

纪念王威琪院士诞辰 85 周年

## 相位感知的水下平台背景宽带噪声干扰抑制

杨 洋<sup>1,2</sup> 郭良浩<sup>1†</sup> 宫在晓<sup>1</sup> 刘建军<sup>1</sup>

(1 中国科学院声学研究所 声场声信息国家重点实验室 北京 100190)

(2 中国科学院大学 北京 100049)

2023 年 2 月 23 日收到

2023 年 5 月 15 日定稿

**摘要** 针对水下航行器等平台存在宽带背景噪声干扰导致目标检测能力下降问题, 提出一种适用于多通道水听器阵列的深度学习宽带背景噪声干扰抑制方法。该方法通过含干扰信号的多通道频域特征级联保留了相位信息, 使用深度复数神经网络建立了一个估计纯净目标信号频谱特征的学习模型实现干扰抑制, 再利用常规波束形成方法实现目标检测与跟踪。仿真结果表明, 在信干比为 $-15$  dB,  $-20$  dB,  $-25$  dB 的双目标信号条件下, 所提干扰抑制方法可以有效减少近场干扰影响, 提高了常规波束形成的检测能力。湖试结果表明, 该方法能够自适应地实现平台宽带背景噪声干扰抑制, 有效提升目标检测和跟踪性能。

**关键词** 干扰抑制, 水下平台, 机器学习, 波束形成

**PACS 数** 43.60

**DOI:** 10.12395/0371-0025.2023025

### Phase-aware underwater platform background broadband noise interference suppression

YANG Yang<sup>1,2</sup> GUO Lianghao<sup>1†</sup> GONG Zaixiao<sup>1</sup> LIU Jianjun<sup>1</sup>

(1 *State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Sciences* Beijing 100190)

(2 *University of Chinese Academy of Sciences* Beijing 100049)

Received Feb. 23, 2023

Revised May 15, 2023

**Abstract** Aiming at the problem of the decline of target detection capability caused by broadband background noise interference in underwater vehicles and other platforms, a deep learning broadband background noise suppression method suitable for multi-channel hydrophone arrays is proposed. This method preserves the phase information by cascading the multi-channel frequency domain characteristics of the interference signal, and a learning model is established to estimate the spectral characteristics of pure target signals by using a deep complex neural network to achieve interference suppression, and then the conventional beamforming method is used to achieve object detection and tracking. The simulation results show that under the dual target signal conditions of  $-15$  dB,  $-20$  dB and  $-25$  dB signal-to-interference ratio, the proposed method can effectively reduce the influence of near-field interference and improve the detection ability of conventional beamforming. The lake test results show that the proposed method can adaptively realize the suppression of platform broadband background noise interference, and effectively improve the target detection and tracking performance.

**Keywords** Interference suppression, Underwater platform, Machine learning, Beamforming

### 引言

潜航器等水下移动平台的宽带强自噪声的严重干扰, 使得平台搭载水下定位系统接收目标信号的

信干噪比受到影响, 降低其对水下目标的有效作用距离。因此, 水下目标定位系统<sup>[1-2]</sup>需要具有较强的近场宽带噪声干扰抑制能力, 其中基于单通道的干扰抑制方法有谱减法、维纳滤波和最小均方差估计<sup>[3-5]</sup>等, 基于多通道的干扰抑制方法有子空间法<sup>[6-8]</sup>、自

† 通讯作者: 郭良浩, [glh2002@mail.ioa.ac.cn](mailto:glh2002@mail.ioa.ac.cn)

适应噪声抵消 (ANC)<sup>[9-10]</sup> 等。已有的子空间方法不适用于目标为远场信号和干扰为近场信号的模型; 自适应噪声抵消需获得有效的干扰参考信号, 但实际复杂水声环境中通常难以稳定获得纯净参考信号, 使自适应噪声抵消的应用效果受到限制。

近年来针对单通道声信号干扰抑制问题, 基于深度神经网络 (DNN) 的算法应用更为广泛。基于 DNN 的声干扰抑制方法可以充分挖掘含噪声信号与纯净目标信号之间的非线性关系, 在低信噪比和非平稳噪声干扰条件下有着明显的优势。根据训练目标的不同, 主要分为两大类方法: 时频掩蔽 (T-F Masking)<sup>[11-14]</sup> 和谱映射方法<sup>[15-16]</sup>。前者的训练目标为每个时频单元中目标信号能量的分布, 即一个乘性的时频掩蔽, 将其与输入的时频表征相乘, 得到预测的时频表征。时频掩蔽方法, 包括理想二值掩蔽 (IBM)<sup>[11-12]</sup>、理想比值掩蔽 (IRM)<sup>[13]</sup> 和幅度谱掩蔽 (SMM)<sup>[14]</sup>, 仅关注信号强度信息。后者将纯净目标信号的对数功率谱作为训练目标, 通过 DNN 得到一个含噪声信号对数功率谱和纯净目标声信号对数功率谱之间的回归映射。上述方法主要针对含噪声信号的短时离散傅里叶变换域 (STFT), 通常只分析和估计幅度谱, 在重构目标波形时直接使用含噪声信号的原始相位, 严重限制了低信噪比声信号的干扰抑制效果<sup>[17]</sup>。

目前大多数深度神经网络中的特征处理都是针对实数域的, 随着复数域深度神经网络如深度复数神经网络 (DCN) 和深度复数卷积神经网络 (DCCN) 等<sup>[18]</sup> 的进一步研究, 利用频域信号作为深度神经网络特征处理成为可能<sup>[19]</sup>。相位敏感掩蔽方法 (PSM)<sup>[20]</sup> 在利用相位信息上做出了尝试。复数比值掩蔽 (CRM)<sup>[21]</sup> 方法通过使用深度神经网络从含噪信号中估计复数比值掩蔽, 得到目标信号复数频谱。复数时频掩蔽的联合学习方法<sup>[22]</sup>, 通过为时频掩蔽提出新的目标函数和参数学习方法来解决幅度和相位谱之间的失配问题。与直接使用复杂的理想比值掩蔽来监督深度神经网络的方法不同, 双流网络<sup>[23]</sup> 中的幅度流和相位流专用于幅度和相位预测, 且在二者之间增加了信息交互机制, 从而在处理过程中能互相利用另外一路的信息作为参考。这些单通道深度神经网络方法因处理对象包含了信号的幅度和相位信息, 理论上可进一步提高声信号的干扰抑制效果。上述基于复数域的深度学习方法目前初步应用于单通道声信号干扰抑制, 对低信噪比下的多通道近场干扰抑制还需要深入研究。

