自导控制周期分别为 0.1 s 和 1 s。每快拍信号时长为 32.77 ms。采集卡的采样频率为 125 kHz。AUV 航行深度约为 2 m, 装有电子罗盘、深度记和 GPS 模块。当 AUV 开始任务后, 会实时对自身位置进行航位推算, 并将数据记录到系统中。实验结果表明, 当 AUV 以 8 kn 速度运动时, 在目标静止的情况下, 平均脱靶量为 0.73 m; 在目标以 2 kn 速度运动的情况下, 平均脱靶量 5.69 m。为了说明导引的正确性, 对其中两次实验的系统数据进行分析 and 讨论。

(1) 目标静止

在此次实验中, 目标经纬度分别为 111.6294°E 和 32.7582°N , AUV 起点的经纬度分别为 111.6195°E

和 32.7641°N , 两者初始距离为 1280 m。AUV 的航行总时间设定为 6 min。AUV 运动速度为 8 kn, 目标保持静止。

AUV 的运动轨迹和目标位置如图 8(a) 所示, AUV 逐渐逼近目标, 经测算此次导引实验的脱靶量为 0.23 m。AUV 航行时电子罗盘所记录的航向角随时间变化关系见图 8(b), 可与 AUV 轨迹相互印证。

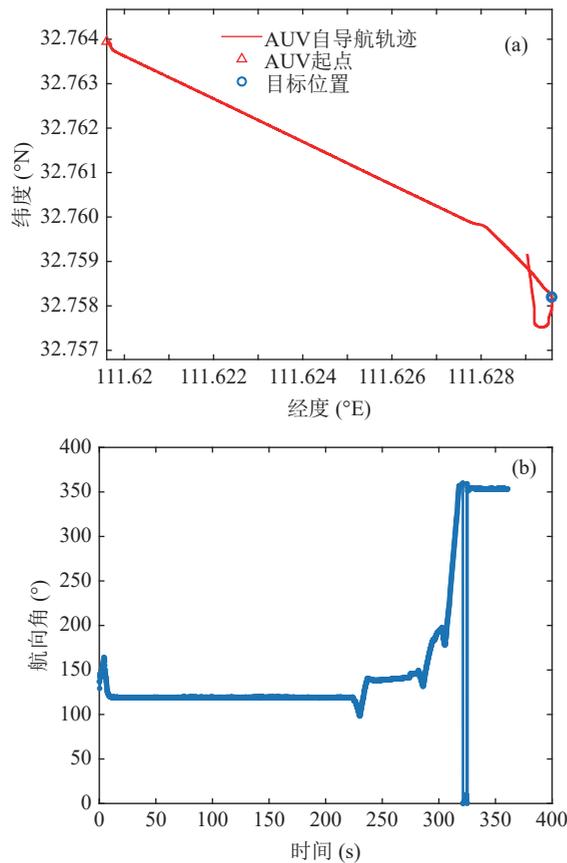


图 8 某次目标静止湖试实验的 AUV 导引轨迹及航向角随时间变化 (a) 导引轨迹; (b) 航向角

图 9(a) 为 AUV 自导过程中接收信号中心频率随时间的变化。在最初的 200 s 左右, 中心频率较高, 目标方位角趋近于 0° , 因此舵角极小, AUV 趋近于直航。而在第 230 秒和第 280 秒, AUV 正在偏离目标, 导致中心频率快速下降。此时, AUV 可以及时自主地改变舵角方向, 再次逼近目标, 使得中心频率上升。在第 310 秒和第 320 秒时, AUV 已经越过目标, 中心频率再次下降, AUV 仍然可以及时变换舵角方向对目标进行逼近。图 9(b) 为接收信号信噪比随时间的变化, 在前 300 s 中, AUV 逐渐逼近目标, 信噪比逐渐上升, 而当 AUV 越过目标后, 信噪比逐渐下降。

