

DOI: 10.3969/j.issn.1671-0673. 2026. 02. 004

基于 AC 网络的扩展目标跟踪认知雷达 波形设计

张雅歌, 胡 涛, 王建涛, 李一鸣, 杨静静

(信息工程大学, 河南 郑州 450001)

摘要: 针对扩展目标机动过程中, 需依据实时交互获得的运动与形状状态信息, 在线生成兼具高精度与高稳定性的发射波形问题, 提出一种基于演员—评论家(AC)网络的智能波形设计方法。首先, 该方法在交互式多模型(IMM)框架下, 结合无迹卡尔曼滤波(UKF)与联合概率数据关联(JPDA), 实现对目标运动状态与形状状态的联合估计; 其次, 构建以高斯瓦瑟斯坦距离(GWD)为核心的多要素奖励函数; 最后, 驱动 AC 网络与环境实时交互, 实现在线波形优化设计。仿真结果表明, 与基于互信息(MI)准则、固定波形和深度 Q 网络(DQN)相比, 所提方法在目标形状估计精度、跟踪稳定性以及波形参数选择方面均表现出显著优势, 验证了其在高精度扩展目标跟踪中的有效性, 具有良好的工程应用价值。

关键词: 演员—评论家网络; 扩展目标跟踪; 认知雷达; 波形设计

中图分类号: TN957.51

文献标识码: A

文章编号: 1671-0673(2026)02-0153-08

Cognitive Radar Waveform Design for Extended Target Tracking Based on AC Networks

ZHANG Yage, HU Tao, WANG Jiantao, LI Yiming, YANG Jingjing

(Information Engineering University, Zhengzhou 450001, China)

Abstract: To address the problem of online generating transmit waveforms with both high precision and high stability for maneuvering extended targets, which requires real-time motion and shape state information obtained through interactive sensing, an intelligent waveform design method based on an Actor-Critic (AC) network is proposed. Firstly, Under the interactive multiple model (IMM) framework, the unscented Kalman filter (UKF) and joint probabilistic data association (JPDA) are integrated to achieve joint estimation of target motion and shape states. Secondly, a multi-component reward function centered on the Gaussian Wasserstein distance (GWD) is constructed. Finally, the AC network is driven to interact with the environment in real time, enabling online waveform optimization design. The simulation results demonstrate that, compared with the mutual information (MI) criterion, fixed waveforms, and the deep Q-network (DQN), the proposed method exhibits significant advantages in target shape estimation accuracy, tracking stability, and waveform parameter selection, verifying its effectiveness in high-precision extended target tracking and highlighting its potential for practical engineering applications.

Key words: Actor-Critic network; extended target tracking; cognitive radar; waveform design

收稿日期: 2025-10-30; 修回日期: 2026-01-30

作者简介: 张雅歌(1982—), 女, 副教授, 博士生, 主要研究方向为雷达信号处理。E-mail: yagebu82@163.com

认知雷达^[1]通过“感知—学习—决策”的闭环机制,能够依据环境与目标特性自适应调整发射波形,从而提升雷达性能,其中,波形优化设计^[2]是认知雷达实现闭环自适应能力的核心环节。文献[3]最早提出了基于最大互信息准则(Mutual Information, MI)的波形优化设计。文献[4]将交互式多模型(Interactive Multiple Model, IMM)、概率数据关联(Probabilistic Data Association, PDA)滤波与 MI 准则相结合,提出了面向机动目标的自适应跟踪算法。然而,上述方法主要依赖于离线设计与先验模型,难以实现与环境的实时交互及在线优化。随着智能算法^[5-6]的发展,强化学习被引入波形选择领域。文献[7]提出了基于熵奖励 Q 学习(Entropy Reward Q Learning, ERQL)的在线波形选择算法,文献[8]探索了基于深度 Q 网络(Deep Q-Network, DQN)的智能波形选择方法。但以上研究均以点目标作为跟踪对象。

随着宽带高分辨率雷达的广泛应用,目标在距离—多普勒域中往往占据多个分辨单元,从而被建模为扩展目标^[9]。针对扩展目标的波形设计研究随之展开。文献[10]基于 MI 准则提出了干扰环境下的网络雷达波形设计模型,能够有效提升扩展目标的检测性能,文献[11]在考虑能量约束的条件下,提出了基于最大相对熵的波形设计,对扩展目标的检测性能有所提升。然而,现有工作大多集中于静止的扩展目标场景。

在现代战场环境中,先进战机、无人机集群等高价值目标均呈现典型的扩展特征。研究能够自适应目标运动状态与形状变化的智能波形设计方法,对于提升高威胁目标的持续跟踪能力具有重大意义。与点目标及静止扩展目标不同,运动扩展目标的波形设计需实时估计目标的运动轨迹与形态变化,并和环境形成动态交互感知,这对雷达波形设计的自适应能力及智能化水平提出了前所未有的要求。近期的研究展示了深度强化学习在解决复杂决策问题上的潜力,如文献[12]将演员—评论家(Actor-Critic, AC)网络架构应用于机载雷达扩展目标跟踪(Extend Target Tracking, ETT)的轨迹规划问题,展现出了优越的性能,为本文的研究提供了重要的思路借鉴。

针对 ETT 需依据实时交互获得的运动与形状状态信息,在线生成兼具高精度与高稳定性的发射波形问题,本文提出一种基于 AC^[13]网络架构的 ETT 波形设计算法。本文主要工作如下:

1)设计了基于交互式多模型—无迹卡尔曼滤波(Unscented Kalman Filter, UKF)—联合概率数据关联(IMM-UKF-JPDA^[14])的联合滤波框架,实现对机动扩展目标运动与形状的同步估计;

2)采用 AC 网络架构与环境实时交互,实现波形参数实时动态变化,提升波形选择的灵活性;

3)设计了以高斯瓦瑟斯坦距离^[15](Gaussian Wasserstein distance, GWD)为核心的多要素奖励函数,综合衡量跟踪精度与形状估计性能。

1 系统模型

基于 AC 网络的扩展目标跟踪波形设计框架如图 1 所示。

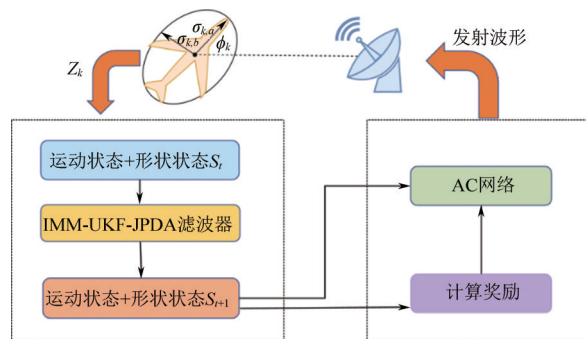


图1 基于AC网络的扩展目标跟踪波形设计框架图

本文基于两坐标雷达建模,对于包含长宽高三维的扩展目标,忽略其高度信息,依据文献[9],将目标建模为二维椭圆。以飞机质心为原点,以飞机长度方向为椭圆长轴,以飞机宽度为短轴,具体建模示意如图1所示。AC网络通过实时感知目标当前状态,选择输出最优波形参数用于下一周期的量测,从而形成一个“感知—决策—优化”的智能闭环框架。

1.1 扩展目标状态模型

在 ETT 领域, Koch^[9]首次提出使用随机矩阵对目标的扩展状态建模,通过逆威沙特(Wishart)分布描述目标形状的不确定性,奠定了扩展目标跟踪的理论基础。该方法将扩展状态建模为一个椭圆,并用一个二维对称正定随机矩阵表示椭圆的大小和方向。

假设扩展目标跟踪在时间 k 时的状态可表示为

$$\zeta_k = (\xi_k, \eta_k) \quad (1)$$

式中:运动状态 $\xi_k = [x_k, y_k, \dot{x}_k, \dot{y}_k]^T$, 包括位置状态 $\xi_k^p = [x_k, y_k]^T$ 和速度状态 $\xi_k^v = [\dot{x}_k, \dot{y}_k]^T$; x_k, y_k 代表 k

时刻目标质心在 x 轴、 y 轴的位置, $\dot{x}_k, \dot{y}_k$ 代表 k 时刻目标质心在 x 轴、 y 轴的速度; η_k 表示扩展目标的形状,建模为椭圆^[9],表达式为 $\eta_k = [\sigma_{k,a}, \sigma_{k,b}, \phi_k]^T$, 其中, $\sigma_{k,a}$ 表示 k 时刻椭圆长半轴值, $\sigma_{k,b}$ 表示 k 时刻椭圆短半轴值, ϕ_k 表示 k 时刻在平面内的旋转角(相对于 x 轴)。文献[9]指出,椭圆的大小和方向由一个二维正定随机矩阵 H_k 表示,有:

$$H_k = R(\phi_k) \text{diag}\{\sigma_{k,a}^2, \sigma_{k,b}^2\} R(\phi_k)^T \quad (2)$$

$$R(\phi_k) = \begin{bmatrix} \cos \phi_k & -\sin \phi_k \\ \sin \phi_k & \cos \phi_k \end{bmatrix} \quad (3)$$

椭圆扩展状态的质心位置由运动状态 ξ_k 的位置分量确定,对 $\sigma_{k,a}^2, \sigma_{k,b}^2$ 取平方根可分别得到椭圆长半轴和短半轴的长度, ϕ_k 通过求解 H_k 的特征向量确定, H_k 服从逆 Wishart 分布^[9]。

1.2 雷达测量模型

雷达发射波形有多种,由于高斯包络线性调频波形具有良好的距离—多普勒分辨特性、易于产生与处理、参数灵活可调等优点^[16],常被作为波形自适应设计的基波形或备选波形集的一员。本文采用包络的线性调频波形设计,其表达式为

$$s_t(t) = A \exp\left(-\frac{t^2}{2T_p^2}\right) \exp(j2\pi(f_c t + \frac{1}{2} K_r t^2)) \quad (4)$$

式中: t 代表时间, $|t| \leq \frac{T_p}{2}$; A 代表发射信号幅度; T_p 是脉冲宽度; f_c 是信号载频; K_r 是调频斜率。

根据最大似然估计理论^[17],雷达匹配接收的克拉美罗下界(Cramer-Rao Lower Bound, CRLB)为 Fisher 信息矩阵 $J(\theta)$ 的逆。KERSHAW 等^[18]提出在高信噪比条件下,估计误差可以达到 CRLB^[9],并推导了高斯包络线性调频信号对于距离速度观测的量测误差协方差 CRLB 表达式:

$$R_k(T_p, K_r) = \begin{bmatrix} \frac{c^2 T_p^2}{2R_{SN}} & \frac{-c^2 K_r T_p^2}{f_c R_{SN}} \\ \frac{-c^2 K_r T_p^2}{f_c R_{SN}} & \frac{c^2}{f_c^2 R_{SN}} \left(\frac{1}{2T_p^2} + 2K_r^2 T_p^2 \right) \end{bmatrix} \quad (5)$$

