

基于强化学习的认知雷达目标跟踪波形挑选方法

朱培坤 梁菁* 罗子涵 沈晓峰

(电子科技大学信息与通信工程学院 成都 611731)

摘要: 认知雷达通过不断与环境互动并从经验中学习, 根据获得的知识不断调整其波形、参数和照射策略, 以在复杂多变的场景中实现稳健的目标跟踪, 其波形设计在提高跟踪性能方面一直备受关注。该文提出了一种用于跟踪高机动目标的认知雷达波形选择框架, 该框架考虑了恒定速度(CV)、恒定加速度(CA)和协同转弯(CT)模型的组合, 在该框架的基础上设计了基于准则优化(CBO)和熵奖励Q学习(ERQL)方法进行最优波形选择。该方法将雷达与目标集成到一个闭环中, 发射波形随目标状态的变化实时更新, 从而达到对目标的最佳跟踪性能。数值结果表明, 与CBO方法相比, 所提出的ERQL方法大大减少了获取最优波形的处理时间, 并实现了与CBO相近的跟踪性能, 相比于固定参数(Fixed-P)方法, 极大地提高了机动目标的跟踪精度。

关键词: 目标跟踪; 认知雷达; 波形挑选; 基于准则优化(CBO); 熵奖励Q学习(ERQL)

中图分类号: TN958

文献标识码: A

文章编号: 2095-283X(2023)02-0412-13

DOI: 10.12000/JR22239

引用格式: 朱培坤, 梁菁, 罗子涵, 等. 基于强化学习的认知雷达目标跟踪波形挑选方法[J]. 雷达学报, 2023, 12(2): 412-424. doi: 10.12000/JR22239.

Reference format: ZHU Peikun, LIANG Jing, LUO Zihan, *et al.* Waveform selection method of cognitive radar target tracking based on reinforcement learning[J]. *Journal of Radars*, 2023, 12(2): 412-424. doi: 10.12000/JR22239.

Waveform Selection Method of Cognitive Radar Target Tracking Based on Reinforcement Learning

ZHU Peikun LIANG Jing* LUO Zihan SHEN Xiaofeng

(School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: Based on the obtained knowledge through ceaseless interaction with the environment and learning from the experience, cognitive radar continuously adjusts its waveform, parameters, and illumination strategies to achieve robust target tracking in complex and changing scenarios. Its waveform design has been receiving attention to improve tracking performance. In this paper, we propose a novel framework of cognitive radar waveform selection for the tracking of high-maneuvering targets. The framework considers the combination of Constant Velocity (CV), Constant Acceleration (CA), and Coordinate Turn (CT) motions. We also design Criterion-Based Optimization (CBO) and Entropy Reward Q-Learning (ERQL) methods to perform waveform selection based on this framework. To provide the optimum target tracking performance, it merges the radar and target into a closed loop, updating the broadcast waveform in real-time as the target state changes. The suggested ERQL technique achieves about the same tracking performance as the CBO while using much less processing time than the CBO, according to numerical results. The proposed ERQL method significantly increases the tracking accuracy of moving targets as compared to the fixed parameter approach.

Key words: Target tracking; Cognitive radar; Waveform selection; Criterion-Based Optimization (CBO); Entropy Reward Q-Learning (ERQL)

收稿日期: 2022-12-21; 改回日期: 2023-02-08; 网络出版: 2023-02-22

*通信作者: 梁菁 liangjing@uestc.edu.cn *Corresponding Author: LIANG Jing, liangjing@uestc.edu.cn

基金项目: 国家自然科学基金(61731006), 四川省自然科学基金(2023NSFSC0450), 111计划(B17008)

Foundation Items: The National Natural Science Foundation of China (61731006), Sichuan Natural Science Foundation (2023NSFSC0450), The 111 Project under Grant (B17008)

责任编辑: 胡卫东 Corresponding Editor: HU Weidong

1 引言

认知雷达根据环境和目标的变化情况调整其操作和处理策略,可实现比常规雷达更好的目标检测效果。这类雷达可以从自身经验中学习和进化,是下一代目标检测的重要研究方向。随着现代科学技术的不断发展,机动性更高、雷达散射面更小的目标越来越多,各种新的电磁干扰策略也应运而生。此外,雷达所在的地理工作环境复杂多变,如何在有限的能量、时间和频谱资源^[1,2]内根据目标和环境的变化合理自适应地设计发射波形、优化资源分配^[3]是现代雷达面临的挑战。Haykin^[4]最先提出认知雷达(Cognitive Radar, CR)的概念,与传统的自适应雷达相比,CR的发射机可以利用接收机反馈的环境信息、现有知识和合适的准则自适应地设计^[5]和发射波形。因此,它可以最大限度地提取非均匀环境中感兴趣目标的信息,提高雷达在动态环境中的性能^[6,7]。此后,各知名学者开始对认知雷达进行深入研究。Guerci^[8]首先提出了认知全自适应雷达的理论框架,并结合知识辅助(Knowledge Aid, KA)和自适应发射^[9]对该框架进行了改进。他最近全面概述了采用KA全自适应方法的认知雷达领域的最新发展^[10]。Bell等人^[11]提出了适用于目标探测和跟踪场景的通用认知雷达系统模型,Smith和Metron创新集团联合开发了认知雷达实验平台^[12]。这些工作在认知雷达设计理论方面具有开创性和启发性,而具体实用的认知雷达系统仍在研发路上。

一般来说,认知发射主要有两个方式^[13]来实现:最优波形设计和最优波形选择。前者是基于雷达获取的环境信息在线实时设计^[14,15],但求最优解的过程往往比较复杂,计算复杂度难以满足实时性要求。而最优波形选择^[16,17]是预先设计一个波形库,按照一定的准则从中选择一组最优的波形或波形参数。如在文献^[18]中,为了最大限度地提高雷达在恶劣环境下的运行性能,Clemente使用分数阶傅里叶变换生成相位编码库波形,并分析它们的模糊度函数以量化所提出波形库的有效性能。在频谱拥挤的环境中,Zhao等人^[19]提出了一种基于波形库的方法来实现实时波形自适应,结果表明,所提出的波形库在小尺度上具有较高的量化精度,并且生成的波形具有令人满意的频谱兼容性。在文献^[20]中,Nguyen提出了一种自适应波形选择算法,通过最小化贝叶斯克拉美罗下界(Cramer-Rao Lower Bound, CRLB)获得跟踪成本函数,选择最小化跟踪均方误差的波形进行目标跟踪。针对机动目标跟踪问题,Roman等人^[21]在自适应雷达背景下,介绍了一种互信息准则函数来选择最优波形参数,结果表明,该判据是自适应SAR系统中波形选择的有效手

段。Cao等人^[22]提出了一种基于间接强化学习的自适应波形选择算法,解决了目标状态空间的不确定性问题。结果表明自适应波形选择比传统固定参数方法具有更好的计算效率和更小的状态估计误差,也提高了跟踪精度。

现代目标形状各异、机动性强,雷达工作的电磁环境也越来越复杂。然而,传统雷达发射波形只有单一的波形,无法有效应对环境变化。此外,通过网格搜索最优波形参数的计算成本较高,且无法保证最优波形的实时性。鉴于此,我们采用IMM和自适应波形选择的思想作为基础工作,提出了一种新颖的认知雷达波形选择框架,以改善高机动性目标跟踪,如图1所示。该框架基于IMM场景结合恒定速度(Constant Velocity),恒定加速度(Constant Acceleration)和协同转弯(Coordinate Turn, CT)运动模型,将它们的滤波结果加权综合得到近似预测误差协方差。基于该波形选择框架,设计了准则优化(Criterion-Based Optimization, CBO)和熵奖励Q学习(Entropy Reward Q-Learning, ERQL)方法分别从以高斯线性调频信号为例的波形库中挑选最优发射波形参数,其中ERQL方法提出了一种基于行为奖惩的熵奖励函数。发射器-目标-接收器形成一个闭环,发射波形参数与目标状态变化实时迭代更新,以达到最佳目标跟踪性能。

2 系统模型

该部分描述了目标运动模型的融合,并分析了波形对目标跟踪的影响。

2.1 目标运动模型的融合

为简单起见,本文研究了3种典型的机动目标运动模型,包括CV, CA和CT运动模型^[23,24]。

目标模型是以下形式的离散时间动态运动模型:

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{w}_k \quad (1)$$

其中, $\mathbf{x}_k = [x_k, \dot{x}_k, \ddot{x}_k, y_k, \dot{y}_k, \ddot{y}_k]^T$, $[x_k, y_k]^T$, $[\dot{x}_k, \dot{y}_k]^T$ 和 $[\ddot{x}_k, \ddot{y}_k]^T$ 分别定义为目标的位置、速度和加速度。 \mathbf{F}_k 为状态转移矩阵,状态噪声 $\mathbf{w}_k \sim \mathcal{N}\{\mathbf{w}_k; \mathbf{0}, \mathbf{Q}_k\}$ 。我们假设非线性观测模型为

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k, k) + \mathbf{v}_k \quad (2)$$

其中, $\mathbf{h}(\cdot)$ 是量测函数,量测噪声 $\mathbf{v}_k \sim \mathcal{N}\{\mathbf{v}_k; \mathbf{0}, \mathbf{R}_k\}$ 。各运动模型的状态转移矩阵 \mathbf{F}_k 和高斯态噪声 \mathbf{w}_k 将在附录中展示。量测向量 \mathbf{z}_k 、量测函数 $\mathbf{h}(\cdot)$ 和误差协方差 \mathbf{R}_k 将会各自展示在第4节和第2节。

