

杂波先验数据缺失条件下基于级联优化处理的雷达波形设计方法

张应奎^① 孙国皓^{*①} 钟苏川^① 余显祥^②

^①(四川大学空天科学与工程学院 成都 610207)

^②(电子科技大学信息与通信工程学院 成都 611731)

摘要: 认知雷达波形设计往往依赖于精准的杂波先验信息, 当先验信息数据存在缺失时, 所构建的杂波模型会严重失配, 进而影响雷达对杂波的抑制能力。该文针对杂波先验数据缺失条件下的雷达波形优化问题, 建立完全随机缺失机制下的点状与块状缺失场景, 设计恒模与相似性约束的波形优化模型, 提出基于优先级填充-强化学习级联优化的雷达波形训练算法: 即采用强化学习智能体与填充算法修复后的杂波环境相交互的级联方法, 以最大化信杂噪比为优化目标, 通过迭代训练得到雷达最佳波形参数配置策略。最后, 仿真验证不同缺失概率条件下所提算法的优越性。结果表明: 相比于传统非级联优化算法, 该文所提算法均可获得更优的杂波抑制性能, 有效提升雷达的探测能力。

关键词: 波形设计; 杂波抑制; 先验数据缺失; 优先级填充; 强化学习; 级联优化

中图分类号: TN958

文献标识码: A

文章编号: 2095-283X(2023)01-0235-12

DOI: 10.12000/JR22166

引用格式: 张应奎, 孙国皓, 钟苏川, 等. 杂波先验数据缺失条件下基于级联优化处理的雷达波形设计方法[J]. 雷达学报, 2023, 12(1): 235–246. doi: 10.12000/JR22166.

Reference format: ZHANG Yingkui, SUN Guohao, ZHONG Suchuan, *et al.* Radar waveform design method based on cascade optimization processing under missing clutter prior data[J]. *Journal of Radars*, 2023, 12(1): 235–246. doi: 10.12000/JR22166.

Radar Waveform Design Method Based on Cascade Optimization Processing under Missing Clutter Prior Data

ZHANG Yingkui^① SUN Guohao^{*①} ZHONG Suchuan^① YU Xianxiang^②

^①(School of Aeronautics and Astronautics, Sichuan University, Chengdu 610207, China)

^②(School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: Cognitive radar waveform design often relies on accurate clutter prior information. When prior information data is missing, the constructed clutter model will be severely mismatched, affecting the radar's ability to suppress clutter. Aiming at the radar waveform optimization problem under missing clutter prior data, this paper establishes point and block-like missing scenarios under the completely random missing mechanism, designs a waveform optimization model with constant modulus and similarity constraints, and proposes a radar waveform training algorithm based on priority filling-reinforcement learning cascade optimization: that is, a cascade method in which the reinforcement learning agent interacts with the clutter environment repaired by a filling algorithm, with the optimization goal of maximizing the signal-to-noise ratio, and the optimal configuration strategy with waveform parameters is obtained through iterative training. Finally, simulations verify the superiority of the proposed algorithm under different missing probability

收稿日期: 2022-08-09; 改回日期: 2022-10-24; 网络出版: 2022-11-03

*通信作者: 孙国皓 sghsjw2005@126.com *Corresponding Author: SUN Guohao, sghsjw2005@126.com

基金项目: 国家自然科学基金(62201371), 四川省自然科学基金(2022NSFSC1952)

Foundation Items: The National Natural Science Foundation of China (62201371), Sichuan Provincial Natural Science Foundation (2022NSFSC1952)

责任编辑: 唐波 Corresponding Editor: TANG Bo

conditions. The results show that the proposed algorithm outperforms the traditional non-cascading optimization algorithm, regarding clutter suppression and effectively improves the detection ability of radar.

Key words: Waveform design; Clutter suppression; Missing clutter prior data; Priority filling; Reinforcement learning; Cascade optimization

1 引言

在对杂波环境信息感知的基础上, 认知雷达可通过主动发射波形的方式, 充分挖掘目标与杂波信息之间的差异性, 并通过优化配置波形参数扩大该差异, 提高杂波抑制和目标检测性能。相较于传统雷达, 认知雷达具有更高的自由度和灵活度, 现已成为雷达杂波抑制的主要技术途径之一。

现有认知雷达波形设计主要利用精确完整的杂波先验数据, 形成两大类准则的设计方法。一是最大化输出信杂噪比(Signal to Clutter plus Noise Ratio, SCNR)准则, 如Tang等人^[1,2]利用精细化的杂波先验数据, 基于松弛方法解决了杂波环境下MIMO (Multiple Inputs Multiple Outputs)雷达波形与滤波器联合设计中的非凸优化问题。为了高效求解上述非凸性难题, Yu等人^[3]提出了基于ADMM (Alternating Direction Method of Multipliers)的MIMO雷达波形多约束优化快速实现方法。在此基础上, Wu等人^[4]研究了基于MM (Majorization-Minimization)的加速算法, 有效提升了发射波形与接收滤波器权联合设计的杂波抑制性能。此外, O'Rourke等人^[5]采用二次优化和半正定松弛方法提升了收发联合设计的目标检测性能。然而上述方法所采用的先验数据需要精确已知各个杂波块的幅相特征, 当杂波先验数据存在缺失时, 所构建的杂波模型会严重失配, 进而影响雷达对杂波的抑制能力。二是基于信息量的优化准则, 通过从回波信号中最大化提取目标信息量, 提升目标检测性能。如Tang等人^[6]研究了杂波环境下基于信息相关熵的MIMO雷达波形设计方法; 何子述团队^[7,8]针对机载雷达杂波环境, 依据最大化互信息量(Mutual Information, MI)准则, 提出了多约束条件下的波形与STAP (Space-Time Adaptive Processing)滤波器联合设计方法。但该类方法仍然依赖于精确完整的先验杂波协方差矩阵, 在数据缺失条件下难以设计出优效的雷达波形。

从上述研究动态可以看出, 现有波形优化方法往往需要精细到每个杂波片的幅频相信息^[9,10]。因此, 在数据信息不完整/缺失条件下, 上述方法存在模型失配效应, 会严重削弱雷达的杂波抑制性能。

信息缺失在雷达信号处理中为常见现象^[11-13], 主要原因包括:

(1) 当可用通道数小于阵列个数时, 雷达通常采用切换器切换数据接收通道。一般情况下, 该切换机制是随机选取的, 从数学角度上看, 这无异于将完整数据“穿孔”, 得到一个尺寸缩小的不完整向量^[11]。

(2) 先验数据信息在采集时, 易受到气候、采集设备和其他电磁频谱设备等干扰的影响, 导致采集到的数据与真实数据之间存在偏差; 用户在使用时需将干扰数据剔除掉, 而剔除过程会造成信息的缺失。

(3) 故障传感器通道内的噪声脉冲、模数转换器故障等导致部分阵列达到饱和状态, 从而造成数据缺失^[14]; 雷达材料、环境、机械故障等情况亦有可能引起数据的失效^[15]。

现有针对数据缺失的雷达信号处理方法主要集中在协方差矩阵估计、波束形成、目标检测等领域^[16-18]。针对数据不完整条件下高维参数的优化问题, 主要有两种解决途径: 一种是补全雷达数据缺失部分^[19], 该类方法在补全过程中通过添加专家的合理性判断, 在一定程度上保证其可信性。如Aubry等人^[11]基于杂波协方差矩阵的结构信息, 采用Expectation-Maximization算法设计了数据缺失条件下的参数估计方法; Pavez等人^[13]针对数据缺失分布的不同机制, 提出了一种无偏协方差估计器, 获得不同缺失条件下的估计误差界限; Liu等人^[20]针对单调数据缺失模式研究了基于最大似然和正则化鲁棒的估计器, 提高了算法的收敛速度和估计精度; 此外, 刘宏伟团队^[21]基于迁移学习算法, 完成了对步进频雷达不完整波形数据的补全。另一种解决途径是通过与复杂环境的交互训练, 直接得到雷达参数的优化策略, 即端到端的波形训练方法^[21,22]。例如Jiang等人^[23]在多约束条件下, 研究了基于端到端的波形与检测器权值联合优化方法, 然而该训练过程依赖于大量完整的数据样本, 尚未考虑数据缺失条件下的优化问题。综上所述, 现有工作多集中在数据缺失参数估计问题上, 鲜有涉及数据缺失条件下波形设计方法的研究。