针对平台背景宽带自噪声强干扰下的目标检测

问题, 提出一种适用于多通道水听器阵列的深度学习宽带背景噪声干扰抑制方法。首先对接收信号进行重叠分帧和短时傅里叶变换; 其次将频域特征拼接成多通道级联特征融合获取声场空间信息; 在特征提取时, 结合宽带随机噪声信号的相位信息, 将相干处理后的级联时频特征作为 DCCN 的输入; 然后设计模型损失函数和优化网络参数; 最后使用 DCCN 实现多通道背景干扰抑制并进行波束形成和信号检测。仿真和湖试结果表明, 平台存在宽带自噪声干扰时, 所提方法可有效提高水平阵对目标信号的检测和跟踪能力。

## 1 信号模型

如图 1 所示, 水下平台搭载一个由  $M$  个各向同性阵元组成的均匀线阵列, 阵元间距为  $\Delta$ 。  $D$  个目标信号和  $Q$  个干扰信源从不同方向入射, 根据平面波和球面波传播特性, 分别表示目标和干扰源信号到达各个阵元的幅度变化和时延, 不考虑声传播的多途效应, 第  $k$  组样本、第  $m$  个阵元接收数据中频率  $\omega_l$  对应的接收信号为

$$x_m(\omega_l, k) = \sum_{d=0}^{D-1} s_d(\omega_l, k) e^{-j\omega_l \tau_m(\phi_d)} + \sum_{q=0}^{Q-1} \rho_m(q) i_q(\omega_l, k) e^{-j\omega_l \mu_m(q)} + n_m(\omega_l, k), \quad m = 1, 2, \dots, M, \quad (1)$$

其中,  $\phi_d$  和  $\tau_m(\phi_d) = (m-1)\Delta \sin \phi_d / c$  分别表示第  $d$  个目标信号的入射角度和到达第  $m$  个阵元的相对时延;  $c$  表示声速;  $s_d(\omega_l, k)$  表示第  $k$  组样本、第  $d$  个目标信号的频域表示;  $\rho_m(q) = A / \sigma_m(q)$  表示第  $q$  个干扰源信号到达第  $m$  个阵元时的信号幅度,  $A$  为常数, 由干扰源的发射功率决定,  $\sigma_m(q)$  表示第  $q$  个干扰源到第  $m$  个阵元的距离;  $i_q(\omega_l, k)$  表示第  $k$  组采样、第  $q$  个干扰源的频域信号;  $\mu_m(q) = \sigma_m(q) / c$  表示第  $q$  个干扰源信

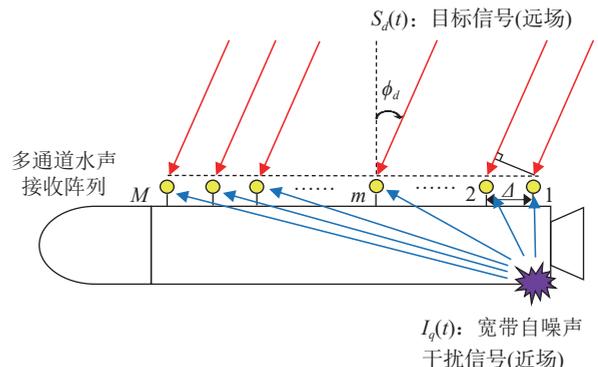


图 1 多通道水声阵列接收信号示意图

号到达第  $m$  个阵元的传播时延;  $n_m(\omega_l, k)$  表示第  $m$  个阵元频域加性高斯白噪声。由阵列各阵元接收数据构成的、对应第  $k$  组采样的频率  $\omega_l$  下的整个阵的接收数据向量为

$$\mathbf{x}(\omega_l, k) = \mathbf{A}(\omega_l, \phi)\mathbf{s}(\omega_l, k) + \mathbf{B}(\omega_l, q)\mathbf{i}(\omega_l, k) + \mathbf{n}(\omega_l, k), \quad (2)$$

其中,  $\mathbf{x}(\omega_l, k) = [x_1(\omega_l, k) \ x_2(\omega_l, k) \ \cdots \ x_M(\omega_l, k)]^T$  为  $M$  个阵元接收数据的向量表示,  $D$  个目标信源的频域向量为  $\mathbf{s}(\omega_l, k) = [s_0(\omega_l, k) \ s_1(\omega_l, k) \ \cdots \ s_{D-1}(\omega_l, k)]^T$ ,  $Q$  个干扰源的频域向量为  $\mathbf{i}(\omega_l, k) = [i_0(\omega_l, k) \ i_1(\omega_l, k) \ \cdots \ i_{Q-1}(\omega_l, k)]^T$ , 背景噪声的频域向量为  $\mathbf{n}(\omega_l, k) = [n_1(\omega_l, k) \ n_2(\omega_l, k) \ \cdots \ n_M(\omega_l, k)]^T$ 。上标 T 表示转置。远场目标源信号的导向矩阵为

$$\mathbf{A}(\omega_l, \phi) = \begin{bmatrix} \mathbf{a}(\omega_l, \phi_0) & \mathbf{a}(\omega_l, \phi_1) & \cdots & \mathbf{a}(\omega_l, \phi_{D-1}) \end{bmatrix}, \quad (3)$$

其中,  $\mathbf{a}(\omega_l, \phi_d)$  为目标信号导向向量。均匀线列阵的导向向量为

$$\mathbf{a}(\omega_l, \phi_d) = \begin{bmatrix} 1 & e^{-j\omega_l \tau_m(\phi_d)} & \cdots & e^{-j\omega_l \tau_M(\phi_d)} \end{bmatrix}^T, \quad (4)$$

$$d = 0, 1, 2, \dots, D-1.$$

近场干扰源信号的导向矩阵为

$$\mathbf{B}(\omega_l, \phi) = \begin{bmatrix} b(\omega_l, \phi_0) & b(\omega_l, \phi_1) & \cdots & b(\omega_l, \phi_{Q-1}) \end{bmatrix}, \quad (5)$$

其中, 导向向量为

$$b(\omega_l, \phi_d) = \begin{bmatrix} \rho_m(q)e^{-j\omega_l \mu_m(q)} & \cdots & \rho_M(q)e^{-j\omega_l \mu_M(q)} \end{bmatrix}^T, \quad (6)$$

$$q = 0, 1, 2, \dots, Q-1.$$

$K$  组采样数据估计得到的互谱密度矩阵为

$$\mathbf{R}_x(\omega_l) = \frac{1}{K} \sum_{k=1}^K \mathbf{x}(\omega_l, k)\mathbf{x}^H(\omega_l, k), \quad (7)$$