对于加速度波动较大的目标,单模型方法无法匹配目标的实际运动状态,跟踪效果不理想。IMM作为一种多模型方法,可以同时使用多种不同形式的运动模型和噪声来估计目标状态,然后自适应地

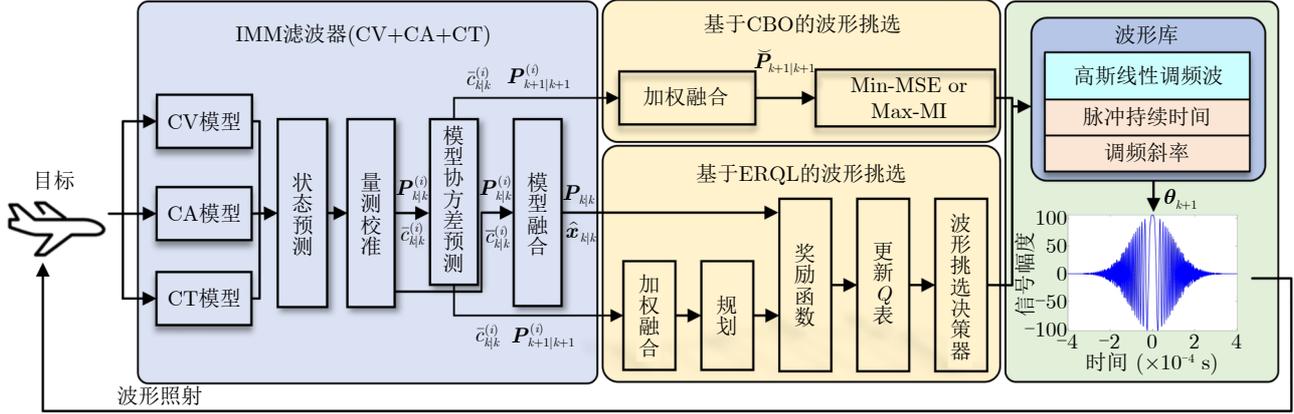


图1 认知雷达波形选择框架

Fig. 1 Cognitive radar waveform selection framework

将各个模型的估计结果折衷，加权和作为最终的目标跟踪结果。目前，它已成功应用于多个跟踪系统^[21]。本节采用CV, CA和CT模型作为IMM的模型集，算法流程如图2所示。

IMM目标跟踪算法假设模型集中的模型转移为1阶马尔可夫过程，具体算法流程如下：

首先，根据转移概率矩阵和模型的先验概率计算每个模型的预测概率：

$$\bar{c}_k^{(i)} = \sum_{j=1}^3 \pi_{ji} \mu_{k-1|k-1}^{(j)} \quad (3)$$

其中，这里 π_{ji} 表示转移概率矩阵 \mathbf{P}_π 中从模型 j 转移到模型 i 的概率， \mathbf{P}_π 将在后面给出。 $\mu_{k-1|k-1}^{(j)}$ 表示第 j 个模型在时间 $k-1$ 的有效概率。可以得到模型 j 到模型 i 的归一化概率：

$$\mu_{k-1|k-1}^{(i|j)} = \frac{\pi_{ji} \mu_{k-1|k-1}^{(j)}}{\bar{c}_k^{(i)}} \quad (4)$$

其次，经过各运动模型的状态输入交互，即根据混合估计重新初始化目标状态和协方差矩阵，如下：

$$\hat{\mathbf{x}}_{k-1|k-1}^{0(i)} = \sum_{j=1}^3 \mu_{k-1|k-1}^{(i|j)} \hat{\mathbf{x}}_{k-1|k-1}^{(j)} \quad (5)$$

$$\mathbf{P}_{k-1|k-1}^{0(i)} = \sum_{j=1}^3 \mu_{k-1|k-1}^{(i|j)} \left[\mathbf{P}_{k-1|k-1}^{(j)} + \left(\hat{\mathbf{x}}_{k-1|k-1}^{(j)} - \hat{\mathbf{x}}_{k-1|k-1}^{0(i)} \right) \left(\hat{\mathbf{x}}_{k-1|k-1}^{(j)} - \hat{\mathbf{x}}_{k-1|k-1}^{0(i)} \right)^T \right] \quad (6)$$

其中，带有帽子的状态向量 $\hat{\mathbf{x}}$ 表示对状态向量 \mathbf{x} 的估计值， $\hat{\mathbf{x}}_{k-1|k-1}^{(j)}$, $\mathbf{P}_{k-1|k-1}^{(j)}$ 分别表示第 j 个运动模型在 $k-1$ 时目标的状态向量和协方差矩阵，式(5)和式(6)分别表示图2中输入交互的结果。

然后，每个滤波器分别根据接收到的测量值

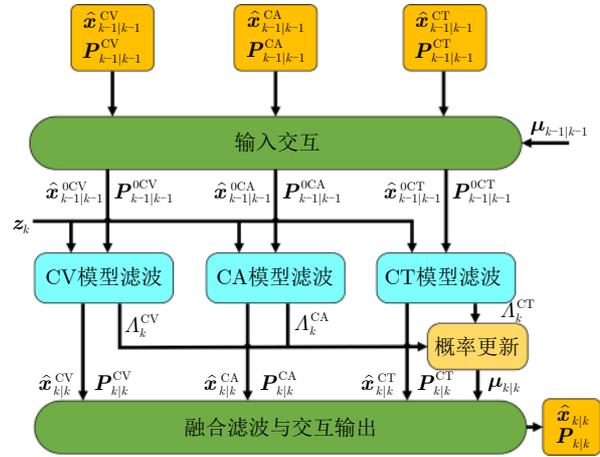


图2 以CV, CA和CT为模型的IMM流程图

Fig. 2 IMM flow chart based on CV, CA and CT models

z_k 进行卡尔曼滤波，更新各自对机动目标状态的估计。此步骤对应于图2中的CA, CV和CT模型滤波。

对于第 i 个模型，具体滤波过程如下：

(1) 根据 $k-1$ 时刻的状态估计值和模型归一化交互概率完成输入交互，如式(5)和式(6)所示。

(2) 根据运动模型，单步预测目标状态和误差协方差矩阵：

$$\hat{\mathbf{x}}_{k|k-1}^{(i)} = \mathbf{F}_{k-1}^{(i)} \hat{\mathbf{x}}_{k-1|k-1}^{0(i)} \quad (7)$$

$$\mathbf{P}_{k|k-1}^{(i)} = \mathbf{F}_{k-1}^{(i)} \mathbf{P}_{k-1|k-1}^{0(i)} (\mathbf{F}_{k-1}^{(i)})^T + \mathbf{Q}_{k-1}^{(i)} \quad (8)$$

(3) 根据 k 时刻的量测值，更新目标的状态估计：

$$\tilde{z}_k^{(i)} = z_k - \mathbf{H}_k^{(i)} \hat{\mathbf{x}}_{k|k-1}^{(i)} \quad (9)$$

$$\mathbf{S}_k^{(i)} = \mathbf{H}_k^{(i)} \mathbf{P}_{k|k-1}^{(i)} (\mathbf{H}_k^{(i)})^T + \mathbf{R}_k^{(i)} \quad (10)$$

$$\mathbf{K}_k^{(i)} = \mathbf{P}_{k|k-1}^{(i)} (\mathbf{H}_k^{(i)})^T (\mathbf{S}_k^{(i)})^{-1} \quad (11)$$

$$\hat{\mathbf{x}}_{k|k}^{(i)} = \hat{\mathbf{x}}_{k|k-1}^{(i)} + \mathbf{K}_k^{(i)} \tilde{z}_k^{(i)} \quad (12)$$

$$\mathbf{P}_{k|k}^{(i)} = [\mathbf{I} - \mathbf{K}_k^{(i)} \mathbf{H}_k^{(i)}] \mathbf{P}_{k|k-1}^{(i)} \quad (13)$$

其中, $\mathbf{H}_k^{(i)} = [\Delta_{\mathbf{x}_k} \mathbf{h}(\mathbf{x}_k, k)]$ 是量测方程 $\mathbf{h}(\mathbf{x}_k, k)$ 关于目标状态 \mathbf{x}_k 的雅可比矩阵, $\Delta_{\mathbf{x}_k}$ 表示对 \mathbf{x}_k 的1阶偏导。 $\tilde{\mathbf{z}}_k^{(i)}$, $\mathbf{S}_k^{(i)}$ 和 $\mathbf{K}_k^{(i)}$ 分别表示第 i 个滤波器的新息, 新息协方差和卡尔曼增益矩阵。

最后, 融合滤波和交互输出, 假设新息服从正态分布, 第 i 个模型的似然概率由式(14)给出:

$$\Lambda_k^{(i)} = \frac{1}{\sqrt{|2\pi\mathbf{S}_k^{(i)}|}} \exp\left[-\frac{1}{2}\left(\tilde{\mathbf{z}}_k^{(i)}\right)^\top \left(\mathbf{S}_k^{(i)}\right)^{-1} \tilde{\mathbf{z}}_k^{(i)}\right] \quad (14)$$

结合式(3)的预测概率, 更新模型 i 的有效概率:

$$\mu_{k|k}^{(i)} = \frac{\Lambda_k^{(i)} \bar{c}_k^{(i)}}{\sum_{j=1}^3 \Lambda_k^{(j)} \bar{c}_k^{(j)}} \quad (15)$$