式中: c 是光速; R_{SN} 是信噪比。从公式(5)可以看出,测量噪声协方差与脉冲持续时间 T_p 和线性调频斜率 K_r 有关。因此,可以通过一定的准则合理调整波形参数 $\theta = [T_p, K_r]^T$,从而有效提高跟踪精度。

2 扩展目标跟踪算法

2.1 扩展目标跟踪算法

对于扩展目标的跟踪,本文采用 IMM-UKF-

JPDA 算法,即在 IMM^[14]框架下利用 UKF-JPDA 滤波并行处理,同时对目标的运动状态和形状状态预测更新。具体流程见图2。

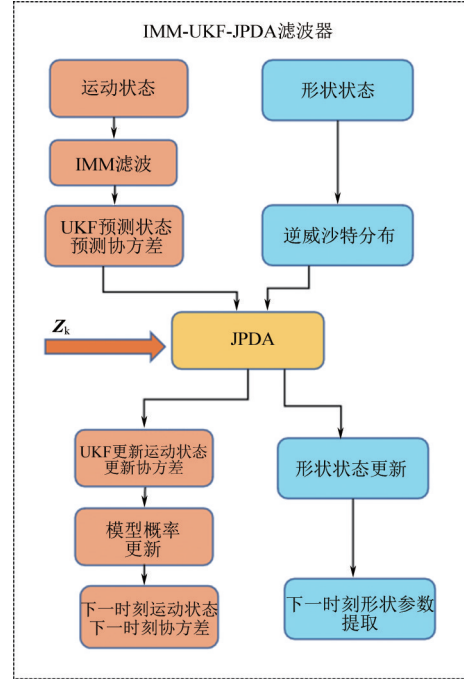


图2 扩展目标跟踪算法

假设 k 时刻的状态、协方差、模型概率分别为 X_k 、 P_k 、 μ_k ,模型的转移矩阵为

$$\Pi = [p_{ij}], p_{ij} = P(M_i \rightarrow M_j) \quad (6)$$

首先,进行模型交互得到每个模型的混合初始值权重 μ_k^{ij} ,据此构造模型 j 的混合先验状态 $\hat{X}_{k-1}^{ij}$ 与协方差 P_{k-1}^{ij} 。

其次,对每个模型独立进行 UKF-JPDA 滤波,以 $(k-1)$ 时刻的 $\hat{X}_{k-1}^{ij}, P_{k-1}^{ij}$ 为输入,通过 UKF 的时间更新步骤,计算预测状态 $\hat{X}_{k|k-1}^{ij}$ 和预测协方差 $P_{k|k-1}^{ij}$,以及预测量测 $\hat{Z}_{k|k-1}^{ij}$ 和新息协方差 S_k^{ij} ,由状态转移与测量函数得到模型条件的后验概率。考虑到量测可能源于杂波,而扩展目标的回波分别落入不同的波门,而彼此间又存在数据互联关系,采用联合概率互联算法 JPDA^[14]计算所有可能的关联事件,并最终计算出关联概率,即量测来自于同一个点目标的概率。

这里将一个扩展目标等价于聚集于椭圆形状内的多个点目标,其中量测函数和卡尔曼增益更新依赖于前一时刻 AC 网络输出的波形。最后,综合考虑所有量测计算每个模型的似然函数,并据此更新模型概率,进行模型组合,得到全局状态估计与协方差。

通过以上流程,“交互—并行滤波—概率更新—

融合”循环进行,自适应地在各个模型间切换。

2.2 更新步骤

2.2.1 时间更新

对于每个模型,利用 UKF 生成 Sigma 点,通过非线性状态转移函数 $f^{(m)}(\cdot)$ 传播,计算预测状态 $\hat{X}_{k|k-1}^j$ 与预测协方差 $\hat{P}_{k|k-1}^j$

形状状态的演化由参数 v_k 、 V_k 的衰减模型描述:

$$v_{k|k-1} = e^{-T_s/\tau} v_{k-1|k-1} \quad (7)$$

$$V_{k|k-1} = \frac{v_{k-1|k-1} - d - 1}{v_{k|k-1} - d - 1} \cdot V_{k-1|k-1} \quad (8)$$

式中: T_s 为采样间隔; τ 为衰减常数; $d = 2$ 为测量维数。

形状的期望为

$$E[H_{k|k-1}] = \frac{V_{k|k-1}}{v_{k|k-1} - 2d - 2} \quad (9)$$

对形状的期望进行特征分解可得目标椭圆的长半轴、短半轴和方向的估计。

2.2.2 测量更新

1) 量测预测。将 Sigma 点通过测量函数 $h(\cdot)$ 传播,得到预测测量 $\hat{z}_k^{(j)}$ 与新息协方差 S_k^j 。

2) 数据关联。考虑到杂波的存在,采用 JPDA 计算各个测量 $z_k^{(i)}$ 与目标的关联概率 $\beta_k^{i(i)}$ 。

3) 状态更新。使用加权新息更新状态与协方差为:

$$\hat{X}_k^j = \hat{X}_{k|k-1}^j + K_k^j \sum_i \beta_k^{i(i)} d_k^{i(i)} \quad (10)$$

$$P_k^j = P_{k|k-1}^j - (1 - \beta_k^{i(0)}) K_k^j S_k^j (K_k^j)^T + \tilde{P}_k^j \quad (11)$$