其中, 上标 H 表示共轭转置。结合互谱密度和导向矩阵, 可利用常规波束形成 (CBF) 计算目标的空间功率谱估计。

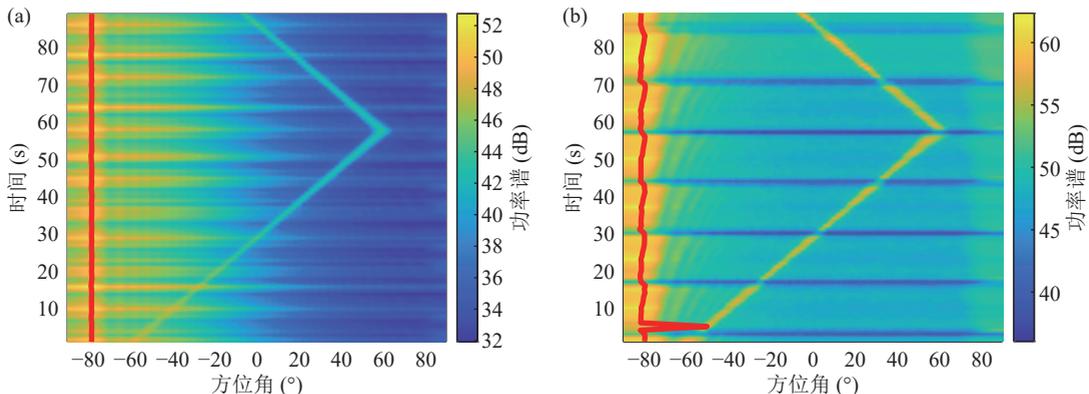


图2 幅度谱处理前后的 CBF 方位功率谱 (信噪比: 10 dB, 信干比: -15 dB) (a) 有干扰信号时; (b) 理想情况下仅恢复接收信号幅度谱

假设水下平台中有 48 阵元水声阵列, 间隔  $\Delta = 0.375$  m, 水中声速为 1500 m/s, 扫描角度为  $[-90^\circ, 90^\circ]$ , 目标和干扰源各一个。目标信号为周期 1 s 的线性调频信号 (LFM), 频带范围 1~2 kHz, 采样频率 4096 Hz, 目标在  $[-60^\circ, 60^\circ]$  范围往复匀速移动, 添加信噪比 10 dB 的白噪声  $n(t)$ , 干扰源在通道 1 下方两倍阵元间隔处 (0.75 m), 干扰信号使用 NoiseX-92 公开噪声数据库<sup>[24]</sup> 中的 Destroyer Engine 噪声, 信干比设置为 -25 dB。如图 2(a) 所示, 红色描线为 CBF 系统目标方向轨迹判断, 其受到近场干扰的不利影响。如果仅处理各通道接收信号幅度谱, 理想情况下恢复出目标信号幅度谱, 再结合接收信号的相位谱得到各通道干扰抑制后的信号频谱:

$$\mathbf{x}_{\text{干扰抑制后}}(\omega_l, k) = |\mathbf{S}(\omega_l, \phi)| e^{j\phi(\mathbf{x}(\omega_l, k))}, \quad (8)$$

其中,  $|\mathbf{S}(\omega_l, \phi)| = |\mathbf{A}(\omega_l, \phi)\mathbf{s}(\omega_l, k)|$  为各通道目标信号幅度谱,  $\phi(\cdot)$  表示相位谱计算。如图 2(b) 所示, CBF 方位谱估计依然受到严重影响。由此可见, 如果仅恢复接收信号幅度谱, 由于期望的目标信号互谱密度矩阵难以得到准确估计, 对目标信号开展方位谱估计时检测和跟踪效果不理想。

## 2 使用 DCCN 的干扰抑制方法

复数卷积神经网络层的前向传递函数定义为

$$\mathbf{x}_i^l = f \left( \sum_{i \in M_j} \mathbf{x}_i^{l-1} * \mathbf{k}_{ij}^l + \mathbf{b}_j^l \right), \quad (9)$$

其中,  $\mathbf{x}_i^l$  表示第  $l$  层的第  $j$  个输出, 首先从第  $l-1$  层的特征图中, 选择若干组成第  $l$  层输入特征图集合  $M_j$ ; 再利用复数卷积核  $\mathbf{k}_{ij}^l$  分别与输入中的每个特征图即  $\mathbf{x}_i^{l-1}$  进行卷积、求和以及增加偏置  $\mathbf{b}_j^l$ ; 最后使用复数激活函数  $f$  得到当前卷积层输出。

算法在处理阵列接收信号时使用了极坐标 (包括幅度和相位) 以及频谱的实部和虚部:

$$\mathbf{x}_r(\omega_l, k) = |\mathbf{x}(\omega_l, k)| \cos(\varphi(\mathbf{x}(\omega_l, k))), \quad (10)$$

$$\mathbf{x}_i(\omega_l, k) = |\mathbf{x}(\omega_l, k)| \sin(\varphi(\mathbf{x}(\omega_l, k))), \quad (11)$$

其中,  $|\mathbf{x}(\omega_l, k)|$ ,  $\varphi(\mathbf{x}(\omega_l, k))$ ,  $\mathbf{x}_r(\omega_l, k)$ ,  $\mathbf{x}_i(\omega_l, k)$  分别为分帧处理含干扰接收信号 STFT 的幅度谱、相位谱、实部谱和虚部谱。声信号处理中的特征表达如信号频谱等大多是复数形式。简单分离实部和虚部, 或者分别处理幅度谱和相位谱会丢失信号原本的内在特征。按照复数计算定义的复数卷积网络层, 对输入数据进行复数卷积、激活、批规范化等操作, 理论上会有更好的处理效果。当使用复数向量  $\mathbf{w} = \mathbf{w}_r + j\mathbf{w}_i$ ; 卷积复数矩阵  $\mathbf{x} = \mathbf{x}_r + j\mathbf{x}_i$  时, 复数卷积表达式定义如下:

$$\begin{bmatrix} \text{Real}(x * w) \\ \text{Imag}(x * w) \end{bmatrix} = \begin{bmatrix} \mathbf{x}_r & -\mathbf{x}_i \\ \mathbf{x}_i & \mathbf{x}_r \end{bmatrix} * \begin{bmatrix} \mathbf{w}_r \\ \mathbf{w}_i \end{bmatrix}, \quad (12)$$

其中,  $\mathbf{w}_r$  和  $\mathbf{w}_i$  分别为卷积核的实部与虚部参数。

首先, 对原始阵列接收信号进行预处理提取特征, 对多通道含干扰信号的重叠帧做短时傅里叶变换, 将幅度谱和相位谱同时作为深度复数卷积网络的输入, 充分利用时空频域信息; 然后, 借鉴传统声信号干扰抑制方法, 在复数卷积处理之前对幅度谱做一次常规卷积处理 (包括卷积、批规范化和激活), 降低了后续联合相位信息的实部谱和虚部谱处理运算量, 加快了网络的收敛速度。再按照复数运算规则卷积处理频谱的实部和虚部, 充分利用相位信息, 局部网络使用跨越连接进一步提高收敛速度; 最后, 通过使用典型噪声干扰数据训练网络, 用于实现多通道目标信号频谱重建。数据处理流程如图 3 所示。