因此, 根据每个模型的概率, 对目标状态估计和误差协方差进行加权融合, 得到IMM滤波的最终估计结果:

$$\hat{\mathbf{x}}_{k|k} = \sum_{j=1}^3 \mu_{k|k}^{(j)} \hat{\mathbf{x}}_{k|k}^{(j)} \quad (16)$$

$$\mathbf{P}_{k|k} = \sum_{j=1}^3 \mu_{k|k}^{(j)} \left[\mathbf{P}_{k|k}^{(j)} + \left(\hat{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k}^{(j)}\right) \left(\hat{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k}^{(j)}\right)^\top \right] \quad (17)$$

2.2 波形对目标跟踪的影响

本文考虑以下的雷达窄带脉冲发射信号^[25]:

$$s_T(t) = \sqrt{2}\text{Re} \left\{ \sqrt{E_T} \tilde{s}(t) e^{j2\pi f_c t} \right\} \quad (18)$$

其中, $\text{Re}(\cdot)$ 表示取实部操作, E_T 为发射信号的能量, f_c 为载频, $\tilde{s}(t)$ 为单位能量复包络信号, 且满足:

$$\int_{-T/2}^{T/2} |\tilde{s}(t)|^2 dt = 1 \quad (19)$$

其中, T 是脉冲重复周期, 那么雷达接收的单个点目标回波信号表示为

$$s_R(t) = \sqrt{2}\text{Re} \left\{ \left[\sqrt{E_R} e^{j\varphi} \tilde{s}(t - \tau) e^{j2\pi vt} + \tilde{n}(t) \right] e^{j2\pi f_c t} \right\} \quad (20)$$

其中, E_R 表示接收信号的能量, φ 为机动目标反射导致的随机相移, τ 为目标时延, v 为机动目标径向运动导致的多普勒平移, $\tilde{n}(t)$ 为接收的高斯白噪声。

在目标跟踪系统中, 当接收信号信噪比足够大, 信号模糊函数的旁瓣可忽略时, 目标时延-多普勒估计误差可以达到CRLB, 即CRLB量测噪声协方差。该值与波形参数有关, 从而建立了参数与跟踪算法之间的联系。

本文以高斯调频信号^[25]为例构建波形参数库, 其复包络的表达式为

$$\tilde{s}(t) = (\pi\lambda^2)^{-\frac{1}{4}} \exp\left[-\left(\frac{1}{2\lambda^2} - j2\pi b\right)t^2\right] \quad (21)$$

令 $a(t) = (\pi\lambda^2)^{-\frac{1}{4}} \exp(-t^2/(2\lambda^2))$, $\Omega(t) = 2\pi(bdt^2/(dt) + f_c)$, 脉冲持续时间 λ 和调频斜率 b 构成的估计矢量为 $\boldsymbol{\theta} = [\lambda, b]^\top$ 。 $s_R(t)$ 对应的时延多普勒估计 $[\tau, v]^\top$ 的费希尔信息矩阵 $\mathbf{J}(\boldsymbol{\theta})$ 的各元素通过以下求得^[26]:

$$\begin{aligned} J_{1,1} &= -\left. \frac{\partial^2 \text{AF}_{\tilde{s}}(\tau, v)}{\partial \tau^2} \right|_{\substack{\tau=0 \\ v=0}} \\ &= \int_{-\lambda/2}^{\lambda/2} (\dot{a}^2(t) + a^2(t)\Omega^2(t)) dt \\ &\quad - \left(\int_{-\lambda/2}^{\lambda/2} a^2(t)\Omega(t) dt \right)^2 \\ J_{1,2} &= -\left. \frac{\partial^2 \text{AF}_{\tilde{s}}(\tau, v)}{\partial \tau \partial v} \right|_{\substack{\tau=0 \\ v=0}} \\ &= \int_{-\lambda/2}^{\lambda/2} 2a^2(t)\Omega(t) dt \\ J_{2,2} &= -\left. \frac{\partial^2 \text{AF}_{\tilde{s}}(\tau, v)}{\partial v^2} \right|_{\substack{\tau=0 \\ v=0}} \\ &= \int_{-\lambda/2}^{\lambda/2} t^2 a^2(t) dt \end{aligned} \quad (22)$$

其中, $\text{AF}_{\tilde{s}}(\tau, v)$ 是 $\tilde{s}(t)$ 的模糊函数且 $J_{1,2} = J_{2,1}$ 。根据检测估计理论, 噪声协方差在时延多普勒域中的CRLB^[26]为 $\mathbf{J}(\boldsymbol{\theta})^{-1}$ 。结合转移矩阵 $\mathbf{T} = \text{diag}(c/2, c/(2f_c))$, 其中 c 是光速, 可将时延多普勒域的CRLB转换到径向距离 r 和径向速度 \dot{r} 域的CRLB, 对应的距离-速度测量噪声协方差的CRLB可表示为

$$\mathbf{R}_{(r,\dot{r})}(\boldsymbol{\theta}_k) = \mathbf{T} \mathbf{J}(\boldsymbol{\theta}_k)^{-1} \mathbf{T} \quad (23)$$

最后可得到高斯调频信号的量测噪声协方差的CRLB为

$$\mathbf{R}_{(r,\dot{r})}(\boldsymbol{\theta}_k) = \begin{bmatrix} c^2\lambda^2/(2\eta) & -c^2b\lambda^2/(f_c\eta) \\ -c^2b\lambda^2/(f_c\eta) & c^2(1/\lambda^2 + 2b^2\lambda^2)/(f_c^2\eta) \end{bmatrix} \quad (24)$$

其中, η 是信噪比。从式(24)可以看出, 测量噪声协方差与脉冲持续时间 λ 和线性调频频率 b 有关。因此, 可以按照一定准则合理调整波形参数 $\boldsymbol{\theta}_k = [\lambda, b]^\top$, 有效提高跟踪精度。

3 波形参数选择准则

本节给出了两个波形选择准则: Max-MI和Min-MSE, 如下所述。

3.1 最大互信息准则

在雷达领域中, 假设量测值 \mathbf{z} 和目标预测状态

\mathbf{x} 都是服从高斯分布的随机变量,那么它们构成的误差协方差矩阵为^[27]

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_x & \mathbf{P}_{xz} \\ \mathbf{P}_{zx} & \mathbf{P}_z \end{bmatrix} \quad (25)$$

量测与目标预测状态的互信息越大,雷达对目标状态估计性能就越好, \mathbf{x} 和 \mathbf{z} 的互信息可表示为

$$\begin{aligned} I(\mathbf{x}; \mathbf{z}) &= \frac{1}{2} \log \frac{|\mathbf{P}_x|}{|\mathbf{P}_x - \mathbf{P}_{xz} \mathbf{P}_z^{-1} \mathbf{P}_{xz}^\top|} \\ &= \frac{1}{2} \log \frac{|\mathbf{P}_z|}{|\mathbf{P}_z - \mathbf{P}_{xz}^\top \mathbf{P}_x^{-1} \mathbf{P}_{xz}|} \end{aligned} \quad (26)$$

由于目标状态的预测值与测量噪声相互独立,将 \mathbf{H}_k 代入式(26)可得到 \mathbf{z}_k 和 $\mathbf{x}_{k|k-1}$ 的互信息为

$$I(\mathbf{x}_{k|k-1}; \mathbf{z}_k) = \frac{1}{2} \log \frac{\det[\mathbf{R}_k(\boldsymbol{\theta}_k) + (\mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^\top)]}{\det[\mathbf{R}_k(\boldsymbol{\theta}_k)]} \quad (27)$$

由于式(27)受到波形参数影响的仅有量测噪声协方差矩阵 \mathbf{R}_k ,且 \log 是单调递增函数,所以基于最大互信息的波形参数选择准则等效为

$$\begin{aligned} \boldsymbol{\theta}_k^{\text{opt}} &= \arg \max_{\boldsymbol{\theta}_k \in \Theta} I(\mathbf{x}_{k|k-1}; \mathbf{z}_k) \\ &= \arg \max_{\boldsymbol{\theta}_k \in \Theta} \frac{\det[\mathbf{R}_k(\boldsymbol{\theta}_k) + (\mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^\top)]}{\det[\mathbf{R}_k(\boldsymbol{\theta}_k)]} \end{aligned} \quad (28)$$

其中, Θ 为所有波形参数 $\boldsymbol{\theta}$ 组成的集合。又因为目标状态的后验协方差 $\mathbf{P}_{k|k}$ 的行列式可以表示为

$$\begin{aligned} \det[\mathbf{P}_{k|k}(\boldsymbol{\theta}_k)] &= \frac{\det[\mathbf{R}_k(\boldsymbol{\theta}_k)]}{\det[(\mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^\top + \mathbf{R}_k(\boldsymbol{\theta}_k))^{-1]} \\ &\quad \cdot \det(\mathbf{P}_{k|k-1}) \end{aligned} \quad (29)$$

其中, $\mathbf{P}_{k|k-1}$ 与雷达发射波形参数 $\boldsymbol{\theta}_k$ 无关,所以式(29)等价于:

$$\boldsymbol{\theta}_k^{\text{opt}} = \arg \min_{\boldsymbol{\theta}_k \in \Theta} \det[\mathbf{P}_{k|k}(\boldsymbol{\theta}_k)] \quad (30)$$

式(27)和式(29)的详细推导见附录。因此,在最大互信息准则下,可通过寻找使得后验估计误差协方差行列式最小的波形参数作为最优参数。

3.2 最小均方差准则

最小均方误差准则^[28]是寻找使目标状态估计在每一时刻的均方根误差最小化的波形参数。其表达式为

$$\min_{\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_k, \dots\} \in \Theta} \left\{ \mathbb{E}[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}][\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}]^\top | \mathbf{z}_k \right\} \quad (31)$$