式中, $\tilde{P}_k^j$ 为关联不确定性引起的散布项。

4) 形状更新。利用关联后的测量质心 $\bar{z}_k$ 与测量散射矩阵 $\bar{Z}_k$ 更新逆 Wishart 分布参数:

$$\bar{Z}_k = \sum_{i=1}^{n_k} (z_k^{(i)} - \bar{z}_k)(z_k^{(i)} - \bar{z}_k)^T \quad (12)$$

$$v_{k|k} = v_{k|k-1} + n_k, V_{k|k} = V_{k|k-1} + L_k + \bar{Z}_k \quad (13)$$

式中, L_k 与 $\bar{Z}_k$ 分别为基于新息^[19]与散射矩阵的修正项,如式(14)~式(16):

$$L_k = H_{k|k-1}^{1/2} S_k^{-1/2} (\bar{z}_k - H_k X_{k|k-1}) \cdot (\bar{z}_k - H_k X_{k|k-1})^T S_k^{-1/2} H_{k|k-1}^{1/2} \quad (14)$$

$$\bar{Z}_k = X_{k|k-1}^{1/2} (Y_k)^{-1/2} \bar{Z}_k (Y_k)^{-1/2} X_{k|k-1}^{1/2} \quad (15)$$

$$Y_k = \frac{\rho E[H_{k|k-1}] + R_k}{n_k} \quad (16)$$

2.2.3 模型概率更新与融合

计算各模型的似然函数:

$$\Lambda_k^{(j)} = N(\bar{z}_k^{(j)}; 0, S_k^{(j)}) \quad (17)$$

更新模型概率:

$$\mu_k^{(j)} = \frac{\Lambda_k^{(j)} \bar{\mu}_k^{(j)}}{\sum_n \Lambda_k^{(n)} \bar{\mu}_k^{(n)}} \quad (18)$$

最终融合输出为

$$X_{k|k} = \sum_m \mu_k^{(j)} X_{k|k}^{(j)}, P_{k|k} = \sum_m \mu_k^{(j)} [P_{k|k}^{(j)} + (X_{k|k}^{(j)} - X_{k|k})(X_{k|k}^{(j)} - X_{k|k})^T] \quad (19)$$

3 波形设计算法

3.1 AC 网络结构

AC 网络是将策略梯度法和值函数近似相结合的一种混合型强化学习架构。其核心思想是分离策略表示与值函数表示,并同时二者进行优化。它包含两个相互作用的神经网络模块:一个是演员网络(Actor Network),一个是评论家网络(Critic Network)。其中,Actor 网络相当于演员,它可以根据目标当前的状态 s_t 通过策略 $\pi_\theta(a_t|s_t)$ 选择对应的动作,即输出雷达波形参数 $a_t = [T_p, K_r]^T$ 。其中状态 $s_t = (\xi_k, P_{k|k}, \eta_k, V_{k|k})^T$,即输入状态包含目标当前位置、速度、形状和协方差矩阵。Critic 网络相当于裁判,对当前 Actor 网络输出的动作参数进行评判打分^[20],输入与 Actor 网络相同,输出为动作价值函数 $Q_\phi(s_t, a_t)$,目的是促进 Actor 网络输出最优动作参数。具体如图 3 所示。

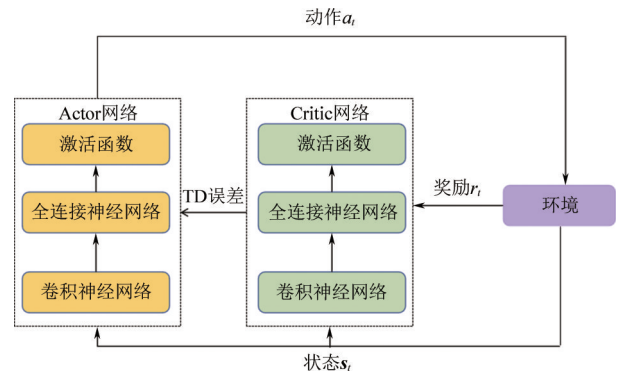


图3 基于AC网络的扩展目标跟踪波形设计框架图

Actor 与 Critic 网络均采用 3 层前馈网络,由卷积神经网络、全连接神经网络和激活函数组成, Critic 网络通过时序差分(Temporal Difference, TD)^[13]误差最小化进行评判打分。

3.2 目标函数设计

文献[12]对于高斯分布的 ETT,利用 GWD^[15]来衡量扩展目标估计值与真实值之间的误差性能。在时刻 k ,若将运动状态的位置分量 $\xi_{k|k-1}^p$ 与表示目标形态的扩展目标扩展状态 η_k 建模为多元高斯分布,则扩展目标在时刻 k 的先验概率分布与后验概

率分布^[21]之间的GWD由下式给出:

$$d_{\text{GWD}}^2(\mathbf{g}_{k|k-1}, \mathbf{g}_{k|k}) = \left\| \xi_{k|k-1}^p - \xi_{k|k}^p \right\|^2 + \text{tr} \left(\mathbf{H}_{k|k-1} + \mathbf{H}_{k|k} - 2\sqrt{\mathbf{H}_{k|k-1}\mathbf{H}_{k|k}} \right) \quad (20)$$