在有监督的训练过程中, 将纯目标信号的频谱特征和网络输出的频谱数据之间的均方误差 (MSE) 作为网络的损失函数, 通过反向误差传播 (BP) 算法更新整个网络的权重和偏置量等参数, 损失函数定义如下:

$$\text{COST} = \frac{1}{2N} \sum_w [(\mathbf{x}_{\text{deNoise},r}(w) - \mathbf{x}_{\text{noNoise},r}(w))^2 + (\mathbf{x}_{\text{deNoise},i}(w) - \mathbf{x}_{\text{noNoise},i}(w))^2], \quad (13)$$

其中,  $N$  代表每一批次输入多通道频谱帧数。线阵列孔径越大即接收阵元通道数量越多, 则  $N$  值也越

大, 导致输入模型特征维度增加。一般来说输入特征维度越高, 模型的拟合能力会更强, 但也需要更多数据来训练模型, 同时也更容易产生过拟合。当输入维度过高时, 会出现维度灾难问题, 训练过程需要更多计算资源, 增大了模型训练难度。因此, 模型的构建和训练需要根据实际情况选择适合的输入维度。 $\mathbf{x}_{\text{deNoise},r}(w)$  和  $\mathbf{x}_{\text{deNoise},i}(w)$  分别表示 DCCN 输出经干扰抑制后信号频谱的实部和虚部,  $\mathbf{x}_{\text{noNoise},r}(w)$  和  $\mathbf{x}_{\text{noNoise},i}(w)$  分别表示目标信号频谱的实部和虚部。网络模型中学习率 (Lr)、权重衰减 (WeightDecay)、动量 (Momentum) 和批量大小 (Batchsize) 分别设置为  $1 \times 10^{-2}$ ,  $1 \times 10^{-4}$ , 0.9, 128, 在训练过程中, 采取小批量模式的随机梯度 (MSGD) 下降法实现并行化处理, 能够加快网络的收敛速度。

使用 DCCN 的多通道信号干扰抑制方法框图如图 4 所示。在训练阶段, 使用有监督的训练方法, 将含干扰的阵列接收信号样本进行特征提取作为训练数据集, 干净目标信号的频谱特征作为训练标签。通过向 DCCN 提供与目标以及干扰源相关联的特征来推断近场干扰源特征信息, 继而得到干扰源与各阵元之间的传递函数, 最优化求解各阵元的目标信号特征。在干扰抑制阶段, DCCN 能充分地学习出含干扰接收信号和干净目标信号之间的相互内在联系, 抑制其中的干扰分量, 输出信号用于目标方位功率谱估计。

## 3 仿真分析

### 3.1 典型噪声的干扰抑制

通过选取多种典型干扰噪声数据, 按照近场模型在多个信干比条件下进行仿真, 分析含噪阵列接收信号经 DCCN 方法干扰抑制后的 CBF 空间谱性能。基于图 1 的仿真模型, 假设接收阵列为 48 阵元的均匀线阵列; 频间非相干宽带处理的频率为 1~2 kHz, 共 251 个特征维度, 采样频率为 4096 Hz, 声速 1500 m/s, 阵元间隔为接收信号最高频率的半波长 (即 0.375 m)。线阵首部方向为  $90^\circ$ , 尾部方向为  $-90^\circ$ , 方向估计中的角度间隔为  $0.1^\circ$ 。假设在阵列接

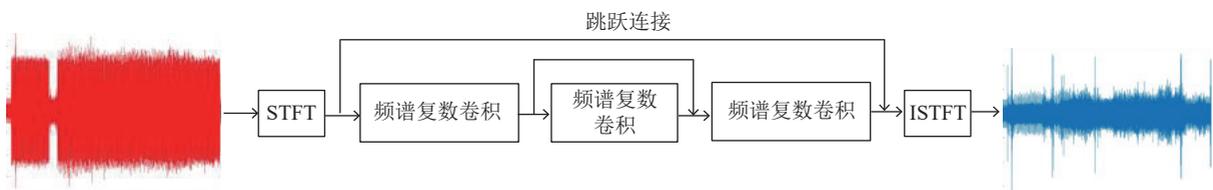


图 3 算法流程框图

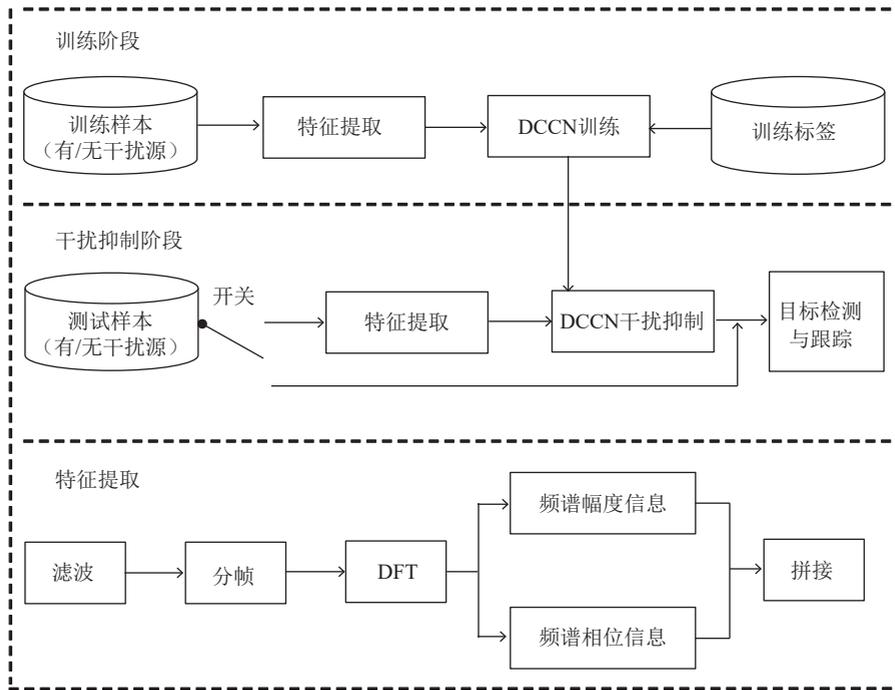


图4 使用DCCN的多通道干扰抑制系统框图

收信号期间有两个目标信源和一个干扰信源。目标信源采用线性调频信号(LFM),两个目标信源分别在 $[-60^\circ, 60^\circ]$ 和 $[-20^\circ, 40^\circ]$ 范围来回匀速移动。在NoiseX-92公开噪声数据库中分别选取Destroyer Engine、Destroyer Ops、M109和Buccaneer1四种类型的信号作为平台宽带自噪声干扰声源,每种类型各自生成信干比(单通道)为 $-15\text{ dB}$ 、 $-20\text{ dB}$ 、 $-25\text{ dB}$ 的干扰声源,干扰源位于尾部阵元(通道1)正下方两倍阵元间距处(0.75 m)。按照近场球面波传播特性计算幅度变化和相移后的各通道干扰信号。同时在各通道添加白噪声,信噪比(单通道)分别为 $-5\text{ dB}$ 、 $0\text{ dB}$ 、 $10\text{ dB}$ 。仿真处理中,快速傅里叶变换样本数为1024,频域采样次数为21,即每次方向估计使用的样本时长为 $1024 \times 21/4096=5.25\text{ s}$ 。