在贝叶斯框架下的滤波器,目标状态估计误差协方差具有如下形式:

$$\mathbf{P}_{k|k}(\boldsymbol{\theta}_k) = \mathbb{E}[\boldsymbol{\varepsilon}_{k|k}(\boldsymbol{\theta}_k)\boldsymbol{\varepsilon}_{k|k}(\boldsymbol{\theta}_k)^\top | \mathbf{z}_k] \quad (32)$$

其中, $\boldsymbol{\varepsilon}_{k|k}(\boldsymbol{\theta}_k) = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}_k)$ 表示目标真实状态与估计状态的差值,对式(31)两边取迹可得

$$\begin{aligned} \text{Tr}(\mathbf{P}_{k|k}(\boldsymbol{\theta}_k)) &= \mathbb{E} \left\{ [\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}_k)][\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}_k)]^\top | \mathbf{z}_k \right\} \end{aligned} \quad (33)$$

其中,式(33)的详细推导详见附录。

因此,在最小均方误差准则下的波形参数可以通过式(34)进行选择:

$$\boldsymbol{\theta}_k^{\text{opt}} = \arg \min_{\boldsymbol{\theta}_k \in \Theta} \text{Tr}[\mathbf{P}_{k|k}(\boldsymbol{\theta}_k)] \quad (34)$$

即通过寻找使得后验估计误差协方差的迹最小的波形参数作为最优参数。

4 最优波形参数选择方法的设计

强化学习^[29]和认知雷达波形参数选择具有相似的交互学习过程。因此,基于RL的雷达波形参数选择是认知雷达的研究热点。然而,在目标跟踪场景中,很难获得状态转移的先验知识,因此我们使用无模型方法来解决这个问题,Q-Learning (QL)方法是无模型方法的典型代表。在QL中,状态动作对 (s, a) 的 Q 值被定义为期望的累积折扣奖励, Q 值可以根据式(35)^[30]更新获得:

$$\begin{aligned} Q_{k+1}(s, a) &= Q_k(s, a) \\ &\quad + \alpha \left[r_k + \gamma \max_{a'} Q_k(s', a') - Q_k(s, a) \right] \end{aligned} \quad (35)$$

其中, $s \in \mathbf{S}$ 表示agent状态, $a \in \mathbf{A}$ 表示agent行为, α 为学习率, $\gamma \in [0, 1]$ 是折扣因子, r_k 是及时奖励,其具体形式稍后给出。

经过足够次数的迭代使 Q 表收敛,选择最大 Q 值的行为作为最佳决策 $\pi^*(s)$:

$$\pi^*(s) = \arg \max_{a \in \mathbf{A}} Q^*(s, a) \quad (36)$$

在单一CV, CA或CT场景下,雷达接收机贝叶斯滤波器估计目标状态,将预测误差协方差 $\mathbf{P}_{k+1|k+1}$ 反馈给雷达发射机,发射机根据指定的准则函数选择最优波形参数。

但是这种单模型波形选择策略在IMM算法中并不适用,主要原因有以下两点:

(1) IMM模型集中单个模型的有效概率并不为1,而且会随着滤波器的迭代而不断更新,不能用单模型的预测误差协方差替代目标的预测误差协方差。

(2) 在 k 时刻, $k+1$ 时刻目标量测值 \mathbf{z}_{k+1} 无法预知,由式(17)可知,IMM算法的总体目标状态估计误差协方差 $\mathbf{P}_{k+1|k+1}$ 无法获取,进而导致无法根据准则函数实现波形参数的选择。

因此,本工作为IMM模型提供了目标预测误

差协方差的近似求解方法。在 k 时刻，通过式(3)一式(6)完成模型间的输入交互，得到各个模型的预测概率 $\bar{c}_k^{(i)}$ 以及输入状态 $\mathbf{P}_{k|k}^{(i)}$ 。然后各模型根据式(8)、式(10)、式(11)和式(13)分别对目标在 $k+1$ 时刻的状态做出预测，得到目标状态预测误差协方差 $\mathbf{P}_{k+1|k+1}^{(i)}$ 。最后结合各模型的预测误差协方差 $\bar{c}_k^{(i)}$ 进行加权求和，作为IMM的估计误差协方差，如下：

$$\bar{\mathbf{P}}_{k+1|k+1} = \Gamma \sum_{i=1}^3 \bar{c}_k^{(i)} \mathbf{P}_{k+1|k+1}^{(i)} \quad (37)$$

其中， Γ 表示使估计误差协方差矩阵单位一致的加权矩阵。该方法考虑了IMM集中的所有目标模型，将每个模型的预测误差协方差与预测概率加权融合，得到IMM算法总体预测误差协方差的近似，然后根据指定的判据函数完成波形参数的选取。

所提出的波形选择方法的框图如图3所示，下面根据图3的描述总结最优波形选择算法的具体步骤。在 k 时刻，雷达发射波形参数为 θ_k 的波形，对目标进行照射，由接收机接收来自目标回波的量测值 z_k ，并交由接收机的IMM滤波器进行数据处理。IMM滤波器完成输入交互、融合滤波和交互输出后，对机动目标在 k 时刻的状态做出估计，得到 $\hat{\mathbf{x}}_{k|k}$ 和 $\mathbf{P}_{k|k}$ 。接下来，为了实现波形参数选择，IMM中的各个模型会根据自己的目标运动模型对 $k+1$ 时刻目标的状态进行预测，得到各自的预测误差协方差 $\mathbf{P}_{k+1|k+1}^{(i)}$ 。随后结合预测概率 $\bar{c}_k^{(i)}$ 通过加权的方式融

合各个模型的预测状态误差协方差得到 $\bar{\mathbf{P}}_{k+1|k+1}$ ，将其作为IMM的预测状态误差协方差的近似，反馈给雷达发射机。雷达发射机根据指定的波形参数选择方法，从高斯线性调频脉冲参数库中选择一组最优波形参数作为 $k+1$ 时刻的发射波形参数，作用于目标和环境。通过雷达和环境的不断交互，动态地调整发射波形参数，改善雷达对目标跟踪的性能。

其中，提出的波形选择算法——CBO/ERQL算法总结如表1，根据误差协方差 $\bar{\mathbf{P}}_{k+1|k+1}(\theta_{k+1})$ 完成波形和参数选择，可通过以下两种方式实现：

(1) 基于准则优化(CBO)方法

通过遍历雷达波形库的所有波形参数，根据CBO选取满足准则的波形参数作为最优发射波形参数。例如，根据Min-MSE准则，选取使 $\bar{\mathbf{P}}_{k+1|k+1}(\theta_{k+1})$ 迹最小的波形参数作为 $k+1$ 时刻的最优发射波形。