因此, 本文主要针对完全随机缺失机制下的杂波数据, 基于强化学习对未知环境的高效探索和学习能力, 提出优先级填充-强化学习级联优化的雷达波形训练方法。该方法将缺失数据恢复与波形设

计有机结合，通过智能体与杂波环境的不断交互，寻求波形参数的最优配置。其难点主要体现为：一是数据缺失条件下的波形设计建模方法；二是雷达波形设计中非凸问题的求解。针对上述难点，本文主要工作与创新点体现在以下几个方面：

(1) 本文建立了数据缺失条件下的波形设计问题模型，提出了基于优先级填充-强化学习两级级联的波形优化求解方法；该方法通过强化学习与填充算法所构建的环境交互训练，决策输出最大信杂噪比下的波形参数；

(2) 本文考虑了完全随机缺失机制，即数据的缺失与否与数据自身值无关，并通过仿真分别验证了点状缺失与块状缺失场景下(类比文献[11]中的缺失场景)所提算法杂波抑制能力；

(3) 本文所提算法可实现恒模、相似性约束下的波形优化。

下面分别从杂波数据缺失条件下波形设计模型、雷达波形级联优化算法，以及算法仿真等方面展开介绍。

2 杂波数据缺失条件下的波形设计问题模型

本节首先根据完全随机缺失机制的定义，设置两种常见的先验数据缺失场景，建立恒模与相似性约束下的雷达波形设计问题模型；然后将波形设计过程与状态-动作-奖励的动态规划结合，进一步构建杂波环境下波形设计的马尔可夫决策模型，为雷达波形的优化设计提供模型基础。

2.1 杂波数据缺失模型

不失一般性，本文考虑单输入单输出的雷达系统，在一个时间相参积累间隔(Coherent Processing Interval, CPI)内发射 N 个脉冲信号。假设发射端发射窄带信号，在时间维度对信号进行离散采样，可得在某一包含目标距离环上的雷达回波信号为

$$y(m) = \alpha \mathbf{w}^H \mathbf{s} + \mathbf{w}^H \mathbf{c}_m + \mathbf{w}^H \mathbf{n} \quad (1)$$

其中， α 表示发射信号的幅度， $\mathbf{s} \in \mathbb{C}^{N \times 1}$ 表示离散发射波形矢量， $\mathbf{c}_m \in \mathbb{C}^{N \times 1}$ 表示第 m 个距离环的杂波矢量， $\mathbf{n} \in \mathbb{C}^{N \times 1}$ 表示零均值方差为 σ_n^2 的高斯白噪声矢量， $\mathbf{w} \in \mathbb{C}^{N \times 1}$ 表示滤波权矢量， $(\cdot)^H$ 为向量/矩阵共轭转置运算符， m 为杂波距离环编号或先验杂波样本编号。

本文重点考虑构造杂波先验协方差矩阵中样本数据的缺失，假设 $\mathbf{C} = [\mathbf{c}_1 \ \mathbf{c}_2 \ \dots \ \mathbf{c}_M] \in \mathbb{C}^{N \times M}$ 为数据完整的杂波样本矩阵， M 为样本数。设置缺失矩阵 $\Delta_{N \times M}$ ， Δ 中仅有两个元素NaN和1，其中NaN表

示对应位置数据缺失，1表示该数据正常，则杂波缺失矩阵 $\tilde{\mathbf{C}}$ 可表示为

$$[\tilde{\mathbf{C}}]_m = [\mathbf{C} \odot \Delta]_m \quad (2)$$

其中， $[\cdot]_m$ 为矩阵的第 m 列， \odot 为Hadamard乘积。

由于气候、采集设备故障等导致的数据缺失完全随机且不可控，因此本文引入完全随机缺失机制。在该机制下，无法获取数据的分布参数，如均值和方差，且数据的缺失位置与数据的先验分布相对独立。类比文献[11]中的缺失场景，本文考虑点状缺失和块状缺失两种杂波先验数据缺失场景，构造缺失矩阵 Δ 的方式分别如下：

场景1：点状缺失。设置点状缺失概率 $p \in (0, 1)$ ，生成 $N \times M$ 个服从 $[0, 1]$ 均匀分布的随机数构成矩阵 Δ ，若 $\Delta(i, j) > p$ ，则 $\Delta(i, j) = 1$ ，反之 $\Delta(i, j) = \text{NaN}$ ，其中 $i = 1, 2, \dots, N$ ， $j = 1, 2, \dots, M$ ；

场景2：块状缺失。同场景1，设置块状缺失概率 $p' \in (0, 1)$ ，在 Δ 中随机生成 n 个缺失块并赋值为NaN，其余数据皆令为1。缺失块行数和列数分别为 r_l 和 c_l ， $l = 1, 2, \dots, n$ 。为保证场景设置的公平性，两种场景下的缺失概率应满足：

$$p' = \sum_{l=1}^n \frac{r_l c_l}{NM} = p \quad (3)$$

点状缺失和块状缺失示意图分别如图1(a)、图1(b)所示，其中，白色代表缺失数据(取值为NaN)，黑色代表正常数据(取值为1)。

2.2 波形设计问题模型

通常情况下，为了提高雷达的目标检测性能，需要在滤波处理输出后获取最大的SCNR。因此，本文采用最佳滤波权矢量，即 $\mathbf{w} = \mathbf{R}^{-1} \mathbf{s}$ ，其中 $\mathbf{R} = \mathbb{E}[(\mathbf{c} + \mathbf{n})(\mathbf{c} + \mathbf{n})^H]$ 为杂波+噪声协方差矩阵。进一步，考虑恒模约束和相似性约束，可将雷达波形优化问题表示为

$$\begin{aligned} & \max_{\mathbf{s}} \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s} \\ & \text{s.t. } |s_n| = 1, n = 1, 2, \dots, N \\ & \quad \|\mathbf{s} - \mathbf{s}_0\|^2 < \varepsilon \\ & \quad \mathbf{R}_c = f(\tilde{\mathbf{C}}) \end{aligned} \quad (4)$$

其中， $\mathbf{R} = \mathbf{R}_c + \sigma_n^2 \mathbf{I}$ ， \mathbf{R}_c 为杂波协方差矩阵， \mathbf{I} 为单位矩阵； $(\cdot)^{-1}$ 表示矩阵求逆运算； $|s_n| = 1$ 表示恒模约束， s_n 表示向量 \mathbf{s} 的第 n 个元素； $\|\mathbf{s} - \mathbf{s}_0\|^2 < \varepsilon$ 为波形相似性约束， ε 为相似度约束， \mathbf{s}_0 为参考波形。

当杂波先验协方差矩阵 \mathbf{R}_c 完全已知时，已有众多学者针对相关波形优化问题展开研究，如Li等

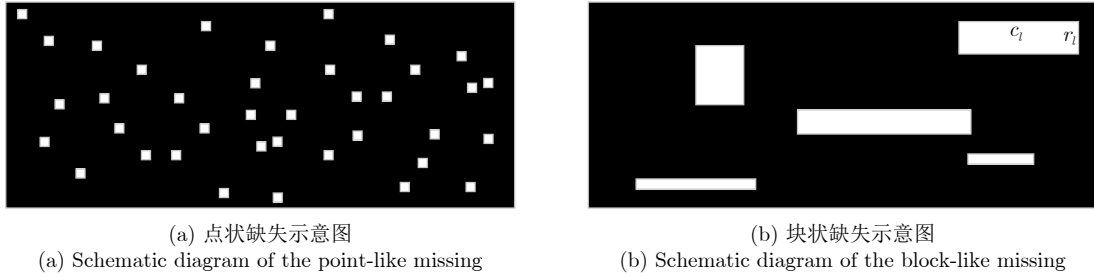


图 1 缺失场景示意图

Fig. 1 Schematic diagram of the missing scene

人^[24]提出的SWORD (Signal Waveform's Optimal-under-Restriction Design)方法等。然而, \mathbf{R}_c 估值的高准确度依赖于大量完整的样本数据。在杂波先验数据缺失的情况下, 样本信息不够精确, 协方差矩阵的估计误差增大, 难以保证优越的杂波抑制性能。因此, 本文考虑样本数据缺失情况下的波形设计问题, 并在式(4)中引入协方差矩阵约束 $\mathbf{R}_c = h(\tilde{\mathbf{C}})$, 用于表征该杂波协方差矩阵由数据缺失样本构成, 其中函数 $h(\cdot)$ 映射了该构成过程。