将上述干扰和噪声信号混入各通道目标信号生成多通道含干扰接收信号。单帧的STFT长为250 ms,帧移为125 ms。训练集共包含近30万帧(约20小时)数据,在训练集之外单独生成测试集,共包含7999帧含噪数据。测试集中信噪比为10 dB下的目标信号CBF方位功率谱估计如图5所示,作为干扰抑制前后性能对比分析使用。

使用训练完毕的DCCN网络模型对测试集中的仿真阵列接收数据开展干扰抑制,其中在干扰源信号为Destroyer Engine类型和信噪比为10 dB下的目标方位功率谱估计如图6—图8所示,红色描线为CBF系统目标方向轨迹判断。同时对比测试集中初

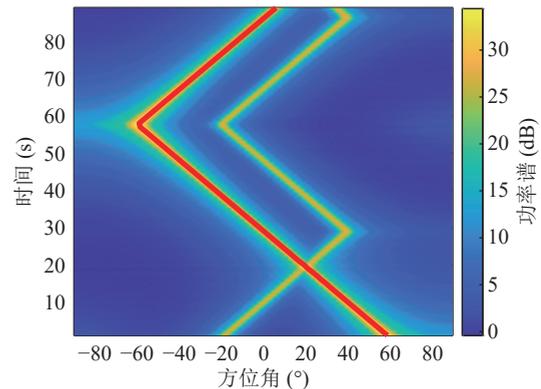


图5 无干扰信号时CBF方位功率谱(信噪比:10 dB)

始时间点处的无干扰、有干扰功率谱和干扰抑制后方位功率谱归一化结果,分析干扰抑制效果。

通过观察干扰抑制前后方位谱估计变化,当信干比为 $-15\text{ dB}$ 时,  $40^\circ\sim 90^\circ$ 阵列首部方向受干扰影响较小,尾部干扰源位置方向出现假目标方位,在 $-90^\circ\sim 20^\circ$ 方向范围出现较高旁瓣展宽,经过干扰抑制后旁瓣得到降低;当信干比为 $-20\text{ dB}$ 时, CBF方位谱估计受到严重干扰,  $-20^\circ$ 目标方位判断已失效,目标跟踪出现大量丢失。经过干扰抑制后,目标方向估计正确,跟踪正常;当信干比进一步降低为 $-25\text{ dB}$ 时,  $-20^\circ$ 方位的目标已经完全被强干扰方向的旁瓣掩盖,  $60^\circ$ 靠近阵列首部方位的波束旁瓣也受干扰影响而升高。经过干扰抑制后旁瓣得到有效降低,目标方位输出信噪比提高,目标检测和跟踪功能恢复正常。目标方向波束峰值 $Y_{\text{peak}}$ 与背景噪声的功

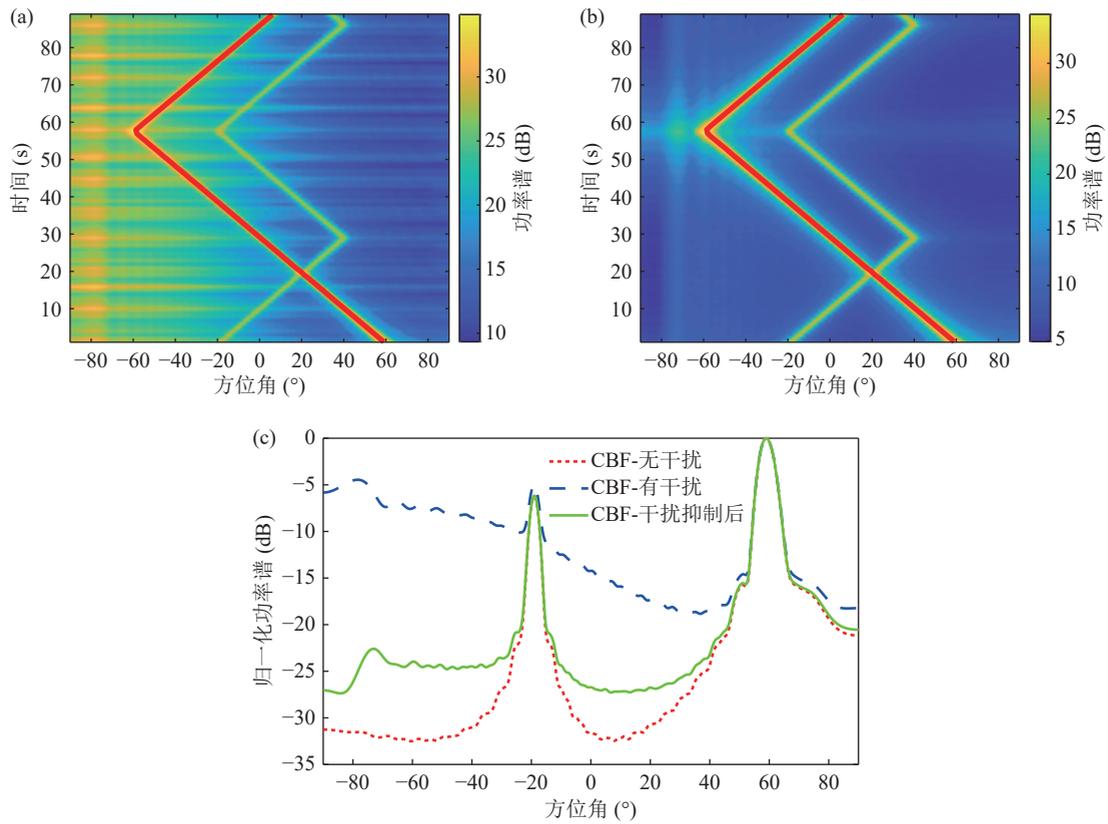


图6 干扰抑制前后 CBF 方位功率谱对比 (信噪比:10 dB, 信干比: -15 dB) (a) 有干扰信号时; (b) 干扰抑制后; (c) 初始时间点处无干扰、有干扰和干扰抑制后的归一化功率谱