(2) 熵奖励Q学习(ERQL)算法

在与目标交互的过程中，借助Q学习试错学习机制得到一个波形参数选择决策器。实现步骤如下：

首先，根据 k 时刻的状态估计误差协方差 $\mathbf{P}_{k|k}$ 评估雷达的跟踪性能。评价标准是熵态，即

$$ES_k = \det(\mathbf{P}_{k|k}) \quad (38)$$

其中， ES_k 表示 k 时刻的熵，通过比较 $k-1$ 时刻的熵和 k 时刻的熵来奖励 k 时刻发出的波形参数 θ_k ，设计的奖励函数如下：

$$r_k = \log(1 + |ES_{k-1} - ES_k|) \text{sign}(ES_{k-1} - ES_k) \quad (39)$$

其中， $\text{sign}(\cdot)$ 是符号函数。当熵在时间 k 小于时间

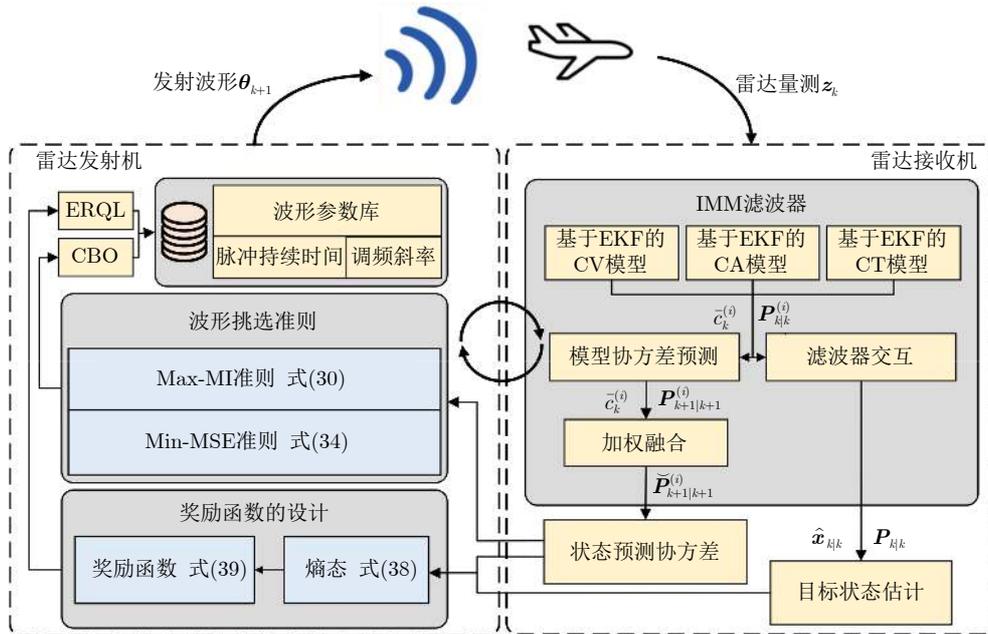


图3 波形选择框图

Fig. 3 Waveform selection block diagram

表 1 CBO/ERQL算法
Tab. 1 CBO/ERQL algorithm

输入: $k-1$ 时刻的状态估计 $\hat{\mathbf{x}}_{k-1 k-1}$, $\mathbf{P}_{k-1 k-1}$, k 时刻的量测 \mathbf{z}_k 。
输出: 最佳发射波形参数 θ_{k+1} 。
(1) 通过IMM滤波器中的交互输入和模型滤波过程, 计算每个模型在时间 k 的估计值 $\hat{\mathbf{x}}_{k k}^{CV}$, $\mathbf{P}_{k k}^{CV}$, $\hat{\mathbf{x}}_{k k}^{CA}$, $\mathbf{P}_{k k}^{CA}$, $\hat{\mathbf{x}}_{k k}^{CT}$, $\mathbf{P}_{k k}^{CT}$ 。
(2) 通过式(8)、式(10)、式(11)、式(13)计算各模型的预测概率 $\bar{c}_k^{(i)}$ 和预测状态估计误差协方差 $\mathbf{P}_{k+1 k+1}^{(i)}$ 。
(3) 通过式(37)的加权融合, 得到 $\tilde{\mathbf{P}}_{k+1 k+1}$ 。
(4) if (CBO)
(5) 通过网格搜索找到式(30)或式(34)的最优波形参数 θ_{k+1} 。
(6) else (ERQL)
(7) 根据式(38)和式(39)计算预测奖励 r_{k+1} , 通过式(35)更新每个波形的 Q 表, 重复此步骤, 直到完成所需的单步预测次数或者 Q 表收敛。
(8) 选择 Q 表中最大 Q 值所对应的策略作为 $k+1$ 时刻的波形选择策略 $\pi_{k+1}^*(s)$ 。
(9) 根据波形选择策略 $\pi_{k+1}^*(s)$ 选择波形参数 θ_{k+1} 。
(10) end if
(11) 根据波形参数 θ_{k+1} , 发射最优波形。

$k-1$ 时, 波形导致估计不确定性减小, 得到积极奖励; 否则, 给予负惩罚。

然后, 通过式(38)和式(39)计算实时奖励 r_k , 并根据式(35)更新 Q 表。

接下来, 根据式(37)单步预测 $\tilde{\mathbf{P}}_{k+1|k+1}(\theta_{k+1})$, 结合式(38)和式(39)计算预测奖励 r_{k+1} , 并根据式(35)再次更新 Q 表, 重复此步骤, 直到完成所需的单步预测次数或 Q 表收敛。

最后, 选择最大 Q 值的行为作为最佳决策, 得到 $k+1$ 时刻的最优波形选择策略 $\pi_{k+1}^*(s)$ 。

5 仿真结果

本节以一个IMM目标跟踪仿真实例为例, 验证所提方法的有效性。假设认知雷达位于坐标原点, 机动目标初始位置位于(3000, 3000) m, 初始速度大小为(0.1, 0.1) m/s, 其运动轨迹分为3个阶段, 如图4所示。机动目标运动总时长为50 s, 采样间隔 $\Delta t = 0.1$ s。在0~20 s, 目标做 10 m/s^2 的匀加速运动; 在21~40 s, 目标做转弯率为 -0.35 rad/s 的恒转弯率运动; 在41~50 s, 目标做匀速运动。雷达载频 $f_c = 10.4 \text{ GHz}$, 噪声系数 $\sigma_{CV} = \sigma_{CA} = \sigma_{CT} = 0.01$ 。

各模型初始目标状态估计 $\hat{\mathbf{x}}_{0|0}$ 、误差协方差 $\mathbf{P}_{0|0}^{(i)}$ 以及加权矩阵 $\mathbf{\Gamma}$ 分别为

$$\hat{\mathbf{x}}_{0|0} = [3010, 4, 0, 3010, 4, 0]^T \quad (40)$$

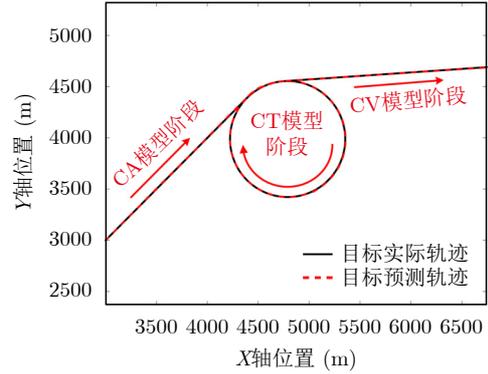


图 4 机动目标运动轨迹

Fig. 4 Trajectory of maneuvering target

$$\mathbf{P}_{0|0}^{(i)} = \text{diag}(10^2, 2^2, 0, 10^2, 2^2, 0) \quad (41)$$

$$\mathbf{\Gamma} = \text{diag}(1, 10, 125, 1, 10, 125) \quad (42)$$

目标状态的量测向量为 $\mathbf{z} = [r, \dot{r}, \beta]$, 量测函数为 $\mathbf{h}(\cdot) = [h_r(\cdot), h_{\dot{r}}(\cdot), h_{\beta}(\cdot)]^T$, 且有

$$\begin{cases} r = h_r(\mathbf{x}_k) = \sqrt{(x_k - x_r)^2 + (y_k - y_r)^2} \\ \dot{r} = h_{\dot{r}}(\mathbf{x}_k) = [x_k - x_r, y_k - y_r] [\dot{x}_k \ \dot{y}_k]^T / r \\ \beta = h_{\beta}(\mathbf{x}_k) = \arctan[(y_k - y_r) / (x_k - x_r)] \end{cases} \quad (43)$$

分别表示径向距离、径向速度和径向角度的量测值。其中, $[x_r, y_r]$ 代表雷达的位置。

角度 β 估计的CRLB是 $\sigma_{\beta} = \beta_3 \text{ dB} / \kappa \sqrt{\eta}$, 其中半功率方位波束宽度 $\beta_3 \text{ dB} = 3$, 单脉冲误差斜率 $\kappa = 1$ 。这里 β 只受信噪比 η 的影响, 与 r 和 \dot{r} 的测量无关, 则测量噪声协方差式(24)可扩展为

$$\mathbf{R}_{(r, \dot{r}, \beta)}(\theta_k) = \begin{bmatrix} c^2 \lambda^2 / (2\eta) & -c^2 b \lambda^2 / (f_c \eta) & 0 \\ -c^2 b \lambda^2 / (f_c \eta) & c^2 (1/2\lambda^2 + 2b^2 \lambda^2) / (f_c^2 \eta) & 0 \\ 0 & 0 & \sigma_{\beta}^2 \end{bmatrix} \quad (44)$$

其中, 信噪比 η 被定义为

$$\eta = R_0^4 / (R_{Tx}^2 R_{Rx}^2) \quad (45)$$

其中, $R_{Tx} = R_{Rx}$, R_0 是雷达接收目标回波信噪比为0 dB时的距离, 此时 $R_{Tx} = R_0$, 仿真实验设置 R_0 为7000 m。

本文采用高斯调频脉冲构建的波形参数库如下:

$$\mathbf{P} = \{\lambda \in [10^{-8}, 1 \times 10^{-6}] \text{ s}, b \in [-10^{12}, 10^{12}] \text{ Hz/s}\} \quad (46)$$

其中, 波形参数取值的步长设置为: $\Delta \lambda = 10^{-8} \text{ s}$, $\Delta b = 2 \times 10^{11} \text{ Hz/s}$ 。

IMM算法中各模型间的概率转移矩阵设置为

$$\mathbf{P}_\pi = \begin{bmatrix} 0.96 & 0.02 & 0.02 \\ 0.02 & 0.96 & 0.02 \\ 0.02 & 0.02 & 0.96 \end{bmatrix} \quad (47)$$

本实验在设定的情境下进行了100次蒙特卡罗仿真，为了分析不同准则函数基于IMM波形参数选择算法的性能，制定了以下3个性能指标：

(1) 目标状态估计均方根误差(Root Means Square Error, RMSE)，以目标位置为例：

$$\text{RMSE}_{\text{pos}}(k) = \sqrt{\frac{1}{N} \sum_{n=1}^N \left[(\mathbf{x}_k^n - \hat{\mathbf{x}}_{k|k}^n)^2 + (\mathbf{y}_k^n - \hat{\mathbf{y}}_{k|k}^n)^2 \right]} \quad (48)$$

(2) 目标状态估计均方根误差的均值(Average Value of Root Means Square Error, ARMSE)，以目标位置为例：

$$\text{ARMSE}_{\text{pos}} = \frac{1}{M} \sum_{n=1}^M \text{RMSE}(k) \quad (49)$$

(3) 熵态(Entropic State, ES)，熵定义为目标状态估计误差协方差的香农熵，可以用来衡量目标整体状态估计的不确定性，其表达式为

$$\text{ES}(k) = \frac{1}{N} \sum_{n=1}^N \frac{1}{2} \log [\det((2\pi\mathbf{e}) \mathbf{P}_{k|k})] \quad (50)$$

由于log函数为单调函数，所以熵态可简化为

$$\text{ES}(k) = \frac{1}{N} \sum_{n=1}^N \det(\mathbf{P}_{k|k}) \quad (51)$$

其中， n 表示第 n 次蒙特卡罗仿真， N 为蒙特卡罗仿真的次数， M 为目标跟踪过程中的采样点数。

仿真实验对所提出的ERQL方法和CBO方法进行了仿真，并采用固定波形参数(Fixed-P)方法作为跟踪性能比较，固定波形的参数是从波形库中选择的一组最佳跟踪性能参数。