2.3 波形优化马尔可夫过程模型

由于强化学习强调智能体在未知环境中以不断试错的方式学习得到最优策略, 故本节将上述先验数据缺失条件下波形优化问题刻画为马尔可夫过程。

将雷达视为智能体, 在 t 时刻, 智能体根据当前时刻的状态 s_t 以及策略 $\pi(a_t|s_t)$ 向环境中发射动作 a_t , 其中策略是从状态到动作概率分布的函数映射。然后, 状态 s_t 根据状态转移函数 $p(s_{t+1}|s_t, a_t)$ 转换到 s_{t+1} , 并获得相应的反馈奖励 r_{t+1} 。假设 t 时刻的状态仅与上一时刻的状态有关, 则整个学习过程可以描述为一个马尔可夫决策过程(Markov Decision Process, MDP)^[25], 用4元组表示为

$$\{S, A, P, R\} \quad (5)$$

其中, S 为状态集合 $s_t \in S$, A 为动作集合 $a_t \in A$, P 为状态转移函数集合 $p_t \in P$, R 为奖励集合 $r_t \in R$ 。

本文假设雷达发射机发射功率不变且载波频率恒定, 波形参数仅与发射相位有关, 故将 t 时刻的离散相位值以及信噪比作为雷达在环境中的状态表征; 若考虑相似性约束, 则添加相似度衡量当前波形与参考波形的偏差, 以期快速获得良好的性能参数。相似度定义为

$$\varsigma = \|s_t - s_0\| \quad (6)$$

其中, s_t 表示 t 时刻的发射波形。

综上, t 时刻的状态可表示为

$$s_t = \{\theta_{t1}, \theta_{t2}, \dots, \theta_{tN}, \varsigma_t, \text{SCNR}_t\} \quad (7)$$

其中, θ_{ti} 表示 t 时刻的离散相位值, $i = 1, 2, \dots, N$ 。

将雷达的动作设计为“指令”操作, 即 $a_t = \{a_{t1}, a_{t2}, \dots, a_{tN}\}$, 其中, a_{ti} 有3种可能取值: 0, 1和2, 不同取值对应不同的操作:

$$\left. \begin{aligned} \theta_{ti} &= \theta_{ti} - \sigma, a_{ti} = 0 \\ \theta_{ti} &= \theta_{ti}, a_{ti} = 1 \\ \theta_{ti} &= \theta_{ti} + \sigma, a_{ti} = 2 \end{aligned} \right\} \quad (8)$$

其中, σ 表示雷达执行一次动作对应相位值的变化量。 σ 的设置与相位的分辨率息息相关, σ 越小, 雷达发射波形精确度越高; 但当 σ 过小时, 雷达需要耗费更多的时间去学习最佳波形, 导致算法收敛速度变慢。需要注意的是, 由于相位的周期性, 在整个决策过程中需保证 θ_{ti} 的值始终位于 $[0, 2\pi]$ 内。

奖励函数是强化学习算法设计的核心, 其作用是针对上一时刻的动作向强化学习模型提供即时反馈。因此, 合理的奖励设计方式有利于智能体最优执行策略的学习。由于目标SCNR是影响检测性能的关键因素, 因此本文以雷达滤波处理后的SCNR作为 t 时刻动作所获得的奖励:

$$r_t = \text{SCNR}_t = \mathbf{s}^H (\mathbf{R}_c + \sigma_n^2 \mathbf{I})^{-1} \mathbf{s} \quad (9)$$

3 基于级联优化的波形设计算法

针对上述波形设计问题, 本文提出一种基于优先级填充-强化学习级联优化的波形设计方法。该级联算法包含缺失数据填充和波形优化两个部分, 分别用函数 $f(\cdot)$ 和 $g(\cdot)$ 表示。如图2所示, 输入数据缺失条件下的先验数据矩阵 \mathbf{X} , 根据缺失值判断数据缺失区域并将其作为缺失规则输入杂波环境中。通过优先级填充算法得到数据完整的杂波数据矩阵 $\hat{\mathbf{C}}$, 进一步估计杂波协方差矩阵 $\hat{\mathbf{R}}_c$, 即 $\hat{\mathbf{R}}_c = \mathbb{E}[\hat{\mathbf{C}}^H \hat{\mathbf{C}}] = f(\mathbf{X})$ 。强化学习波形设计的目的是通过雷达与杂波环境交互, 训练生成最大SCNR的发射波形, 即 $\mathbf{s} = g(\hat{\mathbf{R}}_c)$, 其中, 将优先级填充算法输出结果作为智能体与之交互的杂波环境。本文所提的级联架构亦是指以估计杂波协方差矩阵为基础的强化学

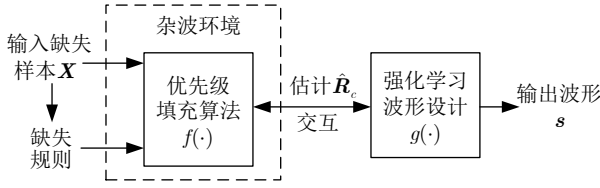


图2 级联优化算法整体框架图

Fig. 2 Overall framework diagram of the cascade optimization algorithm

习波形设计框架，通过优先级填充-强化学习级联的迭代过程，最终获得先验数据缺失条件下的优化波形。

3.1 基于优先级填充算法的杂波数据缺失估计方法

根据级联优化算法整体架构，本节首先介绍基于优先级的填充算法。将杂波缺失矩阵 \tilde{C} 以及缺失区域输入到杂波环境中，通过优先级填充算法的映射关系 $f(\cdot)$ ，完成对杂波缺失数据的修复，估计输出杂波协方差矩阵 \hat{R}_c 。如图3所示，该修复过程为：首先，通过判断杂波数据与缺失值是否匹配，将其分为目标区域 Ω 和源区域 Φ 。目标区域对应数据缺失部分，其轮廓用 $\delta\Omega$ 表示；剩余部分为源区域。然后在目标区域轮廓 $\delta\Omega$ 上选定点 p ，以 p 为中心点确定一待修复的正方形滑窗 Ψ_p ，在源区域内搜索最佳匹配数据 Ψ_q ，令 $\Psi_p = \Psi_q$ ，完成对 Ψ_p 的填充。最后，随着填充进度的推进，不断更新 $\delta\Omega$ 直至所有目标区域 Ω 填充完成。

填充顺序是数据修复质量的关键性因素之一。传统标准同心层填充算法以逆时针顺序将同心层块逐渐填入目标区域，可能导致目标轮廓附近源区域内的完整数据被意外重构，使得目标轮廓区域的填充效果变差。因此，为了确定最佳样本填充顺序，本节引入样本填充优先级的计算^[29]。针对上述目标轮廓区域填充效果不明显的问题，优先级的计算偏向于数据差异大的强边缘区域和高置信度的目标区域。