(2) 目标运动

此次实验中, AUV 起点的经纬度同样为

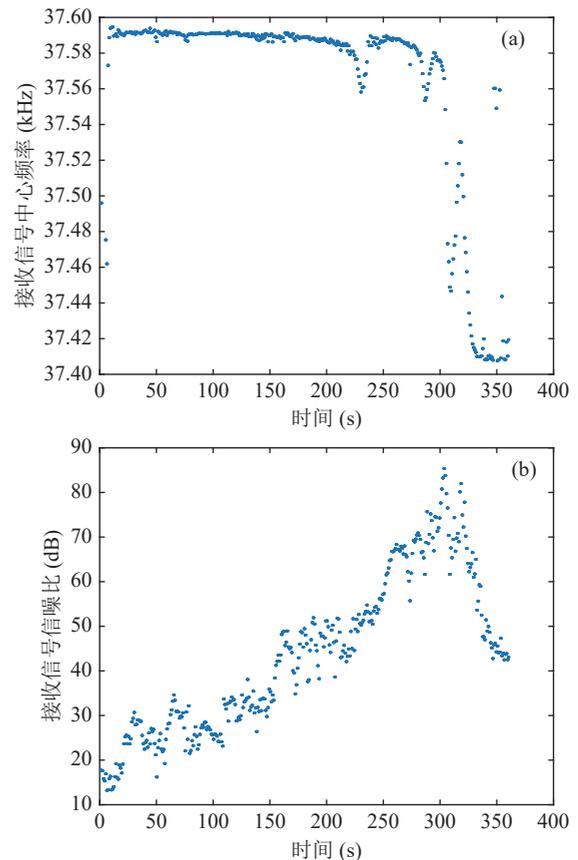


图 9 某次目标静止湖试实验的接收信号中心频率及信噪比随时间变化 (a) 接收信号中心频率; (b) 接收信号信噪比

111.6195°E 和 32.7641°N , 目标起点的经纬度为 111.6281°E 和 32.7579°N , 两者初始距离为 1177 m。AUV 的航行总时间设定为 350 s。AUV 的运动速度同样为 8 kn, 目标以 2 kn 的速度做直线运动。图 10 和图 11 为此次实验的内记数据结果, 在目标运动的情况下, AUV 依然可以对目标进行有效的跟踪逼近。当接收信号中心频率下降时, 可以及时改变舵角方向, 再次逼近目标使得中心频率上升。经测算, 此次实验的脱靶量为 5.3 m。

最后需要说明的是, 本文方法在湖中进行了验证。对于深海环境, 水体虽然会对航行轨迹会产生扰动, 但是多普勒频移的测量与估计并不会受到影响, 始终能够实时测得准确的水听器接收到的发射换能器的信号频率及多普勒频移, 可以反映出在各种水体运动情况下两者之间的相对位置关系。例如实际中海洋领域常用的用于测量各类航行器速度的多普勒计程仪并不受海水运动的影响。理论上本文方法在深海环境同样适用, 相关理论证明和海试将在后续工作中展开。此外, 在湖试实验时, 湖域比较开阔, 单水听器的指向性有限, 可以保证在一定程度上减轻反射和折射声线对测频精度的影响。实验信