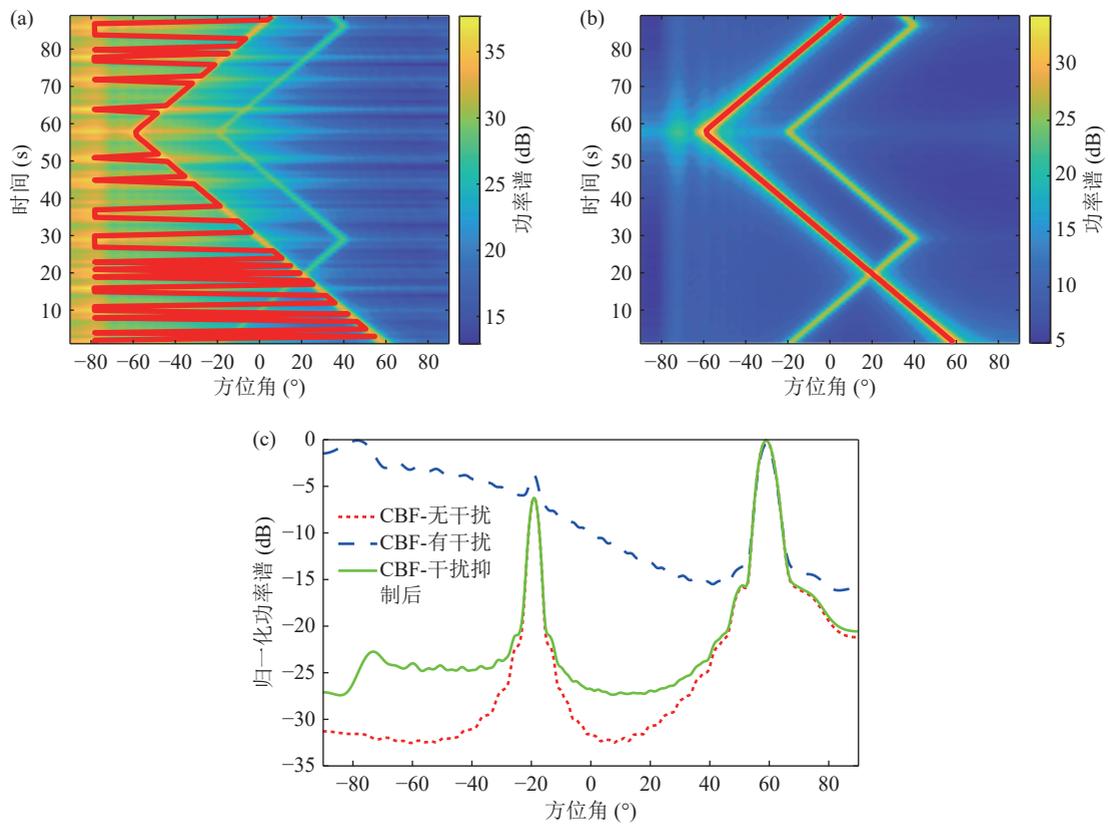


图7 干扰抑制前后 CBF 方位功率谱对比 (信噪比:10 dB, 信干比: -20 dB) (a) 有干扰信号时; (b) 干扰抑制后; (c) 初始时间点处无干扰、有干扰和干扰抑制后的归一化功率谱

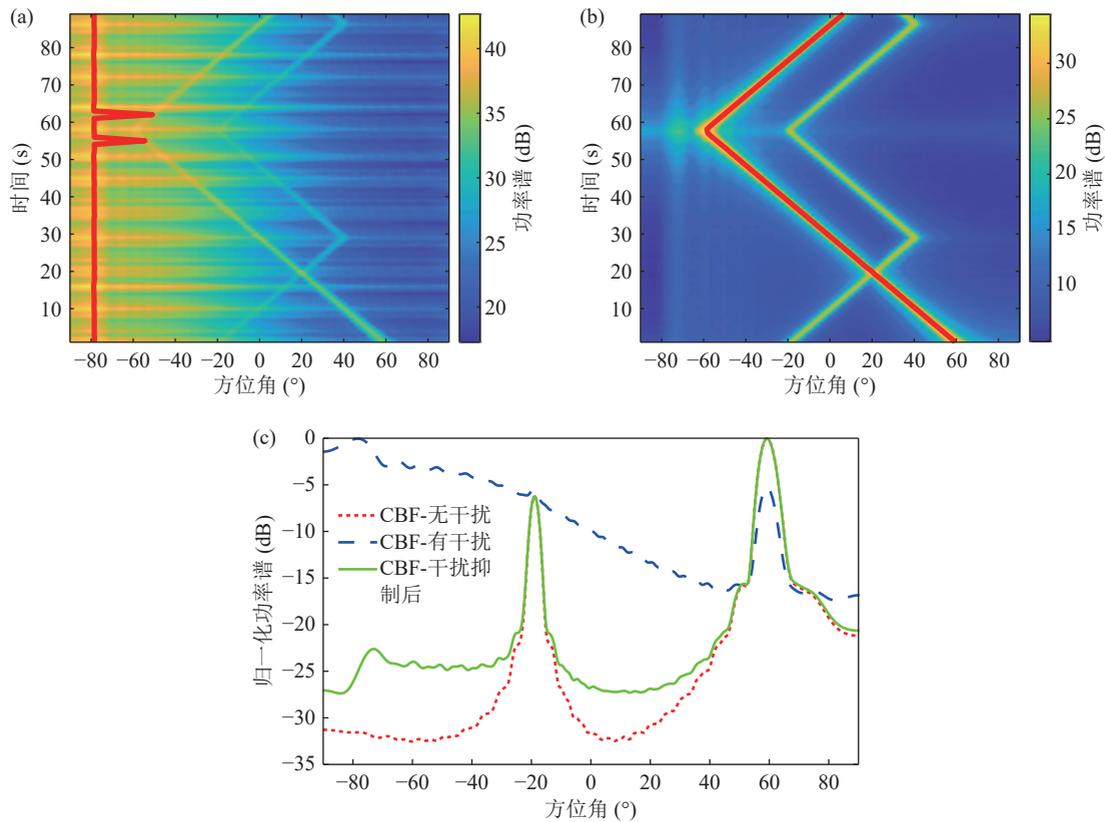


图 8 干扰抑制前后 CBF 方位功率谱对比 (信噪比:10 dB, 信干比: -25 dB) (a) 有干扰信号时; (b) 干扰抑制后; (c) 初始时间点处无干扰、有干扰和干扰抑制后的归一化功率谱

率谱  $Y_{\text{ground}}$  的差值  $Y_{\text{diff}}$  反映了对目标的检测和跟踪能力, 将干扰抑制前后的差值  $Y_{\text{diff}}$  变化作为检验干扰抑制效果的标志量。汇总各个仿真干扰源信号类型在信噪比为 -5 dB, 0 dB, 10 dB 下的干扰抑制效果, 如表 1 所示。仿真结果表明, DCCN 干扰抑制方法可以有效减少近场干扰信号的影响, 在降低干扰波束旁瓣的同时, 提高了波束形成对目标的检测能力。

在仿真实验中, 使用 DCCN 方法的干扰抑制效果受数据集质量和大小、阵元数量以及信干噪比的限制。在上述同等条件下, 使用 Destroyer Engine 类型噪声作为干扰源, 当信噪比为 -10 dB, 信干比分别为 -15 dB, -20 dB, -25 dB, 使用 DCCN 干扰抑制时, 初始时间处 -20° 目标方位干扰抑制效果不明显。