给定中心点 p 以及目标填充样本 Ψ_p ，将优先级函数 $P(p)$ 表示为

$$P(p) = C(p)D(p) \quad (10)$$

其中， $C(p)$ 为置信度项， $D(p)$ 为数据项。置信度 $C(p)$ 表征中心点 p 附近可靠数据的度量，即优先填充目标区域轮廓周围的样本并不断向内填充；数据项 $D(p)$ 体现 $\delta\Omega$ 两边数据信息的差异大小，优先填充差异较大的样本。计算公式分别为

$$C(p) = \frac{\sum_{p \in \Psi_p \cap \Phi} C(p)}{|\Psi_p|} \quad (11)$$

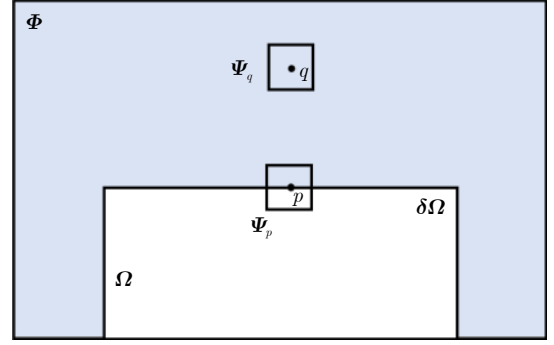


图3 数据修复结构原理图

Fig. 3 Schematic diagram of data repair structure

$$D(p) = \frac{|\nabla I_p^\perp \cdot J_p|}{\kappa} \quad (12)$$

其中， $|\Psi_p|$ 表示 Ψ_p 的面积， κ 是归一化因子， ∇I_p 为 $\Psi_p \cap I$ 区间内的最大梯度值， J_p 表示 p 点处正交于目标区域轮廓 $\delta\Omega$ 的单位向量。 \perp 表示正交。初始化时，设置置信度函数 $C(p)$ 为

$$\left. \begin{aligned} C(p) &= 0, \quad \forall p \in \Omega \\ C(p) &= 1, \quad \forall p \in \Phi \end{aligned} \right\} \quad (13)$$

获得目标区域内所有待填充样本的优先级后，按照优先级递减的方向从源区域内搜寻最相似样本进行填充。上述过程可通过最小化平方差之和算法实现：

$$\Psi_{\hat{q}} = \arg \min_{\Psi_q \in \Phi} d(\Psi_p, \Psi_q) \quad (14)$$

其中， $d(\Psi_a, \Psi_b)$ 表示样本 Ψ_a 和 Ψ_b 中数据的平方差之和。

将最相似样本 $\Psi_{\hat{q}}$ 内的数据复制到 Ψ_p 内，更新置信度 $C(p)$ ：

$$C(p) = C(\hat{p}), \quad \forall p \in \Psi_{\hat{p}} \cap \Omega \quad (15)$$

此时杂波数据矩阵的目标区域轮廓 $\delta\Omega$ 发生相应变化，重复优先级的计算，逐层更新直至目标区域填充完成，即通过 $f(\cdot)$ 的映射关系，实现从杂波缺失矩阵 \tilde{C} 到协方差矩阵 \hat{R}_c 的估计过程。算法步骤的伪代码见表1。

3.2 基于强化学习算法的波形设计方法

由3.1节可知，将优先级填充算法补偿后的杂波矩阵作为强化学习优化算法中智能体的交互环境。雷达发射波形作用到杂波环境中并获取反馈奖励，通过最大化累积奖励和的方式决策出最优波形相位。整个过程基于端到端的学习方式，用函数 $g(\cdot)$ 表示，即 $s = g(\hat{R}_c)$ 。本文采用DDPG (Deep Deterministic Policy Gradient)算法实现函数 $g(\cdot)$ 的映射过程。DDPG算法是一种基于Actor-critic框架

表 1 优先级填充算法

Tab. 1 Priority filling algorithm

输入：杂波缺失矩阵 \tilde{C} ，滑窗维度 M
输出：杂波修复矩阵 \hat{C}
1: 将杂波缺失矩阵 \tilde{C} 分为目标区域 Ω 和源区域 Φ
2: 根据式(13)初始化置信度 $C(p)$
3: 识别目标区域轮廓 $\delta\Omega$
4: 根据式(10)计算优先级 $P(p)$, $\forall p \in \delta\Omega$
5: 找到优先级最高的待填充样本 $\Psi_{\hat{p}} \in \mathbb{C}^{M \times M}$ ，即 $\hat{p} = \arg \max_{p \in \delta\Omega} P(p)$
6: 根据式(14)得到最相似样本 $\Psi_{\hat{q}}$
7: 将最相似样本 $\Psi_{\hat{q}}$ 内的数据信息复制到 $\Psi_{\hat{p}}$ 内
8: 根据式(15)更新置信度 $C(p)$
9: 判断 Ω 是否为空集，如果是，算法结束；否则跳转3

的深度确定性策略梯度算法^[27]，由DPG (Deterministic Policy Gradient)算法改进而来，得益于其在连续动作空间中的有效决策性能，广泛用于解决各领域复杂连续的决策问题。在波形优化设计问题中，由于动作空间和状态空间的高维连续性，使用DDPG算法可以更好地输出最佳动作策略。

Actor-critic框架由Actor网络和Critic网络构成。Actor网络输入状态 s_t ，输出唯一确定性动作 a_t 与环境交互，返回下一时刻的状态 s_{t+1} 和奖励 r_{t+1} ，其网络权重参数用 θ 表示；Critic网络输入状态 s_t 和动作 a_t ，通过价值函数评估动作 a_t 的优劣性，其网络权重参数用 ω 表示。为了避免网络输出的震荡导致不稳定的学习行为，增强算法的稳定性，DDPG分别复制Actor网络和Critic网络生成相对应的目标网络，其网络权重参数分别用 θ' 和 ω' 表示。

在DDPG算法中，为了增加学习过程的随机性，提高智能体对环境的探索概率，针对Actor网络决策机制引入随机噪声，设计输出动作为

$$\mu'(s_t) = \mu(s_t | \theta) + \mathcal{N} \quad (16)$$

其中， $\mu(s_t | \theta)$ 表示Actor网络在状态 s_t 时的输出动作， \mathcal{N} 表示随机噪声过程。

Actor网络采用off-policy的训练方法，用于生成动作的行为策略和评估策略不同。引入策略梯度下降法更新其网络参数：

$$\nabla_{\theta} J_{\theta} \approx \frac{1}{K} \sum_i \nabla_a Q(s, a | \omega) \Big|_{s=s_i, a=\mu'(s_i)} \nabla_{\theta} \mu'(s_i | \theta) \quad (17)$$

其中， K 表示从经验池中采样的样本数。更新Critic网络时，采用均方误差计算其损失函数：

$$\mathcal{L} = \frac{1}{K} \sum_i (r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1} | \theta') | \omega') - Q(s_i, a_i | \omega))^2 \quad (18)$$

其中， γ 表示奖励衰减因子。

通过软更新的方式分别更新Actor网络和Critic网络的目标网络参数：

$$\left. \begin{aligned} \theta' &\leftarrow \tau\theta + (1-\tau)\theta' \\ \omega' &\leftarrow \tau\omega + (1-\tau)\omega' \end{aligned} \right\} \quad (19)$$

其中， τ 表示软更新系数，用于控制迭代中每一步更新的幅度，一般取 $\tau = 0.01$ 。

结合前文分析，智能体通过Actor网络输出波形相位，将当前所处状态和动作矢量信息输入到Critic网络中；Critic网络输出评估状态-动作对的 Q 值。利用梯度下降法和均方误差分别计算Actor网络和Critic网络的损失函数，结合软更新的方式，实现策略网络、评估网络以及对应目标网络参数的更新。DDPG伪代码如表2所示，算法结构框图如图4所示。

4 仿真分析

为了分析雷达杂波先验数据缺失对发射波形优化过程的影响，验证本文所提级联优化算法的有效性，本节针对杂波先验数据点状与块状缺失场景进行仿真试验。

杂波数据缺失模型参数设置如下： $N = 16$ ， $M = 2000$ ，数据缺失概率 $p = 0.1$ 。为方便观察数据修复效果，将杂波数据矩阵可视化，如图5所示，为杂波参考矩阵 C 的俯视图。据2.1节所述，设置点状缺失和块状缺失如图6所示，其中，黄色部分表示数据缺失。由于修复完成的杂波矩阵将进一步计算杂波协方差矩阵 \hat{R}_c ，为提高数据缺失的代表性，将大多数缺失数据集中设置于 N 个脉冲编码内（图6中第41行到第56行内），保证 \hat{R}_c 内耦合更多缺失数据信息。