在目标跟踪过程中，CA, CT和CV的有效概率变化曲线分别如图5(a)、图5(b)和图5(c)所示。如3幅图中红色虚线框所示，分别对应图4中运动轨迹的3个阶段。所提出的方法可以在目标的3个运动阶段中以最大概率选择对应的运动模型，图中，曲线ERQL-10和ERQL-40分别表示使用ERQL方法单步预测10次和40次。所提出方法与Fixed-P方法相比，可以提高模型与目标轨迹匹配的有效概率。其中，Min-MSE方法表现最好，在每个阶段都能以最高的概率匹配到正确的目标运动模型，这也是Min-MSE方法在目标跟踪中误差最小的原因之一。

限于篇幅，位置和速度的RMSE以X轴为例。图6给出了波形参数选择策略对目标位置跟踪的RMSE变化曲线。可以看出，与Fixed-P方法相比，所提出的Min-MSE, Max-MI和ERQL波形参数选择方法可以有效提高雷达目标位置的跟踪精度；Min-MSE和Max-MI方法在目标跟踪中的RMSE几乎相同。同时，还可以看到ERQL方法单步预测的次数越多，跟踪性能越好，但随着目标距离的增加，CBO的跟踪性能优于ERQL方法。

5种波形参数选择策略下目标速度的RMSE曲线如图7所示。从图中可以看出，基于CBO和ERQL方法的波形参数动态调整可以有效提高雷达目标速度估计精度。在21~40 s期间，机动目标处于CT运动阶段，速度变化较大，导致雷达目标速度估计出现较大波动。而本文提出的方法对波形参数进行动态调整，可以有效减小这种波动，提高系统跟踪的稳定性。当目标离雷达较近时，CBO和ERQL方法在目标速度估计上的误差相似，但随着目标的远离，CBO的跟踪精度优于ERQL方法。

表2显示了各波形参数选择方法的位置和速度ARMSE比较。其中， \bar{X}_{pos} 和 \bar{Y}_{pos} 分别表示X轴和Y轴

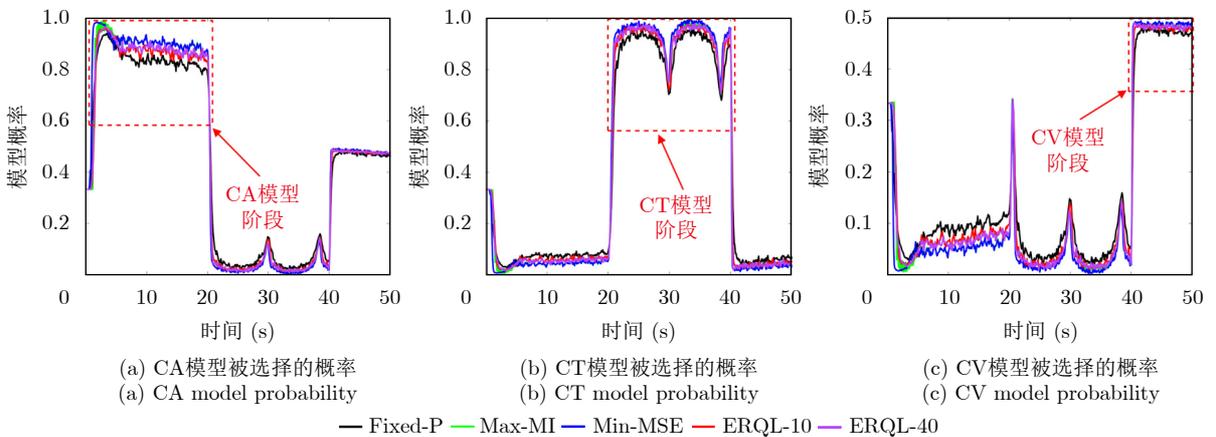


图5 各运动模型在不同运动阶段被选择的概率

Fig. 5 Probability of each motion model being selected in different motion stages

的位置跟踪ARMSE, \bar{X}_{vel} 和 \bar{Y}_{vel} 分别表示X轴和Y轴的速度跟踪ARMSE。可以看出, CBO和ERQL方法与Fixed-P方法相比, 目标跟踪性能有了显著提高。其中, Min-MSE的性能最好, Min-MSE与Fixed-P相比, X轴和Y轴位置跟踪误差分别降低了23.38%和24.04%, X轴和Y轴速度跟踪误差分别降低了47.92%和52.93%。ERQL-40与Fixed-P相比, X轴和Y轴位置跟踪误差分别降低了21.05%和22.08%, X轴和Y轴速度跟踪误差分别降低了40.63%和43.41%。表2也清楚地表明, CBO的跟踪性能优于ERQL方法, ERQL-40的跟踪性能优于ERQL-10。

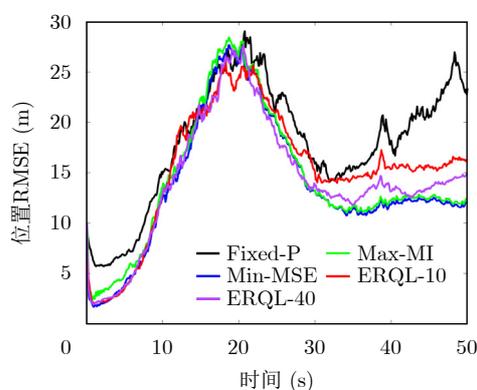


图6 目标位置跟踪RMSE曲线(X轴)

Fig. 6 Target position tracking RMSE curve (X axis)

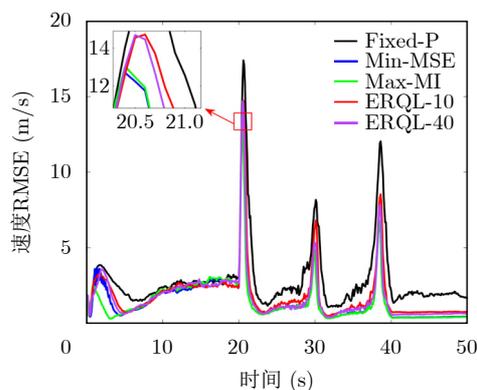


图7 目标速度跟踪RMSE曲线(X轴)

Fig. 7 Target velocity tracking RMSE curve (X axis)

表2 不同方法的ARMSE对比结果

Tab. 2 ARMSE comparison results of different methods

方法	\bar{X}_{pos}	\bar{Y}_{pos}	\bar{X}_{vel}	\bar{Y}_{vel}
Fixed-P	18.05 m	20.47 m	2.88 m/s	4.10 m/s
Min-MSE	13.83 m	15.55 m	1.50 m/s	1.93 m/s
Max-MI	14.44 m	15.79 m	1.46 m/s	1.92 m/s
ERQL-10	15.40 m	17.98 m	1.87 m/s	2.55 m/s
ERQL-40	14.25 m	15.95 m	1.71 m/s	2.32 m/s

图8和图9是目标跟踪波形参数变化曲线结果。从图8可以看出, CBO更愿意选择较大的脉冲持续时间, 在运动轨迹变换时, 选择较小的脉冲持续时间以提高雷达对目标位置的估计精度。但是ERQL方法的脉冲持续时间在整个目标跟踪期间变化不大, 只是在CT运动阶段脉冲持续时间有较大的波动来应对目标的变化。对于调频斜率的变化, 从图9可以看出, 在整个运动过程中, Max-MI总是选择较大的调频斜率, 而ERQL方法选择较小的调频斜率, 两种方法在整个目标运动过程中调频斜率变化不大。而Min-MSE方法在目标改变运动轨迹时, 线性调频率有较大变化, 其也是Min-MSE的跟踪性能最好的原因。

熵态变化曲线如图10所示, 从整体熵态变化的角度, 比较了各种方法下的雷达目标的跟踪性能。可以看到在自适应调整波形参数后, CBO和ERQL方法与Fixed-P方法相比显著降低了熵态大小。不同波形挑选方法的熵态变化规律类似于图6, 即在目标距离雷达较近时, CBO和ERQL-40方法的熵态没有明显区别, 但随着目标的远离, CBO方法的跟踪性能更好。

由上面分析可知, CBO方法的目标跟踪精度

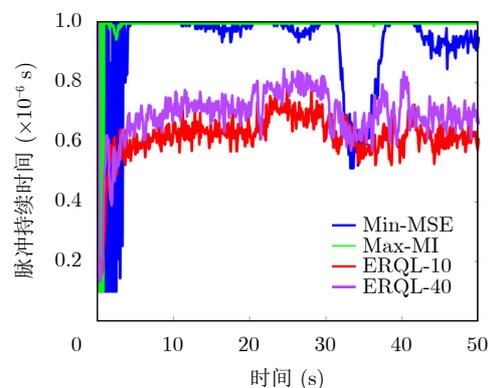


图8 目标跟踪脉冲持续时间变化曲线

Fig. 8 Target tracking pulse duration variation curve

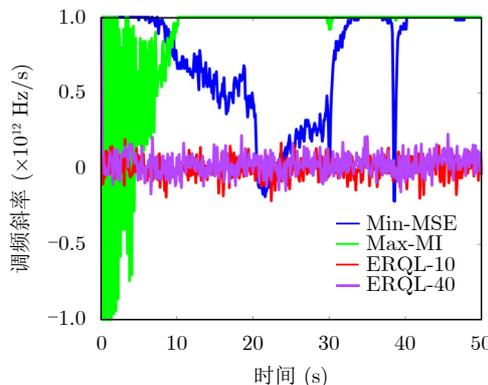


图9 目标跟踪调频斜率变化曲线

Fig. 9 Target tracking frequency modulation slope variation curve

最高，ERQL方法次之。但CBO方法是通过网格搜索实现的。它会遍历波形库中的所有参数来寻找最优的波形参数，所以这种方法非常耗时。图11显示了各种波形参数选择算法的平均耗时结果。可以看到，在相同的跟踪条件下，CBO方法所需的CPU时间约为ERQL-10方法的22倍，ERQL-40方法的7倍，Fixed-P方法的84倍。