### 4 试验验证

为验证所提方法的有效性, 于 2020 年 9 月在千岛湖开展湖试。模拟水下平台和声源布放深度约 12 m, 水中声速约 1495 m/s, 收发位置平均水深 40 m, 距离约 1 km, 模拟水下平台侧面水平 0.3 m 等间隔布放 5 个阵元。发射船携带声源到达预定地点后关闭主机进入停车飘航状态, 并按计划吊放宽带声源以 1 s 为周期发射 1~2 kHz 线性调频信号 (LFM), 声源

表 1 峰值背景比效果对比表 (阵元数: 48, 目标方位: -20°, 单位: dB)

信噪比	信干比	干扰噪声类型			
		Destroyer Engine	Destroyer Ops	M109	Buccaneer1
-10	-15	1.5	1.1	0.8	0.9
	-20	2.8	2.6	2.1	2.2
	-25	3.1	3	3.2	3.3
-5	-15	4.2	5	4.8	4.9
	-20	5.4	5.9	5.5	5.9
	-25	7.3	7.5	7.1	7.2
0	-15	6.2	6.5	6.1	6.4
	-20	8.4	8.7	8.4	8.5
	-25	10.8	10.2	9.8	10.1
10	-15	10.3	13.1	12.8	11.2
	-20	15.2	14.5	14.1	13.8
	-25	18.1	18.2	17.9	17.8

级约 160 dB。放置一台可调振动干扰声源于模拟水下平台首部附近, 以 0.5 s 为周期发射 2~1 kHz (从高到低) LFM 干扰信号, 信干比约为 -5 dB。数据处理的频率范围为 1~2 kHz, 共 257 个频点, 采样频率为 4000 Hz。一次处理的数据长度约为 5.25 s, 每 250 ms 进行一次频域采样, 即频域采样次数为 21, 方位搜索的角度间隔为 1°。实测水中声速剖面如图 9 所示,

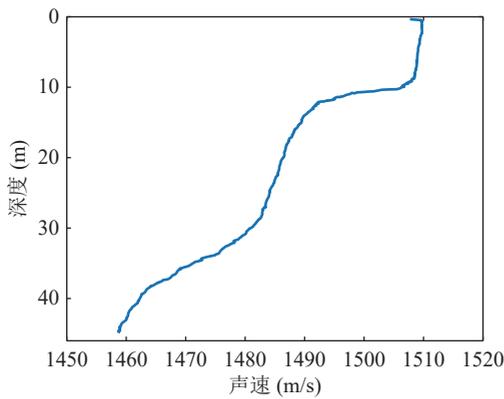


图 9 实测湖试水下声速剖面图

在深度 5~15 m 之间存在一个温跃层, 为典型的夏季浅海(湖)水文条件。

在实际场景中使用 DCCN 时, 首先设计和建立 DCCN 干扰抑制网络, 之后在无目标信号存在时采样各通道的纯干扰信号作为本平台干扰特征数据, 然后在其中各通道混入不同方向和信干比的单目标仿真 LFM 信号作为数据集, 最后训练完成针对此次实验环境的 DCCN 干扰抑制网络。干扰抑制前后的 CBF 方位功率谱估计结果如图 10 所示。红色线条表示目标方向轨迹判断, 图 10(a)(b) 分别为搭载接收线阵的水下平台在静止和运动状态下的 CBF 结果,

处理时长分别为 48 s 和 68 s。受近场干扰的影响, 波束旁瓣升高, 干扰抑制前的 CBF 结果难以得到可靠的目标方位, 目标没有被正确检测。由图 10(c)(d) 可见, 经 DCCN 干扰抑制后, 干扰造成的高旁瓣被有效抑制, 提高了输出信干噪比, CBF 系统实现了正确的目标检测和稳定的目标跟踪, 测试时间段内湖试目标方位估计平均误差为 9.8°。试验过程中水下平台由手动操作旋转模拟目标运动, 在操作转动过程中线阵处于运动状态并产生噪声。由于在波束形成积分时间内出现方位变化, 且此次湖上初步实验有效实验样本数量有限, 阵元数较少造成输入数据维度偏小, 阵元位置可能存在误差, 以及数据处理中忽略了多途效应等, 限制了 DCCN 干扰抑制方法性能, 上述多种因素导致 CBF 方位谱估计存在不同程度的旁瓣展宽。

以上实验结果表明, 在近场强干扰的实际使用环境中, 使用 DCCN 的多通道干扰抑制方法可以自适应地有效抑制近场强干扰, 为获得可靠的目标方位估计和目标跟踪提供了一种有效的技术途径。同等条件下, 使用更多样的目标信号、采集更多高质量的平台自噪声干扰信号特征扩大训练数据集, 理论上可进一步提高 DCCN 干扰抑制方法的稳健性和泛