式中,描述扩展目标估计的先验概率分布和后验概率分布定义为:

$$\mathbf{g}_{k|k-1} \sim \mathcal{N}(\xi_{k|k-1}^p, \mathbf{H}_{k|k-1}) \quad (21)$$

$$\mathbf{g}_{k|k} \sim \mathcal{N}(\xi_{k|k}^p, \mathbf{H}_{k|k}) \quad (22)$$

$$\mathbf{H}_{k|k-1} = \mathbf{R}(\phi_{k|k-1}) \text{diag}\{\sigma_{k|k-1,a}^2, \sigma_{k|k-1,b}^2\} \mathbf{R}^T(\phi_{k|k-1}) \quad (23)$$

$$\mathbf{H}_{k|k} = \mathbf{R}(\phi_{k|k}) \text{diag}\{\sigma_{k|k,a}^2, \sigma_{k|k,b}^2\} \mathbf{R}^T(\phi_{k|k}) \quad (24)$$

式中, $\mathcal{N}(\cdot)$ 表示符合高斯分布。从式(20)可以看出, GWD公式包括目标质心位置的平方欧式距离 $\left\| \xi_{k|k-1}^p - \xi_{k|k}^p \right\|^2$, 同时包含了两个扩展矩阵的迹和, 以及衡量两个扩展矩阵“相似度”的项, 能够更公平、更全面地评价扩展目标跟踪的整体性能。

由此设计综合性目标奖励函数, 包含GWD^[15], 信息熵、距离误差、速度误差等, 设置如下:

$$r_t = -\alpha d_{\text{GWD}}^2(\mathbf{g}_{k|k-1}, \mathbf{g}_{k|k}) - \beta r_h - \gamma r_{\text{meas}} - \delta r_{\text{dist}} - \chi r_{\text{vel}} \quad (25)$$

式中: $r_h = \left| \log_2(1 + |\Delta h_k|) \right| \times \text{sign}(\Delta h_k)$ 表示信息熵; $r_{\text{meas}} = -\frac{1}{N} \sum_{i=1}^N |z_i - \hat{z}_i|$ 表示测量更新误差, N 表示当前时刻进入门限的有效量测数量; r_{dist} 表示距离误差, 用均方根误差表示; r_{vel} 表示速度误差, 用均方根误差表示。

该奖励函数通过综合考虑多种因素, 平衡了熵、形状、速度、距离和量测的估计质量, 更有利于跟踪性能的精确和稳定。

3.3 基于AC网络的波形选择算法

算法1 基于AC网络的扩展目标跟踪波形选择算法

输入: $(k-1)$ 时刻的扩展目标的状态估计, k 时刻的量测 $\mathbf{Z}_k$ 。

输出: 最佳发射波形参数。

1. 初始化参数: IMM滤波器状态、模型概率、AC网络参数。

2. for 每个跟踪时刻 do

(1) 使用当前波形参数发射波形, 获取量测集合 $\mathbf{Z}_k$;

(2) 执行IMM-UKF-JPDA滤波, 更新目标状态 $\hat{\mathbf{x}}_k$ 、协方差 $\mathbf{P}_k$ 及模型概率 $\boldsymbol{\mu}_k$;

(3) 从IMM-UKF-JPDA滤波算法输出中构建目标状态向量 $\mathbf{s}_t$;

(4) AC网络的Actor部分根据状态 $\mathbf{s}_t$ 通过策略函数选

择动作 $\mathbf{a}_t = [T_r, K_r]^T$;

(5) 执行动作 $\mathbf{a}_t$, 更新雷达波形参数;

(6) 计算即时奖励 r_t ;

(7) 更新Critic网络(最小时序差分误差)和Actor网络(策略梯度上升);

3. 判断是否满足终止条件: 达到最大步长或跟踪误差阈值; 如果是结束, 如果否重复2。

该算法通过将Actor网络输出的波形参数发射至目标, 并获取量测信息, 完成目标状态与模型概率的更新, 然后进行IMM-UKF-JPDA滤波。同时, 从这个过程中获得新的状态向量 $\mathbf{s}_{t+1}$, Actor网络根据新的状态输出具体的雷达发射参数, 用于下一个周期的量测, 从而形成一个“感知—决策—优化”的智能闭环框架。

4 仿真实验

4.1 仿真环境和参数设置

4.1.1 仿真环境

本文的仿真环境如下: 采用Matlab R2023a语言进行实验仿真, 整个算法在Windows 11系统上运行, 该系统配备了16核12th Gen Intel(R) Core(TM) i9-12900 CPU @2.40 GHz处理器。

4.1.2 扩展目标参数设置

本文依据文献[12]设计机动ETT场景。雷达初始位置为 $[-200 \text{ m}, 500 \text{ m}]$, 扩展目标初始位置为 $[0 \text{ m}, 0 \text{ m}]$ 。采样时间为 $T_s = 1 \text{ s}$, 共 $N = 55$ 个跟踪采样周期。扩展目标的长轴和短轴分别记为 $a = 10 \text{ m}$ 和 $b = 5 \text{ m}$ 。方位角定义为椭圆长轴与笛卡尔坐标系轴线之间的夹角, 其中逆时针方向定义为正方向, 初始值为 0° 。