为进一步验证ERQL方法的性能，又分别进行了单步预测多次的ERQL实验，实验对比结果如表3所示。以X轴位置跟踪为例，虽然Min-MSE方法相比Fixed-P方法的跟踪精度更高，但是Min-MSE方法的CPU时间是Fixed-P方法的84倍，而ERQL-40方法仅为Fixed-P方法的11.8倍。为了获得较高的跟踪精度，CBO遍历参数库的时间成本是非常巨大的。然而，ERQL方法通过单步预测与更新，不仅保证了目标跟踪性能，而且显著减少了波形参数选择过程的计算时间。表3还列出了各种预测次数的ERQL方法对目标跟踪性能的改进。可以看出，当需要更高的跟踪性能并考虑时间成本时，ERQL-40是最佳选择。

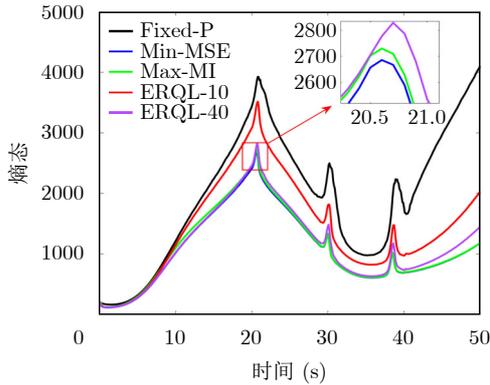


图 10 目标跟踪熵态变化曲线

Fig. 10 Target tracking entropy state variation curve

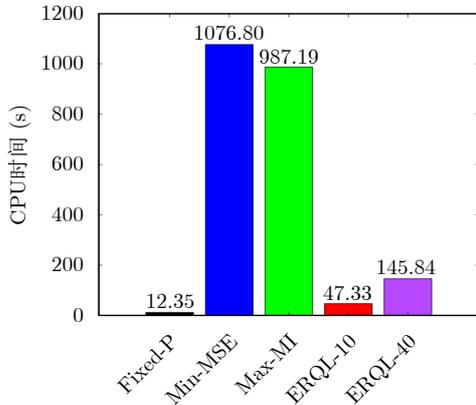


图 11 各波形参数选择算法的平均耗时结果

Fig. 11 The average time-consuming results of each waveform parameter selection algorithm

表 3 CBO和ERQL方法相比于Fixed-P方法的跟踪性能改善与CPU时间比较(%)

方法	X_{pos}	Y_{pos}	X_{vel}	Y_{vel}	CPU time
Min-MSE	23.38	24.04	47.92	52.93	8619
Max-MI	20.61	22.86	49.13	53.17	7893
ERQL-10	14.68	12.16	34.84	37.80	283
ERQL-20	16.01	16.76	37.28	40.73	545
ERQL-40	21.05	22.08	40.63	43.41	1081
ERQL-80	15.51	15.68	41.11	47.07	2016

6 结语

为了提高机动目标的跟踪精度，本文提出了一种认知雷达波形选择框架。并基于该框架设计了基于准则优化(CBO)或熵奖励Q学习(ERQL)方法以迭代智能地挑选波形参数，将雷达发射机-目标-雷达接收机集成一个闭环结构。结果表明，CBO方法虽然具有较高的跟踪精度，但其计算时间约为固定参数(Fixed-P)方法的84倍。当对跟踪性能和成本有更高要求时，单步预测40次的ERQL方法是最好的选择。与Fixed-P方法相比，ERQL-40比CBO节省约71.8倍的时间，对目标位置和速度的跟踪精度分别提高了21.05%和40.63%。未来的工作将改进强化学习和卡尔曼滤波器，以进一步提高跟踪性能。

附录

CV, CA, CT的状态转移矩阵 F_k 分别为

$$F_k^{\text{CV}} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, F_k^{\text{CA}} = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix},$$

$$F_k^{\text{CT}} = \begin{bmatrix} 1 & \frac{\sin \omega T}{\omega} & 0 & -\frac{1 - \cos \omega T}{\omega} \\ 0 & \cos \omega T & 0 & -\sin \omega T \\ 0 & \frac{1 - \cos \omega T}{\omega} & 1 & \frac{\sin \omega T}{\omega} \\ 0 & \sin \omega T & 0 & \cos \omega T \end{bmatrix}$$

其中， T 为测量数据的采样周期。

CV, CA, CT的高斯态噪声 w_k 的协方差 Q_k 分别为

$$Q_k^{\text{CV}} = \sigma_{\text{CV}}^2 \begin{bmatrix} T^2/3 & T^3/2 \\ T^3/2 & T^2 \end{bmatrix},$$

$$Q_k^{\text{CA}} = \sigma_{\text{CA}}^2 \begin{bmatrix} T^2/20 & T^2/8 & T^3/6 \\ T^4/8 & T^3/6 & T^2/2 \\ T^3/6 & T^2/2 & T \end{bmatrix},$$

$$\mathbf{Q}_k^{\text{CT}} = \sigma_{\text{CT}}^2 \begin{bmatrix} \frac{2(\omega T - \sin \omega T)}{\omega^3} & \frac{1 - \cos \omega T}{\omega^2} & 0 & \frac{\omega T - \sin \omega T}{\omega^2} \\ \frac{1 - \cos \omega T}{\omega^2} & T & -\frac{\omega T - \sin \omega T}{\omega^2} & 0 \\ 0 & -\frac{\omega T - \sin \omega T}{\omega^2} & \frac{2(\omega T - \sin \omega T)}{\omega^3} & \frac{1 - \cos \omega T}{\omega^2} \\ \frac{\omega T - \sin \omega T}{\omega^2} & 0 & \frac{1 - \cos \omega T}{\omega^2} & T \end{bmatrix}$$

其中, ω 表示转向角速度, σ_{CV}^2 , σ_{CA}^2 , σ_{CT}^2 为可以调节的状态噪声强度。

式(27)的详细推导为

$$\begin{aligned} I(\mathbf{x}_{k|k-1}; \mathbf{z}_k) &= \frac{1}{2} \log \frac{|\mathbf{P}_z|}{|\mathbf{P}_z - \mathbf{P}_{xz}^T \mathbf{P}_x^{-1} \mathbf{P}_{xz}|} \\ &= \frac{1}{2} \log \frac{|\mathbf{P}_{k|k-1}|}{|\mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}^T (\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k)^{-1} \mathbf{H} \mathbf{P}_{k|k-1}|} \\ &= \frac{1}{2} \log \frac{|\mathbf{P}_{k|k-1}|}{|\mathbf{H}^{-1} [(\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k) - \mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T] (\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k)^{-1} \mathbf{H} \mathbf{P}_{k|k-1}|} \\ &= \frac{1}{2} \log \frac{|\mathbf{P}_{k|k-1}|}{|\mathbf{H}^{-1} \mathbf{R}_k (\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k)^{-1} \mathbf{H} \mathbf{P}_{k|k-1}|} \\ &= \frac{1}{2} \log \frac{|\mathbf{I}|}{|\mathbf{H}^{-1} \mathbf{R}_k (\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k)^{-1} \mathbf{H}|} \\ &= \frac{1}{2} \log \frac{|\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k|}{|\mathbf{R}_k|} = \frac{1}{2} \log \frac{\det [\mathbf{R}_k(\boldsymbol{\theta}_k) + (\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T)]}{\det [\mathbf{R}_k(\boldsymbol{\theta}_k)]} \end{aligned} \quad (\text{A-1})$$

式(29)的详细推导为

$$\begin{aligned} \det [\mathbf{P}_{k|k}(\boldsymbol{\theta}_k)] &= \det [(\mathbf{I} - \mathbf{K}_k(\boldsymbol{\theta}_k) \mathbf{H}) \mathbf{P}_{k|k-1}] \\ &= \det \left\{ \left[\mathbf{I} - \mathbf{P}_{k|k-1} \mathbf{H}^T (\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k(\boldsymbol{\theta}_k))^{-1} \mathbf{H} \right] \mathbf{P}_{k|k-1} \right\} \\ &= \det \left\{ \left[\mathbf{H}^{-1} (\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k(\boldsymbol{\theta}_k)) - \mathbf{P}_{k|k-1} \mathbf{H}^T \right] \cdot \left[\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k(\boldsymbol{\theta}_k) \right]^{-1} \mathbf{H} \mathbf{P}_{k|k-1} \right\} \\ &= \det \left[\mathbf{H}^{-1} \mathbf{R}_k(\boldsymbol{\theta}_k) (\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k(\boldsymbol{\theta}_k))^{-1} \mathbf{H} \mathbf{P}_{k|k-1} \right] \\ &= \frac{\det [\mathbf{R}_k(\boldsymbol{\theta}_k)]}{\det [(\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k(\boldsymbol{\theta}_k))^{-1}]} \det (\mathbf{P}_{k|k-1}) \end{aligned} \quad (\text{A-2})$$

式(33)的详细推导为

$$\begin{aligned} \text{Tr} (\mathbf{P}_{k|k}(\boldsymbol{\theta}_k)) &= \text{Tr} \left\{ \mathbf{E} \left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}_k)) (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}_k))^T | \mathbf{z}_k \right] \right\} \\ &= \mathbf{E} \left\{ \text{Tr} \left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}_k)) (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}_k))^T | \mathbf{z}_k \right] \right\} \\ &= \mathbf{E} \left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}_k)) (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}_k))^T | \mathbf{z}_k \right] \\ &= \mathbf{E} \left[\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}_k)\|^2 | \mathbf{z}_k \right] \end{aligned} \quad (\text{A-3})$$

参 考 文 献

- [1] YUAN Ye, YI Wei, HOSEINNEZHAD R, *et al.* Robust power allocation for resource-aware multi-target tracking

with colocated MIMO radars[J]. *IEEE Transactions on Signal Processing*, 2021, 69: 443–458. doi: [10.1109/TSP.2020.3047519](https://doi.org/10.1109/TSP.2020.3047519).