根据样本优先级填充算法，对杂波数据缺失矩阵目标区域进行修复。考虑到杂波数据缺失对后续波形优化的影响，将杂波缺失矩阵 \tilde{C} 中的NaN值设为500，并采用KNN填充算法作为对比算法进行仿真试验，结果如图7、图8所示（黑色方框表示杂波数据缺失矩阵的目标区域轮廓）。由图7可看出，针对完全随机缺失机制，杂波数据缺失部分在一定程度上得到了修复，然而点状缺失数据分布比较分散，无法得出明显的效果图；由图8可看出，在块状缺失场景下，优先级填充算法结合了优先级排序和样本相似度的计算，修复效果明显优于KNN填充算法，与参考矩阵 C 差异较小。由此可得出初步结论：优先级填充算法考虑了缺失点周围数据的关联性，可以从源区域内全局搜索出最相似样本块进行填充，并取得良好的修复效果。

表 2 基于DDPG的波形优化算法

Tab. 2 Algorithm for waveform optimization based on DDPG

输入：Actor策略网络及其目标网络，Critic评估网络及其目标网络，网络参数分别为 $\theta, \theta', \omega, \omega'$ ，奖励衰减因子 γ ，软更新系数 τ ，最大迭代次数 T ，经验回放池R，采样样本数 K

输出：最佳Actor策略网络 $\pi_*(a|s)$

- 1: 随机初始化Actor策略网络参数 θ 和Critic评估网络参数 ω
- 2: 初始化目标网络参数 $\theta' = \theta, \omega' = \omega$
- 3: 初始化经验回放池R
- 4: **for** 回合 $e \in \{1, 2, \dots, T\}$ **do**
- 5: 初始化随机噪声 \mathcal{N} ，初始化状态 s
- 6: 根据式(16)得到Actor网络的输出动作 a_t
- 7: 执行动作 a_t ，获得下一时刻状态 s_{t+1} ，反馈奖励 r_t
- 8: 将 $\{s_t, a_t, r_t, s_{t+1}\}$ 存入经验回放池R
- 9: 从经验回放池中随机采样 K 个经验样本 $\{s_i, a_i, r_i, s_{i+1}\}, i = 1, 2, \dots, K$
- 10: 根据式(17)和式(18)更新Actor策略网络和Critic评估网络
- 11: 根据式(19)更新目标网络参数 θ' 和 ω'
- 12: 判断 s_{t+1} 是否为终止状态，如果是，迭代完毕，否则跳转步骤5

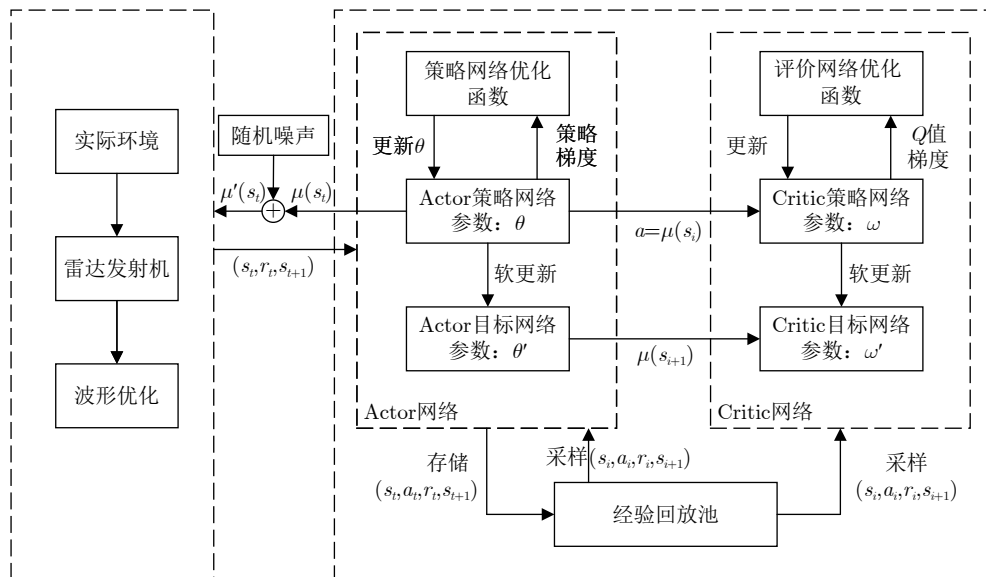


图 4 雷达波形设计的DDPG算法结构图

Fig. 4 Structure diagram of DDPG algorithm for radar waveform design

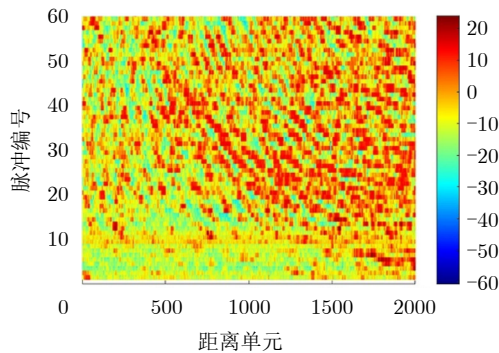


图 5 杂波参考矩阵可视化图

Fig. 5 Visualization of the clutter reference matrix

为进一步验证上述优先级填充算法对缺失数据修复的有效性，设置不同的缺失概率，将修复后的杂波矩阵与参考杂波矩阵的均方根误差(Root Mean Square Error, RMSE)作为衡量指标，定量分析算法的优劣性。如图9所示，RMSE随着缺失概率 p 的增加不断增大。在点状缺失场景下，优先级填充算法修复误差明显低于KNN填充算法；在块状缺失场景下，当缺失概率 $p < 0.3$ 时，优先级填充算法略优，随着 p 继续增加，其算法误差亦快速增大，填充性能下降且劣于KNN填充算法。对比不同的缺失场景，根据缺失数据分布的均匀性可