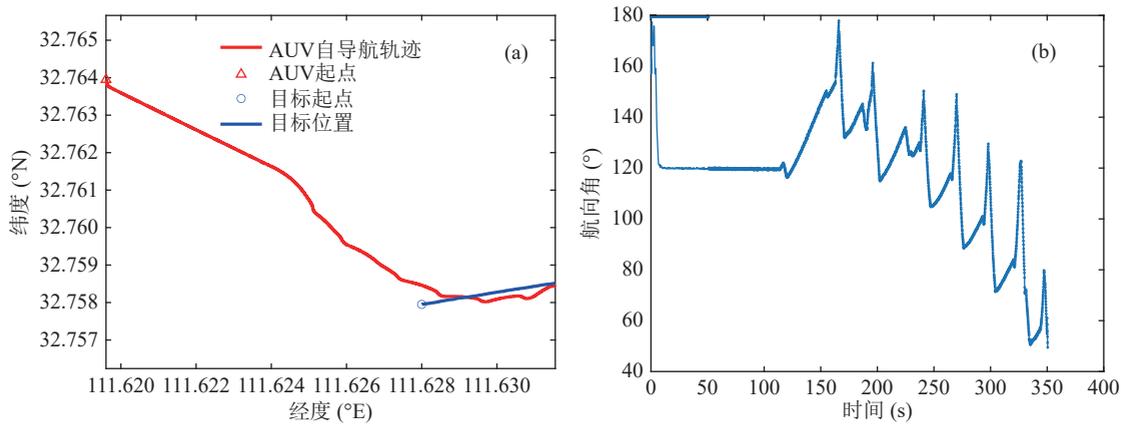


图 10 某次目标运动湖试实验的 AUV 导引轨迹及航向角随时间变化 (a) 导引轨迹; (b) 航向角

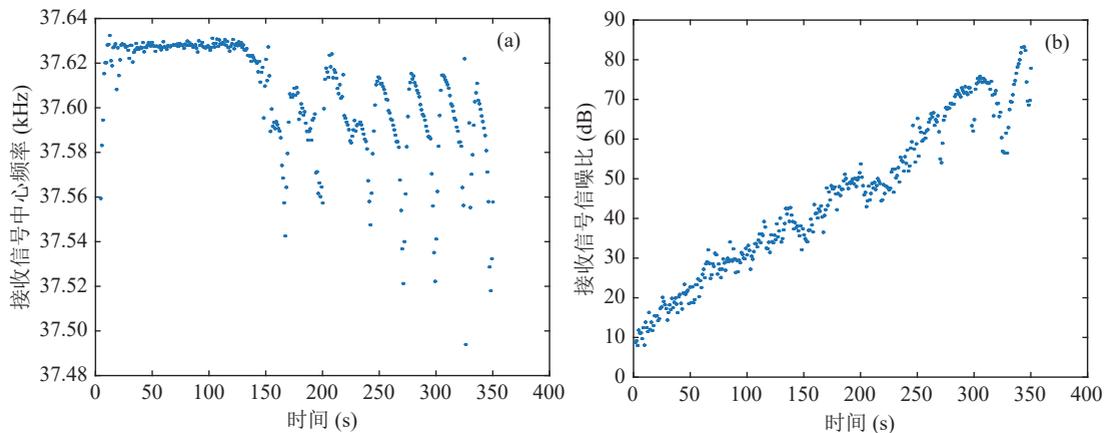


图 11 某次目标运动湖试实验的接收信号中心频率及信噪比随时间变化 (a) 接收信号中心频率; (b) 接收信号信噪比

道的多途效应对导引方法的影响及解决措施,也是未来的工作方向之一。

4 结论

本文基于强化学习 Actor-Critic 算法设计了一种单水听器无源自导方法,该方法可从单个声压水听器接收的声多普勒信息中提取关键特征,利用价值函数和策略函数估计目标方位和相对运动趋势,引导 AUV 跟踪逼近声源目标。仿真实验结果表明,减小采样周期能够降低平均脱靶量,但模型训练时长会增加。当目标静止或运动时,该方法均能获得较为精确的跟踪效果。在丹江口水库进行了湖试,并就其中两次湖试结果进行了详细的讨论分析,进一步验证了所提算法的有效性。

参 考 文 献

- 1 黄颖淞,葛辉良,王付印,等.蛙人探测声呐系统发展综述. *水下无人系统学报*, 2020; **28**(1): 1-9
- 2 杜召平,陈刚,王达.国外声呐技术发展综述. *舰船科学技术*, 2019; **41**(1): 145-151
- 3 Breivik M, Fossen T I. Guidance laws for autonomous underwa-

- ter vehicles. In: *Underwater Vehicles*, IntechOpen, 2009
- 4 Scudere A R. Assessing the feasibility of integrating an inertial navigation system and DVL into an a-sized AUVs. Master's thesis, USA: University of Rhode Island, 2019: 1-20
- 5 康帅,俞建成,张进.微小型自主水下机器人研究现状. *机器人*, 2023; **45**(2): 219-237
- 6 王超,笄良龙,韩梅,等.单矢量水听器稀疏近似最小方差方位估计算法. *声学学报*, 2021; **46**(6): 1050-1058
- 7 戚聿波,周士弘,任云,等.浅海中利用单水听器的声源被动测距. *声学学报*, 2015; **40**(2): 144-152
- 8 Zhang W D, Guan L G, Zhang G J, *et al.* Research of DOA estimation based on single MEMS vector hydrophone. *Sensors*, 2009; **9**(9): 6823-6834
- 9 Tichavsky P, Wong K T, Zoltowski M D. Near-field/far-field azimuth and elevation angle estimation using a single vector hydrophone. *IEEE Trans. Signal Process.*, 2001; **49**(11): 2498-2510
- 10 James D A. Principles of Doppler homing guidance. *Radar Homing Guid. Tactical Missiles*, 1986; **5**(1): 56-61
- 11 Bai Y, Yang Z, Li R, *et al.* Research on single channel Doppler direction finding system. International Conference on Digital Image Processing, SPIE, Guangzhou, China, 2019: 1220-1226
- 12 Sutton R S, Barto A G. Reinforcement learning: An Introduction. MIT Press, 1998
- 13 Xu J, Huang F, Wu D, *et al.* A learning method for AUV collision avoidance through deep reinforcement learning. *Ocean Eng.*, 2022; **260**: 112038
- 14 Cao X, Sun H, Guo L. Potential field hierarchical reinforcement learning approach for target search by multi-AUV in 3-D under-