- [2] SUN Zhichao, YEN G G, WU Junjie, *et al.* Mission planning for energy-efficient passive UAV radar imaging system based on substage division collaborative search[J]. *IEEE Transactions on Cybernetics*, 2023, 53(1): 275–288. doi: [10.1109/TCYB.2021.3090662](https://doi.org/10.1109/TCYB.2021.3090662).
- [3] LIANG Jing and LIANG Qilian. Design and analysis of distributed radar sensor networks[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2011, 22(11): 1926–1933. doi: [10.1109/TPDS.2011.45](https://doi.org/10.1109/TPDS.2011.45).
- [4] HAYKIN S. Cognitive radar: A way of the future[J]. *IEEE Signal Processing Magazine*, 2006, 23(1): 30–40. doi: [10.1109/MSP.2006.1593335](https://doi.org/10.1109/MSP.2006.1593335).
- [5] LUO Zihan, LIANG Jing, and XU Zekai. Intelligent waveform optimization for target tracking in radar sensor networks[C]. 10th International Conference on Communications, Signal Processing, and Systems (CSPS), Changbaishan, China, 2021: 165–172. doi: [10.1007/978-981-19-0386-1_21](https://doi.org/10.1007/978-981-19-0386-1_21).
- [6] HAYKIN S. Cognition is the key to the next generation of radar systems[C]. 2009 IEEE 13th Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop, Marco Island, USA, 2009: 463–467. doi: [10.1109/DSP.2009.4785968](https://doi.org/10.1109/DSP.2009.4785968).
- [7] HAYKIN S, ZIA A, ARASARATNAM I, *et al.* Cognitive tracking radar[C]. 2010 IEEE Radar Conference, Arlington, USA, 2010: 1467–1470. doi: [10.1109/RADAR.2010.5494383](https://doi.org/10.1109/RADAR.2010.5494383).
- [8] GUERCI J R. Cognitive radar: A knowledge-aided fully adaptive approach[C]. 2010 IEEE Radar Conference, Arlington, USA, 2010: 1365–1370. doi: [10.1109/RADAR.2010.5494403](https://doi.org/10.1109/RADAR.2010.5494403).
- [9] GUERCI J R, GUERCI R M, RANAGASWAMY M, *et al.* CoFAR: Cognitive fully adaptive radar[C]. 2014 IEEE Radar Conference, Cincinnati, USA, 2014: 984–989. doi: [10.1109/RADAR.2014.6875736](https://doi.org/10.1109/RADAR.2014.6875736).
- [10] GUERCI J R. Cognitive Radar: The Knowledge-Aided Fully Adaptive Approach[M]. 2nd ed. Norwood, USA: Artech House, 2020.
- [11] BELL K L, BAKER C J, SMITH G E, *et al.* Cognitive radar framework for target detection and tracking[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2015, 9(8): 1427–1439. doi: [10.1109/JSTSP.2015.2465304](https://doi.org/10.1109/JSTSP.2015.2465304).
- [12] SMITH G E, CAMMENGA Z, MITCHELL A, *et al.* Experiments with cognitive radar[C]. 2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Cancun, Mexico, 2015: 293–296. doi: [10.1109/CAMSAP.2015.7383794](https://doi.org/10.1109/CAMSAP.2015.7383794).
- [13] ZHANG Lingzhao and JIANG Min. Cognitive radar target tracking algorithm based on waveform selection[C]. 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 2021: 1506–1510. doi: [10.1109/IAEAC50856.2021.9390757](https://doi.org/10.1109/IAEAC50856.2021.9390757).
- [14] HULEIHEL W, TABRIKIAN J, and SHAVIT R. Optimal adaptive waveform design for cognitive MIMO radar[J]. *IEEE Transactions on Signal Processing*, 2013, 61(20): 5075–5089. doi: [10.1109/TSP.2013.2269045](https://doi.org/10.1109/TSP.2013.2269045).
- [15] ALDAYEL O, MONGA V, and RANGASWAMY M. Successive QCQP refinement for MIMO radar waveform design under practical constraints[J]. *IEEE Transactions on Signal Processing*, 2016, 64(14): 3760–3774. doi: [10.1109/TSP.2016.2552501](https://doi.org/10.1109/TSP.2016.2552501).
- [16] FENG Shuo and HAYKIN S. Cognitive risk control for transmit-waveform selection in vehicular radar systems[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(10): 9542–9556. doi: [10.1109/TVT.2018.2857718](https://doi.org/10.1109/TVT.2018.2857718).
- [17] SAVAGE C O and MORAN B. Waveform selection for maneuvering targets within an IMM framework[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2007, 43(3): 1205–1214. doi: [10.1109/TAES.2007.4383612](https://doi.org/10.1109/TAES.2007.4383612).
- [18] CLEMENTE C, SHOROKHOV I, PROUDLER I, *et al.* Radar waveform libraries using fractional Fourier transform[C]. 2014 IEEE Radar Conference, Cincinnati, USA, 2014: 855–858. doi: [10.1109/RADAR.2014.6875710](https://doi.org/10.1109/RADAR.2014.6875710).
- [19] ZHAO Dehua, WEI Yinsheng, and LIU Yongtan. Real-time waveform adaptation in spectral crowded environment using a sub-waveforms-based library[C]. 2016 CIE International Conference on Radar, Guangzhou, China, 2016: 1–5. doi: [10.1109/RADAR.2016.8059310](https://doi.org/10.1109/RADAR.2016.8059310).
- [20] NGUYEN N H, DOGANCA Y K, and DAVIS L M. Adaptive waveform selection for multistatic target tracking[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2015, 51(1): 688–701. doi: [10.1109/TAES.2014.130723](https://doi.org/10.1109/TAES.2014.130723).
- [21] ROMAN J. R., GARNHAM J. W. and ANTONIK P., Information Theoretic Criterion for Waveform Selection. Fourth IEEE Workshop on Sensor Array and Multichannel Processing, 2006., Waltham, MA, USA, 2006, 444–448, doi: [10.1109/SAM.2006.1706172](https://doi.org/10.1109/SAM.2006.1706172).
- [22] CAO Xin, ZHENG Zhe, and AN Di. Adaptive waveform selection algorithm based on reinforcement learning for cognitive radar[C]. 2019 IEEE 2nd International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), Shenyang, China, 2019: 208–213. doi: [10.1109/AUTEEE48671.2019.9033413](https://doi.org/10.1109/AUTEEE48671.2019.9033413).
- [23] HAN Bo, HUANG Hanqiao, LEI Lei, *et al.* An improved IMM algorithm based on STSRCKF for maneuvering target tracking[J]. *IEEE Access*, 2019, 7: 57795–57804. doi: [10.1109/ACCESS.2019.2912983](https://doi.org/10.1109/ACCESS.2019.2912983).
- [24] BLACKMAN S S, DEMPSTER R J, BUSCH M T, *et al.*

- IMM/MHT solution to radar benchmark tracking problem[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 1999, 35(2): 730–738. doi: [10.1109/7.766953](https://doi.org/10.1109/7.766953).
- [25] KERSHAW D J and EVANS R J. Optimal waveform selection for tracking systems[J]. *IEEE Transactions on Information Theory*, 1994, 40(5): 1536–1550. doi: [10.1109/18.333866](https://doi.org/10.1109/18.333866).
- [26] SIRA S P, PAPANDREOU-SUPPAPPOLA A, and MORRELL D. *Advances in Waveform-Agile Sensing for Tracking*[M]. Cham: Springer, 2009: 59–60.
- [27] WILLIAMS J L. Information theoretic sensor management[D]. [Ph. D. dissertation], Massachusetts Institute of Technology, 2007: 41–42.
- [28] ATHANS M and TSE E. A direct derivation of the optimal linear filter using the maximum principle[J]. *IEEE Transactions on Automatic Control*, 1967, 12(6): 690–698. doi: [10.1109/TAC.1967.1098732](https://doi.org/10.1109/TAC.1967.1098732).
- [29] THORNTON C E, KOZY M A, BUEHRER R M, *et al.* Deep reinforcement learning control for radar detection and tracking in congested spectral environments[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2020, 6(4): 1335–1349. doi: [10.1109/TCCN.2020.3019605](https://doi.org/10.1109/TCCN.2020.3019605).
- [30] WANG Qing, QIAO Yanming, and GAO Lirong. A cognitive radar waveform optimization approach based on deep reinforcement learning[C]. 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Chongqing, China, 2019: 1–6. doi: [10.1109/ICSIDP47821.2019.9173348](https://doi.org/10.1109/ICSIDP47821.2019.9173348).

作者简介

朱培坤，博士生，主要研究方向包括雷达波形设计、雷达传感器网络和分布式协同信号处理等。

梁 菁，教授，博士生导师，主要研究方向包括雷达传感器网络、分布式协同信号处理、模糊逻辑与机器学习等。

罗子涵，硕士生，主要研究方向包括雷达波形设计、机器学习 and 智能信号处理。

沈晓峰，研究员，主要研究方向包括雷达探测与目标识别、智能感知与信息系统、先进信号与信息处理。

(责任编辑：于青)