每个时间索引内的测量次数服从强度为 $\lambda_k = 10 \times \sqrt{a \times b} + 0.5$ 的泊松分布。扩展目标的运动状态设定为 $\xi_k = [(\xi_k^p)^T, (\xi_k^v)^T]^T$, 包含位置和速度信息, 初始目标运动状态设定为 $\xi_0 = [5, 5, 17, -17]^T$ 。机动目标的运动模型由匀速(Constant Velocity, CV)模型和协同转弯(Coordinated Turning, CT)模型两种模型描述。

4.1.3 雷达参数设置

雷达参数设置如下: 载波频率 $f_c = 3 \text{ GHz}$, 光速 $c = 3 \times 10^8 \text{ m/s}$, 脉冲重复周期 $T_r = 0.001 \text{ s}$, 发射功率 $P_t = 20 \text{ kW}$, 天线增益 $G_t = G_r = 30 \text{ dB}$, 噪声系数 $F_r = 3 \text{ dB}$, 目标雷达截面积 $\sigma = 10 \text{ m}^2$, 波束宽度 $\beta_{3\text{dB}} = 3.5^\circ$ 。

波形参数选择范围为:脉冲宽度 $[1\ \mu\text{s}, 20\ \mu\text{s}]$ 、间隔 $1\ \mu\text{s}$ 和调频带宽 $[10\ \text{MHz}, 55\ \text{MHz}]$ 、间隔 $1\ \text{MHz}$ 。

4.1.4 AC网络参数设置

Actor和Critic网络均采用3层前馈网络,具体参数设置如表1。

表1 AC网络参数设置								
网络模块	输入层	隐藏层	激活函数	输出层	学习率	正则化系数	目标函数	网络训练
Actor网络	$s_t = (\xi_k, P_{k k}, \eta_k, V_{k k})^T$	200(两层)	ReLU	波形参数	0.001	0.2	策略梯度和熵正则化	缓冲区大小: 10 000; 批次大小: 32; 目标网络更新频率: 100
Critic网络	$s_t = (\xi_k, P_{k k}, \eta_k, V_{k k})^T$	200(两层)	ReLU	1维(状态价值)	0.001	0.2	TD误差最小化	

4.1.5 奖励函数权重设置

本文将各权重设置为: $\alpha = 0.7$ 、 $\beta = 0.1$ 、 $\gamma = 0.1$ 、 $\delta = 0.05$ 、 $\chi = 0.05$ 。

4.2 仿真结果

为验证基于AC网络的波形设计算法的有效性,通过仿真对比实验来评估算法的性能。对比的算法有固定波形(Fixed)、MI和DQN网络算法。

4.2.1 距离误差和速度误差

图4展示了不同算法对目标跟踪时的距离均方根误差。从图4可以看出,无论是CV运动还是CT运动,本文所提的AC算法在整个跟踪过程中始终维持较低的误差水平。经计算表明,与Fixed算法相比较,距离均方误差提升了约54.58%。

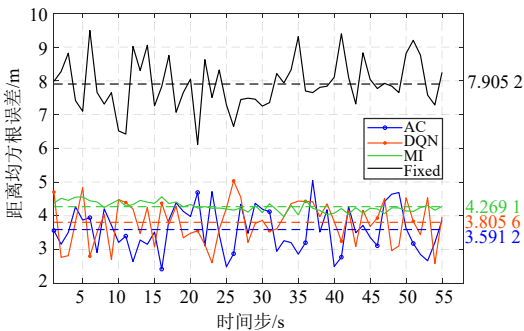


图4 距离均方根误差对比图

图5展示了不同算法对目标跟踪时的速度均方根误差。

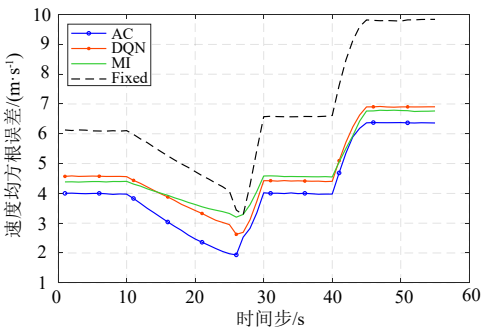


图5 速度均方根误差对比图

图5表明,所有算法在约28 s时达到速度均方根误差最小,但随着目标速度和朝向发生变化,速度均方根误差有所提升。但整体上,AC方法在速度跟踪中始终低于其他3种方法的误差水平。计算表明,与Fixed算法相比较,速度跟踪精度提高了约37.81%。

4.2.2 形状预估误差

在形状估计方面,图6~图9展示了形状估计方面的误差对比情况。图6反映了各方法在GWD评价指标中的对比情况。整体来看,基于DQN和AC网络的GWD误差低于MI和Fixd,表明基于强化学习的方法能够更精准地捕获目标的扩展特征。对比的方法中,基于AC网络的误差处于较低值,展现了AC算法在跟踪中良好的稳定性能。