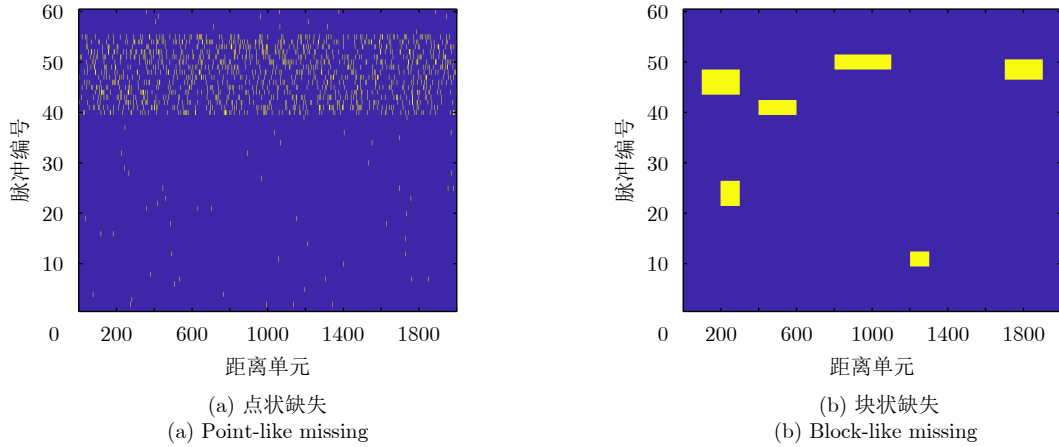


图 6 杂波数据信息缺失图

Fig. 6 Missing information of clutter data

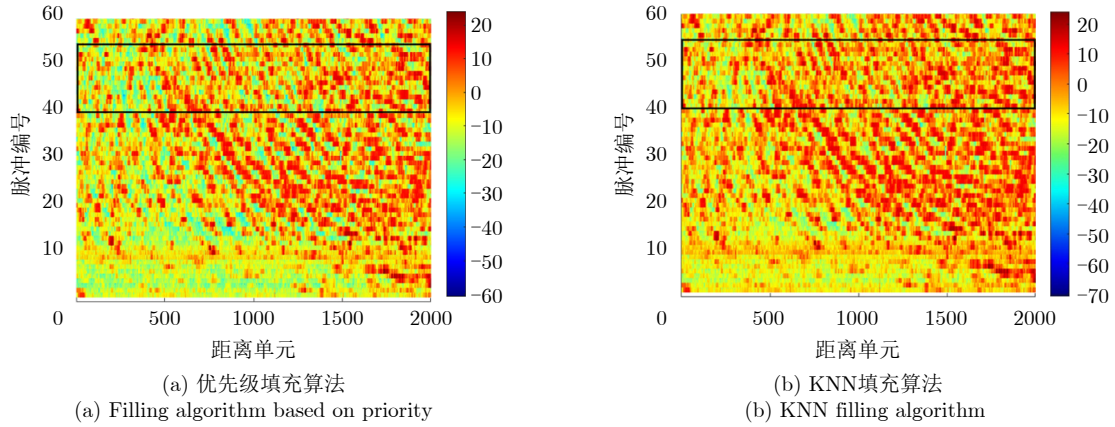


图 7 点状缺失场景下缺失数据修复图

Fig. 7 Missing data repair diagram in the point-like missing scene

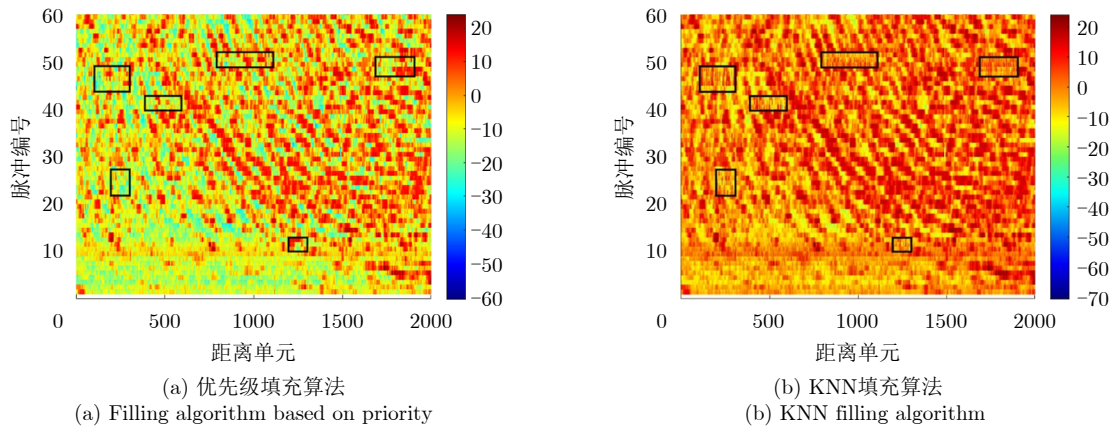


图 8 块状缺失场景下缺失数据修复图

Fig. 8 Missing data repair diagram in the block-like missing scenario

知，块状缺失的数据相较于点状缺失更为集中，故点状缺失的修复性能更佳，误差更小。

通过上述仿真实现 $\hat{R}_c = f(\mathbf{X})$ 的估计过程，进一步通过强化学习算法对雷达发射端波形相位动作进行决策，实现发射波形的最优化，提高信杂噪

比。本文采用DDPG的强化学习决策算法，相关参数如表3所示。

设置相位初始状态为 $\mathbf{0}_{1 \times 16}$ ，运行步骤 $\text{step} = 1000$ ，单个回合结束。训练曲线如图10所示，图10(a)、图10(b)分别表示点状缺失和块状缺失两

种杂波缺失场景下的强化学习奖励曲线图。由图10可以看出，奖励曲线皆由初始的负奖励值快速上升至400左右，然后逐渐收敛并趋于稳定，对应最大信噪比可达16 dB。说明在杂波先验数据缺失的条件下，雷达发射机通过Critic网络对动作策略的评估及奖励反馈，成功学习到较优的波形相位动作。对比两种缺失场景，由于块状缺失场景下缺失块中心填充误差较大，奖励曲线收敛速度相对较慢。

为验证完全随机缺失机制下基于优先级填充算法-强化学习级联优化的训练效果，以雷达接收信号的SCNR作为参考指标进行仿真分析。如图11所示，横坐标表示输入信杂噪比 $\alpha^2/\text{tr}(\mathbf{R}_c + \sigma^2)$ ，其中 α^2 表示信号的能量， σ^2 为噪声方差，依据不同输入条件设计噪声能量值。在仿真中，设置缺失概率为0.1和0.3，将协方差矩阵无缺失条件下SWORD算法所得结果作为性能上限，将数据缺失条件下

SWORD算法以及KNN填充后的SWORD算法与本文所提算法进行仿真对比。由图11可知，针对完全随机缺失机制下的缺失场景，随着杂噪比值的变化，所提算法能够较好地逼近无缺失条件的波形设计性能，且优于KNN补偿算法。其中，随着缺失概率 p 的增大，杂波有效先验数据信息减少，输出性能略有下降。对比不同缺失场景，点状缺失场景下的优化性能相较于块状缺失场景更优，但本文所提算法的优势在块状缺失场景中更为明显。

为验证不同约束对算法输出性能的影响，在雷达波形优化方程中添加相似性约束，并根据式(7)在强化学习 t 时刻的状态中添加相似度参数 ζ 。仿真结果如图12和图13所示。从图12可以看出，不同场景下的奖励曲线有明显的提升和收敛表现；相较于无相似性约束下的仿真结果，由于具备参考波形的引导，Actor网络能在相似度的衡量指标下迅速逼近最佳输出动作，算法收敛速度加快且奖励值小幅提升。从图13可以看出，本文所提算法在恒模与相似性约束条件下，仍然具有较优的杂波抑制性能。由于对比实验均考虑了相同的波形相似性约束，虽然在输出性能曲线上不同算法的仿真结果差异变小，但本文所提算法与无缺失条件下SWORD

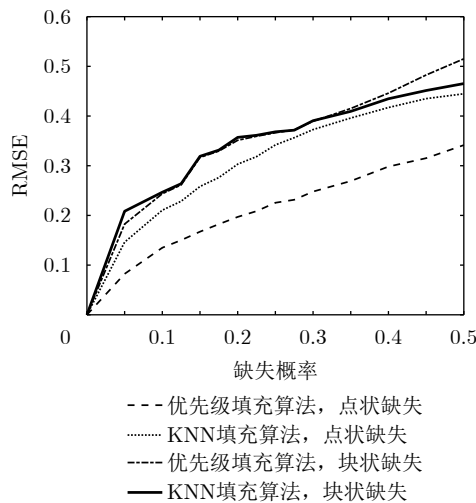


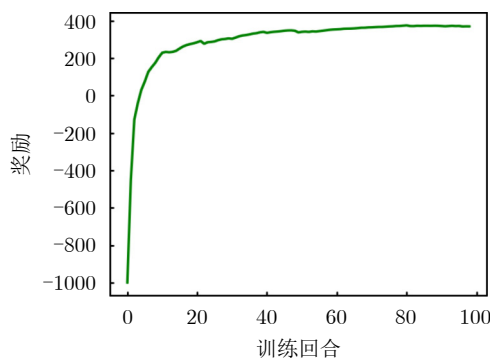
图9 不同缺失概率下数据修复性能分析

Fig. 9 Data repair performance analysis under different missing probability

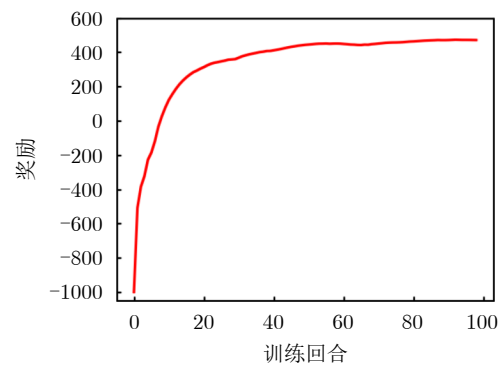
表3 强化学习参数表

Tab. 3 Reinforcement learning parameters table

参数	数值
经验池大小	200000
训练批次大小	64
训练总次数	100000
回合训练次数	1000
学习率	0.001
惩罚因子	0.98
Actor, Critic网络层节点数	[400, 300]