- water environments. *Int. J. Control*, 2020; **93**(7): 1677–1683
- 15 Gaskett C, Wettergreen D, Zelinsky A. Reinforcement Learning applied to the control of an Autonomous Underwater Vehicle. Australian Conference on Robotics and Automation, Brisbane, Australia, 1999: 1–7
- 16 Konda V, Tsitsiklis J. Actor-critic algorithms. Advances in Neural Information Processing Systems, Denver, Colorado, USA, 1999: 1009–1014
- 17 Peters J, Schaal S. Natural actor-critic. *Neurocomputing*, 2008; **71**(7-9): 1180–1190
- 18 Ezeafulukwe U A, Darus M, Fadipe-Joseph O. On analytic properties of a sigmoid function. *Int. J. Math. Comput. Sci.*, 2018; **13**(2): 171–178
- 19 Hoyer E, Stork R. The zoom FFT using complex modulation. International Conference on Acoustics, Speech, and Signal Processing, IEEE, Hartford, CT, USA, 1977: 78–81
- 20 Pei D, Yang S, Yang H, *et al.* High efficient and real-time realization of Zoom FFT based on FPGA. International Conference on Computer Application and System Modeling, IEEE, Taiyuan, China, 2010: 669–673

声学消息

《声学学报》2023 年度优秀论文

经《声学学报》编辑委员会和青年编辑委员会评议,“声速剖面对深海远程声脉冲结构的影响”(吴双林, 秦继兴, 李整林, 吴禹沈, 张仁和. 声学学报, 2023, 48(4): 774–784)和“基于超快超声多普勒的三维脑损伤成像方法研究”(许凯亮, 付亚鹏, 闫少渊, 隋怡晖, 他得安, 王威琪. 声学学报, 2023, 48(1): 173–181)两篇论文获选《声学学报》2023 年度优秀论文。

“声速剖面对深海远程声脉冲结构的影响”一文及其相关工作利用东印度洋和南海海域的深海远程声传播实验数据, 比较分析了声道轴深度发射的声信号在两个海域不同声速剖面结构下的远程传播损失和脉冲时间到达结构, 发现东印度洋与南海的脉冲时间到达结构存在显著差异。在实验观测的基础上, 该论文进一步结合深海声道的参数化数学模型, 分析了声速剖面对远程脉冲传播时间到达结构的影响机理, 理论解释了实验中观测到的脉冲声信号时间到达结构现象, 并确定了造成两个海域实验结果差异的关键参数。该研究对水声探测与通信技术在不同深海环境中的应用具有很好的理论意义与工程应用前景。

“基于超快超声多普勒的三维脑损伤成像方法研究”一文针对脑血管的三维成像问题, 将超快超声成像、超快功率多普勒和机械扫描相结合, 实现了微型位移平台精准控制、相控超声超快发射、信号高速采集和压缩存储的三维超声成像系统; 该系统可在 GPU 并行运算的支持下, 实现高效波束合成计算和三维图像重建。在体实验表明, 该系统可用于三维大鼠脑血管网络成像, 以及脑血管损伤区域定位与量化评价。该研究为脑损伤的无创诊断提供了新方法, 相关思路对脑病检测技术发展及诊断方法进步具有很好的借鉴意义。

《声学学报》优秀论文评选迄今已进行了 24 届, 旨在促进声学领域基础研究和应用研究的发展, 鼓励作者开展创新、系统、深入的研究工作, 写出高质量论文。获知研究团队论文获奖, 两文作者均发来感谢信。秦继兴研究员表示:“团队的论文能够获此殊荣我们倍感荣幸, 我谨代表全体作者向编辑部和评委表示衷心感谢!《声学学报》是国内声学领域的顶级期刊, 为声学领域的科研人员提供了宝贵交流平台, 也让我们团队多年来的研究成果能够呈现在广大从业者面前。在论文发表过程中,《声学学报》的编辑和审稿专家给予了科学严谨、细致入微的指导, 多角度的修改意见和建议对我们的研究工作帮助很大, 同时在学术论文写作方面也令我们收获良多。我们将珍惜荣誉, 继续努力做好水声学基础研究及相关工作, 多贡献创新性成果。”许凯亮研究员表示:“非常荣幸我们团队的成果能获此殊荣。首先对《声学学报》的长期支持和帮助表示衷心感谢! 近日,‘快速超分辨超声成像’相关研究也入选了中国工程院发布的《全球工程前沿 2024》报告。在此, 由衷地感谢各位评委专家对医学超声领域相关进展的肯定, 以及对团队工作的支持和指导!《声学学报》是声学科科研人员共同的精神家园。在未来工作中, 我们一定会加倍努力, 坚持把高水平声学论文写在祖国大地上, 坚持把声学科技成果应用在实际现代化的伟大事业中。”

《声学学报》编辑委员会
2024 年 12 月 22 日