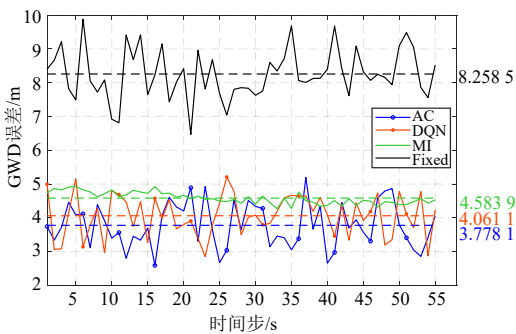


图6 GWD误差对比图

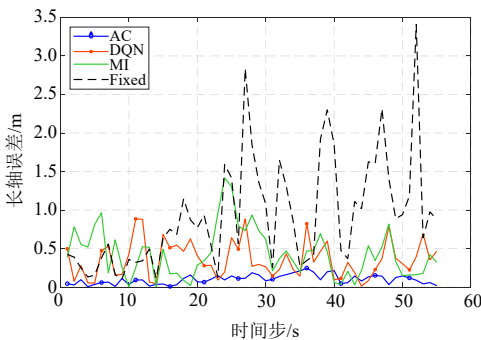


图7 长轴误差对比图

图7和图8分别展示了目标长轴和短轴的估计误差。与其他3种方法相比,基于AC网络的误差显著降低,无论目标是CV还是CT,误差比较稳定。表明了AC网络在这两项关键几何参数上的估计精度更高,验证了其在目标形态精细重建方面的能力。图9表明,在整个跟踪时间内,对于目标的朝向角度误差,AC和DQN算法始终处于较低水平,均小于 1° ,而MI和Fixed算法在初期误差值波动较大,甚至超过了 20° ,虽然后期逐渐稳定,但相比AC网络还是具有巨大的偏差,充分证明了强化学习方法对目标的朝向角度跟踪具有显著的优势。

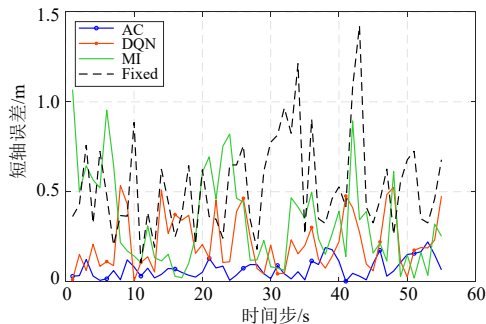


图8 短轴误差对比图

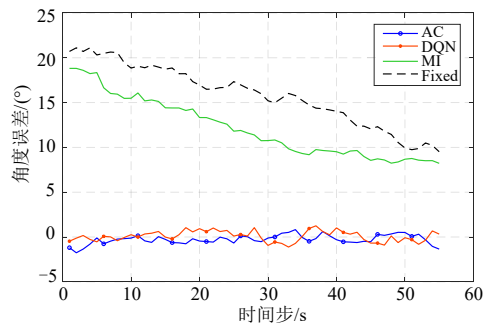


图9 角度误差对比图

4.2.3 波形参数设计

图10~图11展示了3种算法的波形设计结果。图10给出了实时交互的带宽选择结果,相较

于DQN相对平稳的带宽选择,AC网络在初期倾向选择大的带宽以探索目标的运动规律,随着时间推进,选择了相对较小的带宽以节约频谱资源,在目标状态发生改变时选择大的带宽以稳定跟踪。而MI算法更倾向随着目标状态改变迅速调整带宽参数,参数选择波动较大。

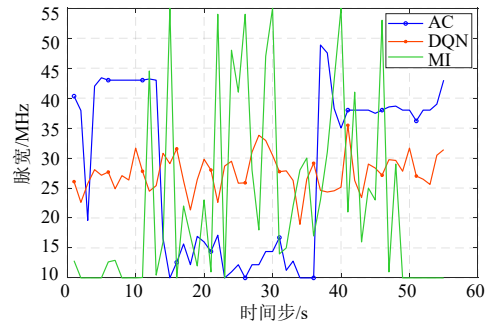


图10 带宽选择

由图11可以看出,DQN网络倾向于选择相对平稳的脉宽,MI选择的脉宽变化较为剧烈,而AC网络初期选择较大的脉宽,随着跟踪逐渐稳定,选择较小的脉宽,当运动状态发生改变时选择较大的脉宽。

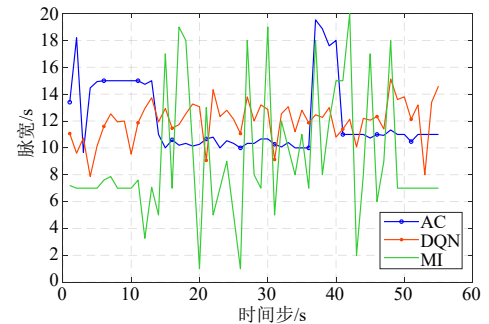


图11 脉宽选择

4.2.4 综合性能对比

表2列出了4种算法在距离、速度、形状等误差方面的详细对比情况。

表2 指标性能对比

算法	长轴平均 误差/m	短轴平均 误差/m	角度平均 误差/($^\circ$)	GWD平均 误差/m	长轴最大 误差/m	短轴最大 误差/m	角度最大 误差/($^\circ$)	GWD最大 误差/m	平均距离 误差/m	平均速度误 差/($\text{m}\cdot\text{s}^{-1}$)	最大距离 误差/m	最大速度误 差/($\text{m}\cdot\text{s}^{-1}$)
AC	0.099 5	0.361 8	0.072 9	0.216 7	0.463 3	0.486 5	1.667 2	4.061 1	0.247 0	0.886 2	0.206 3	0.533 1
DQN	0.361 8	0.216 7	0.486 5	4.061 1	0.886 2	0.533 1	1.736 7	5.434 8	3.805 6	4.744 8	5.033 9	7.077 3
MI	0.437 8	0.330 7	12.146 1	4.583 9	1.417 4	1.067 6	18.817 1	6.470 2	4.269 1	4.774 5	4.562 0	6.991 2
Fixed	0.933 4	0.528 9	15.855 4	8.258 5	3.409 6	1.422 8	21.111 5	9.886 1	7.905 2	6.687 8	9.492 5	9.840 0