(a) 点状缺失
(a) Point-like missing



(b) 块状缺失
(b) Block-like missing

图10 仅恒模约束下强化学习奖励曲线图

Fig. 10 Reinforcement learning reward curves under constant modulus constraint

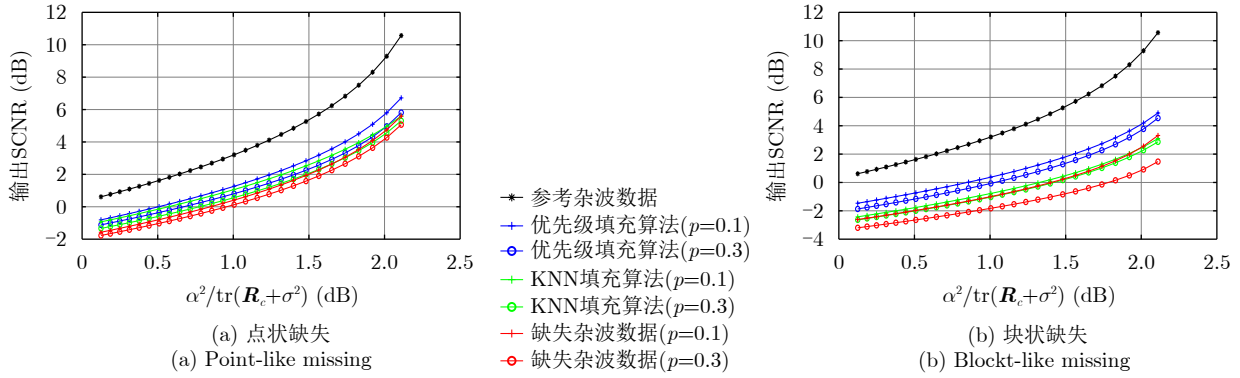


图 11 仅恒模约束下不同场景的输出SCNR性能曲线图

Fig. 11 Output SCNR performance curves of different scenarios under constant modulus constraint

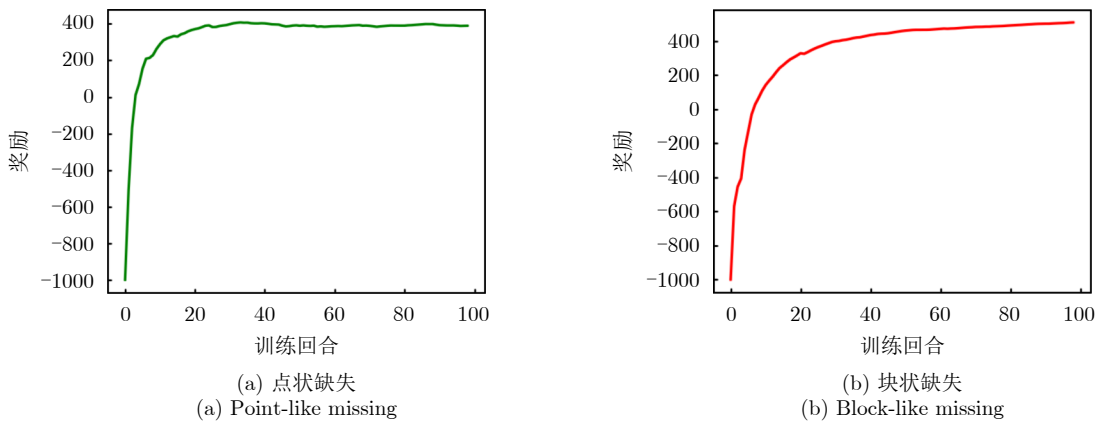


图 12 相似性约束下强化学习奖励曲线图

Fig. 12 Reinforcement learning reward curves under similarity constraints

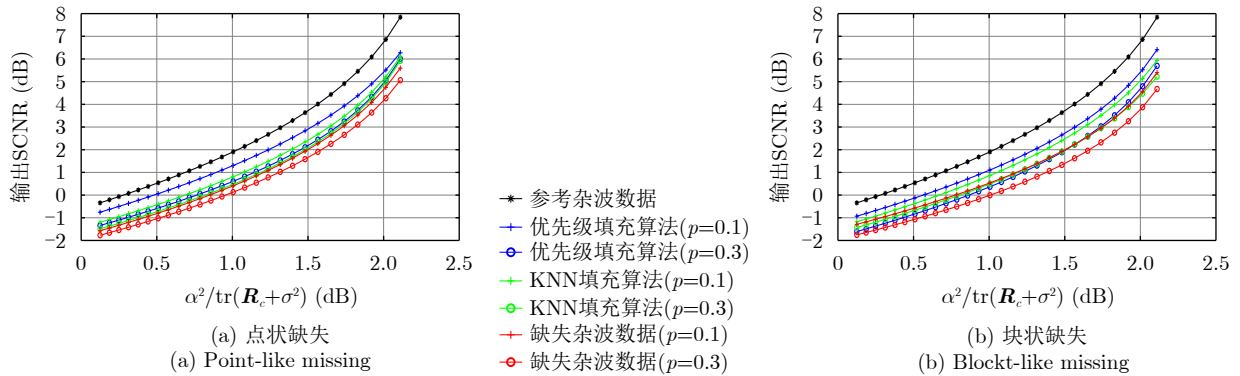


图 13 相似性约束下不同场景的输出SCNR性能曲线图

Fig. 13 Output SCNR performance curves of different scenarios under similarity constraints

算法的性能表现更为接近，且依然优于KNN算法补偿后的传统优化算法。

5 结语

针对杂波先验数据缺失条件下的波形优化问题，本文提出了一种优先级填充-强化学习级联优化的波形设计算法。为求解恒模约束和相似性约束下的非凸非线性优化问题，本文首先根据缺失数据

的优先度在源区域内搜索最佳匹配数据进行填充，实现随机缺失机制下的数据修复，再结合DDPG强化学习决策算法，通过Actor-Critic网络架构训练决策出雷达最佳相位选择策略，实现最大信杂噪比下的波形优化。最后，本文针对杂波先验数据不同缺失场景，采用数值仿真对比KNN算法填充后的SWORD优化算法，分析结果表明：(1)级联算法无需考虑先验数据的分布函数，对杂波数据的完全