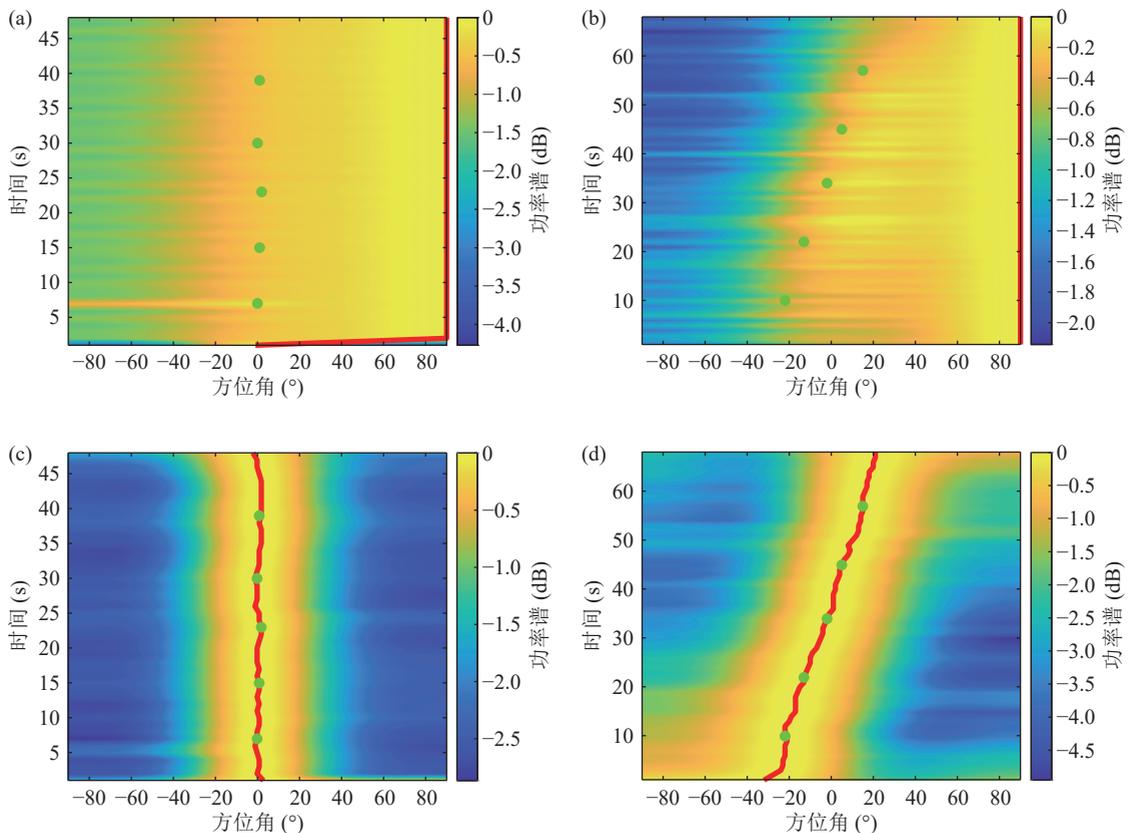


图 10 两种状态干扰抑制前后的 CBF 方位功率谱对比(实际目标方位经 GPS 记录如绿色圆点所示) (a) 静止状态; (b) 运动状态; (c) 静止状态干扰抑制后; (d) 运动状态干扰抑制后

化性。

## 5 结论

针对在近场低信干比条件下的水下目标检测与跟踪问题,提出了一种适用于多通道水听器阵列的深度学习宽带背景噪声干扰抑制方法。该方法提取和级联包含空间信息的各通道接收信号频域特征作为输入,利用深度复神经网络对信号频域复数特征非线性映射关系的强建模能力,通过学习大量含干扰声信号和纯净目标声信号多通道样本数据,构造出含干扰声信号特征和能够表示纯净目标声信号的训练目标之间的一个复杂非线性函数。仿真和湖上实验结果表明,该方法可以有效抑制平台背景宽带自噪声干扰,在低信干比条件下提升了 CBF 方位谱估计的目标检测和跟踪性能。

在实际场景中使用 DCCN 方法的干扰抑制效果受到数据集质量、信噪比、阵元数量(输入维度)、阵元位置误差、目标方位与干扰源相对平台位置以及多途等因素限制。本文模型目前只考虑了干扰源的直达声,且局限于试验样本数量偏少,实际复杂水声环境中的多途、混响以及多个干扰源等的干扰抑制问题需进一步研究。

## 参 考 文 献

- Cox H, Zeskind R, Owen M. Robust adaptive beamforming. *IEEE Trans. Acoust. Speech Signal Process.*, 1987; **35**(10): 1365–1376
- Niu H, Gong Z, Ozanich E, et al. Deep-learning source localization using multi-frequency magnitude-only data. *J. Acoust. Soc. Am.*, 2019; **146**(1): 211–222
- Boll S F. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Trans. Acoust. Speech Signal Process.*, 1979; **27**(2): 113–120
- Chen J, Benesty J, Huang Y, et al. New insights into the noise reduction Wiener filter. *IEEE Trans. Acoust. Speech Signal Process.*, 2006; **14**(4): 1218–1234
- Ephraim Y, Malah D, et al. Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator. *IEEE Trans. Acoust. Speech Signal Process.*, 1984; **32**(6): 1109–1121
- Ephraim Y, Trees H V. A signal subspace approach for speech enhancement. *IEEE Trans. Speech Audio Process.*, 1995; **3**(4): 251–266
- 李超, 刘文学. 基于 F 范数的信号子空间维度估计的多通道语音增强算法. *声学学报*, 2005; **33**(10): 1786–1789
- 任岁玲, 葛凤翔, 郭良浩. 基于特征分析的自适应干扰抑制. *声学学报*, 2013; **38**(3): 272–280
- 蒋国健, 任克明. 噪声抵消法估计和抑制声呐部位主要自噪声. *声学学报*, 1996; **21**(4): 289–296
- Chang K M, Liu S H. Gaussian noise filtering from ECG by Wiener filter and ensemble empirical mode decomposition. *J. Signal Process. Syst.*, 2011; **64**(2): 249–264
- Wang D L. On ideal binary mask as the computational goal of auditory scene analysis. In: Divenyi P (ed). *Speech separation by humans and machines*. Springer, New York, 2005: 181–197
- Lin K W E, Balamurali B T, Koh E, et al. Singing voice separation using a deep convolutional neural network trained by ideal binary mask and cross entropy. *Neural Comput. Appl.*, 2020; **32**(4): 1037–1050
- Narayanan A, Wang D. Ideal ratio mask estimation using deep neural networks for robust speech recognition. *IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, Vancouver, BC, Canada, 2013: 7092–7096
- Wang Y, Narayanan A, Wang D L. On training targets for supervised speech separation. *IEEE/ACM Trans. Audio Speech Lang. Process.*, 2014; **22**(12): 1849–1858
- Xu Y, Du J, Dai L R, et al. An experimental study on speech enhancement based on deep neural networks. *IEEE Signal Process. Lett.*, 2014; **21**(1): 65–68
- Xu Y, Du J, Dai L R, et al. A regression approach to speech enhancement based on deep neural networks. *IEEE/ACM Trans. Audio Speech Lang. Process.*, 2015; **23**(1): 7–19
- Paliwal K, Wójcicki K, Shannon B. The importance of phase in speech enhancement. *Speech Commun.*, 2011; **53**(4): 465–494
- Trabelsi C, Bilaniuk O, Zhang Y, et al. Deep complex networks. *The International Conference on Learning Representations*, Vancouver, Canada, 2018
- Pandey A, Wang D L. Exploring deep complex networks for complex spectrogram enhancement. *IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, Brighton, United Kingdom, 2019: 6885–6889
- Erdogan H, Hershey J R, Watanabe S, et al. Phase-sensitive and recognition-boosted speech separation using deep recurrent neural networks. *IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, South Brisbane, Queensland, Australia, 2015: 708–712
- Williamson D S, Wang Y, Wang D L. Complex ratio masking for monaural speech separation. *IEEE/ACM Trans. Audio Speech Lang. Process.*, 2016; **24**(3): 483–492
- Lee J, Kang H G. A joint learning algorithm for complex-valued T-F masks in deep learning-based single-channel speech enhancement systems. *IEEE/ACM Trans. Audio Speech Lang. Process.*, 2019; **27**(6): 1098–1108
- Yin D, Luo C, Xiong Z, et al. PHASEN: A phase-and-harmonics-aware speech enhancement network. *Proceedings of the AAAI Conference on Artificial Intelligence*, New York, USA, 2020; **34**(5): 9458–9465
- Varga A, Steeneken H J. Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems. *Speech Commun.*, 1993; **12**(3): 247–251