从平均误差来看,AC网络在速度、距离、长短轴、GWD误差方面均优于其他算法,在角度估计误差中显著高于MI和Fixed,体现出其在目标形状与方位联合估计中的更高精度。

在最大误差方面,AC网络在所有4项指标上均取得更低数值,特别是在角度最大误差上($1.667 2^\circ$)相较于Fixed($21.111 5^\circ$),相差约 19.5° ,进一步说明其在极端情况下的估计稳定性更强,不易出现

大幅偏离。

该结果进一步验证了 AC 网络在扩展目标跟踪任务中具有更优的精度一致性和有效性。

5 结束语

随着宽带雷达的迅速发展,对扩展目标的跟踪越来越成为研究的关注点。本文针对 ETT 中需依据实时交互获得的运动与形状状态信息,在线生成兼具高精度与高稳定性的发射波形问题,提出一种基于 AC 网络架构的智能波形设计方法。首先基于 IMM-UKF-JPDA 实现扩展目标的运动状态和形状状态预测估计,然后依据 AC 网络架构的特点分别对运动状态和形状状态进行交互,并依据反馈结果调整波形参数用于下一周期的量测,从而形成一个“感知—决策—优化”的智能闭环框架。仿真实验表明,本文提出的方法可以显著降低目标形状和运动的误差,随着目标状态的变化,带宽和脉宽随之改变,这与波形设计的目标相一致,具有一定的工程应用价值。

参考文献:

- [1] HAYKIN S. Cognitive radar: a way of the future[J]. IEEE Signal Processing Magazine, 2006, 23(1): 30-40.
- [2] 崔国龙,余显祥,杨婧,等. 认知雷达波形优化设计方法综述[J]. 雷达学报, 2019, 8(5): 537-557.
- [3] BELL M R. Information theory and radar waveform design[J]. IEEE Transactions on Information Theory, 1993, 39(5): 1578-1597.
- [4] 王建涛. 面向参数估计的认知雷达自适应发射波形优化技术研究[D]. 长沙: 国防科学技术大学, 2014: 1-132.
- [5] 王海军,胡涛,张雅歌,等. 距离跟踪雷达的改进遗传算法干扰波形设计[J]. 信息工程大学学报, 2025, 26(3): 261-265.
- [6] 刘毅,朱雨,许瑾晨,等. 分布式强化学习驱动的量子编译自动调优方法[J]. 信息工程大学学报, 2025, 26(4): 462-469.
- [7] ZHU P K, LIANG J, LUO Z H, et al. Cognitive radar target tracking using intelligent waveforms based on reinforcement learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: No.5107315.
- [8] 熊芳勇. 认知雷达导引头波形选择技术研究[D]. 郑州: 信息工程大学, 2023: 1-72.
- [9] KOCH J W. Bayesian approach to extended object and cluster tracking using random matrices[J]. IEEE Transactions on Aerospace and Electronic Systems, 2008, 44(3): 1042-1059.
- [10] DING W J, ZHANG W W, ZHANG M, et al. Networked radar waveform design for detecting extended target in the presence of jamming[J]. Physical Communication, 2025, 68: No.102539.
- [11] 王文杰. 面向认知雷达目标检测的波形设计研究[D]. 西安: 西安电子科技大学, 2023: 1-90.
- [12] ZHANG H Y, CHEN H, ZHANG W X, et al. Trajectory planning for airborne radar in extended target tracking based on deep reinforcement learning[J]. Digital Signal Processing, 2024, 153: No.104603.
- [13] YANG Z R, CHEN Y X, HONG M Y, et al. Provably global convergence of actor-critic: a case for linear quadratic regulator with ergodic cost[C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems. Red Hook, USA: Curran Associates Inc., 2019: 8353-8365.
- [14] 何友,修建娟,刘瑜,等. 雷达数据处理及应用[M]. 4版. 北京: 电子工业出版社, 2022: 1-546.
- [15] 张虹芸,陈辉,张文旭. 扩展目标跟踪中基于深度强化学习的传感器管理方法[J]. 自动化学报, 2024, 50(7): 1417-1431.
- [16] 吴小丽. 用于网络化雷达目标跟踪的波形认知选择算法研究[D]. 西安: 西安电子科技大学, 2023: 1-95.
- [17] VAN TREES H L. Detection, estimation, and modulation theory, part III[M]. New York, USA: John Wiley & Sons, Inc., 2001: 1-622.
- [18] KERSHAW D J, EVANS R J. Optimal waveform selection for tracking systems[J]. IEEE Transactions on Information Theory, 1994, 40(5): 1536-1550.
- [19] HORN R A, JOHNSON C R. Matrix analysis[M]. Cambridge, UK: Cambridge University Press, 2012: 1-662.
- [20] 毛德辉,李鹏,吴建德. 基于深度强化学习的电机 S-曲线转速控制方法[J]. 机床与液压, 2025, 53(5): 88-94.
- [21] 徐如燕,鲁汉榕,郭齐胜. 基于信息论的连续属性离散化[J]. 计算机工程与设计, 2002, 23(2): 62-64.

(编辑: 冯 春)