随机缺失具有较好的修复效果；(2)针对点状缺失和块状缺失场景，本文所提算法均可获得优越的杂波抑制性能，进一步提升雷达的探测能力。

参 考 文 献

- [1] TANG Bo and TANG Jun. Joint design of transmit waveforms and receive filters for MIMO radar space-time adaptive processing[J]. *IEEE Transactions on Signal Processing*, 2016, 64(18): 4707–4722. doi: [10.1109/TSP.2016.2569431](https://doi.org/10.1109/TSP.2016.2569431).
- [2] TANG Bo, TUCK J, and STOICA P. Polyphase waveform design for MIMO radar space time adaptive processing[J]. *IEEE Transactions on Signal Processing*, 2020, 68: 2143–2154. doi: [10.1109/TSP.2020.2983833](https://doi.org/10.1109/TSP.2020.2983833).
- [3] YU Xianxiang, CUI Guolong, YANG Jing, *et al.* Wideband MIMO radar waveform design[J]. *IEEE Transactions on Signal Processing*, 2019, 67(13): 3487–3501. doi: [10.1109/TSP.2019.2916732](https://doi.org/10.1109/TSP.2019.2916732).
- [4] WU Linlong and PALOMAR D P. Radar Waveform Design Via the Majorization-Minimization Framework[M]. CUI Guolong, DE MAIO A, FARINA A, *et al.* Radar Waveform Design Based on Optimization Theory. London: The Institution of Engineering and Technology, 2020: 185–220. doi: [10.1049/SBRA533E_ch7](https://doi.org/10.1049/SBRA533E_ch7).
- [5] O'ROURKE S M, SETLUR P, RANGASWAMY M, *et al.* Quadratic semidefinite programming for waveform-constrained joint filter-signal design in STAP[J]. *IEEE Transactions on Signal Processing*, 2020, 68: 1744–1759. doi: [10.1109/TSP.2020.2977271](https://doi.org/10.1109/TSP.2020.2977271).
- [6] TANG Bo, NAGHSH M M, and TANG Jun. Relative entropy-based waveform design for MIMO radar detection in the presence of clutter and interference[J]. *IEEE Transactions on Signal Processing*, 2015, 63(14): 3783–3796. doi: [10.1109/TSP.2015.2423257](https://doi.org/10.1109/TSP.2015.2423257).
- [7] WANG Yikai, XIA Wei, HE Zishu, *et al.* Polarimetric detection in compound Gaussian clutter with Kronecker structured covariance matrix[J]. *IEEE Transactions on Signal Processing*, 2017, 65(17): 4562–4576. doi: [10.1109/TSP.2017.2716912](https://doi.org/10.1109/TSP.2017.2716912).
- [8] SUN Guohao, HE Zishu, TONG Jun, *et al.* Mutual information-based waveform design for MIMO radar space-time adaptive processing[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(4): 2909–2921. doi: [10.1109/TGRS.2020.3008320](https://doi.org/10.1109/TGRS.2020.3008320).
- [9] 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. doi: [10.12000/JR19072](https://doi.org/10.12000/JR19072).
CUI Guolong, YU Xianxiang, YANG Jing, *et al.* An overview of waveform optimization methods for cognitive radar[J]. *Journal of Radars*, 2019, 8(5): 537–557. doi: [10.12000/JR19072](https://doi.org/10.12000/JR19072).
- [10] 王珽, 赵拥军, 胡涛. 机载MIMO雷达空时自适应处理技术研究进展[J]. *雷达学报*, 2015, 4(2): 136–148. doi: [10.12000/JR14091](https://doi.org/10.12000/JR14091).
WANG Ting, ZHAO Yongjun, and HU Tao. Overview of space-time adaptive processing for airborne MIMO radar[J]. *Journal of Radars*, 2015, 4(2): 136–148. doi: [10.12000/JR14091](https://doi.org/10.12000/JR14091).
- [11] AUBRY A, DE MAIO A, MARANO S, *et al.* Structured covariance matrix estimation with missing-(complex) data for radar applications via expectation-maximization[J]. *IEEE Transactions on Signal Processing*, 2021, 69: 5920–5934. doi: [10.1109/TSP.2021.3111587](https://doi.org/10.1109/TSP.2021.3111587).
- [12] HIPPERT-FERRER A, EL KORSO M N, BRELOY A, *et al.* Robust low-rank covariance matrix estimation with a general pattern of missing values[J]. *Signal Processing*, 2022, 195: 108460. doi: [10.1016/j.sigpro.2022.108460](https://doi.org/10.1016/j.sigpro.2022.108460).
- [13] PAVEZ E and ORTEGA A. Covariance matrix estimation with non uniform and data dependent missing observations[J]. *IEEE Transactions on Information Theory*, 2021, 67(2): 1201–1215. doi: [10.1109/TIT.2020.3039118](https://doi.org/10.1109/TIT.2020.3039118).
- [14] ZHANG Ying, LIE J P, NG B P, *et al.* Robust minimum ℓ_1 -norm adaptive beamformer against intermittent sensor failure and steering vector error[J]. *IEEE Transactions on Antennas and Propagation*, 2010, 58(5): 1796–1801. doi: [10.1109/TAP.2010.2044353](https://doi.org/10.1109/TAP.2010.2044353).
- [15] XIONG Can, XIAO Gaobiao, HOU Yibei, *et al.* A compressed sensing-based element failure diagnosis method for phased array antenna during beam steering[J]. *IEEE Antennas and Wireless Propagation Letters*, 2019, 18(9): 1756–1760. doi: [10.1109/LAWP.2019.2929353](https://doi.org/10.1109/LAWP.2019.2929353).
- [16] GAO Yongchan, LIAO Guisheng, and LIU Weijian. High-resolution radar detection in interference and nonhomogeneous noise[J]. *IEEE Signal Processing Letters*, 2016, 23(10): 1359–1363. doi: [10.1109/LSP.2016.2597738](https://doi.org/10.1109/LSP.2016.2597738).
- [17] LIM D, GIANELLI C D, and LI Jian. Automatic target recognition in missing data cases[J]. *IEEE Aerospace and Electronic Systems Magazine*, 2017, 32(7): 40–49. doi: [10.1109/MAES.2017.150273](https://doi.org/10.1109/MAES.2017.150273).
- [18] SHEN Lei, LIU Zhiwen, XU Yougen, *et al.* Robust polarimetric adaptive detector against target steering matrix mismatch[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2020, 56(1): 442–455. doi: [10.1109/TAES.2019.2916708](https://doi.org/10.1109/TAES.2019.2916708).
- [19] LOUNICI K. High-dimensional covariance matrix estimation with missing observations[J]. *Bernoulli*, 2014, 20(3): 1029–1058. doi: [10.3150/12-BEJ487](https://doi.org/10.3150/12-BEJ487).
- [20] LIU Junyan and PALOMAR D P. Regularized robust estimation of mean and covariance matrix for incomplete data[J]. *Signal Processing*, 2019, 165: 278–291. doi: [10.1016/j.sigpro.2019.108460](https://doi.org/10.1016/j.sigpro.2019.108460).

- j.sigpro.2019.07.009.
- [21] XU Danlei, DU Lan, LIU Hongwei, *et al.* Compressive sensing of stepped-frequency radar based on transfer learning[J]. *IEEE Transactions on Signal Processing*, 2015, 63(12): 3076–3087. doi: [10.1109/TSP.2015.2421473](https://doi.org/10.1109/TSP.2015.2421473).
- [22] LV Qinzhe, QUAN Yinghui, WEI Feng, *et al.* Radar deception jamming recognition based on weighted ensemble CNN with transfer learning[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5107511. doi: [10.1109/TGRS.2021.3129645](https://doi.org/10.1109/TGRS.2021.3129645).
- [23] JIANG Wei, HAIMOVICH A M, and SIMEONE O. Joint design of radar waveform and detector via end-to-end learning with waveform constraints[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2022, 58(1): 552–567. doi: [10.1109/TAES.2021.3103560](https://doi.org/10.1109/TAES.2021.3103560).
- [24] LI Jian, GUERCI J R, and XU Luzhou. Signal waveform's optimal-under-restriction design for active sensing[J]. *IEEE Signal Processing Letters*, 2006, 13(9): 565–568. doi: [10.1109/LSP.2006.874465](https://doi.org/10.1109/LSP.2006.874465).
- [25] BELLMAN R. A Markovian decision process[J]. *Journal of Mathematics and Mechanics*, 1957, 6(5): 679–684.
- [26] CRIMINISI A, PEREZ P, and TOYAMA K. Region filling and object removal by exemplar-based image inpainting[J]. *IEEE Transactions on Image Processing*, 2004, 13(9): 1200–1212. doi: [10.1109/TIP.2004.833105](https://doi.org/10.1109/TIP.2004.833105).
- [27] LILLICRAP T P, HUNT J J, PRITZEL A, *et al.* Continuous control with deep reinforcement learning[EB/OL]. <http://arxiv.org/abs/1509.02971>, 2015.

作者简介

张应奎，硕士生，主要研究方向为基于机器学习的雷达信号处理方法。

孙国皓，博士，副研究员，主要研究方向为认知雷达信号处理、分布式雷达信号处理、机载/星载雷达信号处理、天域态势感知。

钟苏川，副研究员，主要研究方向为随机动力系统、随机信号处理等。

余显祥，博士后，主要研究方向为雷达波形设计与处理、最优化理论算法以及阵列信号处理等。

(责任编辑：于青)