

波束域特征融合的浅海水平阵目标方位估计*

王珍珠^{1,2,3} 王文博^{1,2†} 李 赫^{1,2} 任群言^{1,2,3} 郭圣明^{1,2}

(1 中国科学院声学研究所 北京 100190)

(2 中国科学院水声环境特性重点实验室 北京 100190)

(3 中国科学院大学 北京 100049)

2023 年 2 月 27 日收到

2023 年 5 月 19 日定稿

摘要 针对实际海洋环境影响下水平阵目标方位估计性能降低的问题,提出了一种基于水平阵波束域特征融合的卷积神经网络方位估计方法。首先建立数据计算仿真模型用于生成数据集,利用常规波束形成和最小方差无失真响应两种波束形成器得到波束域特征数据。然后,利用两种波束域特征数据分别训练卷积神经网络模型,仿真和海试数据均表明,训练后的卷积神经网络模型方位估计精度优于常规波束形成、最小方差无失真响应和匹配波束处理,尤其是在端射方向性能提升明显。最后,将两种波束域特征进行前端特征融合后再训练卷积神经网络模型,海试数据测试结果表明,特征融合后的卷积神经网络模型方位估计性能进一步提升,在 5°误差范围内的可靠测向概率比单一特征卷积神经网络估计结果高约 4%,均方根误差降低约 0.2°。

关键词 卷积神经网络, 水平阵, 波达方向估计, 波束域, 特征融合

PACS 数 43.30, 43.60

DOI: 10.12395/0371-0025.2023029

Direction of arrival estimation using a horizontal array with beamforming domain feature fusion in shallow water

WANG Zhenzhu^{1,2,3} WANG Wenbo^{1,2†} LI He^{1,2} REN Qunyan^{1,2,3} GUO Shengming^{1,2}

(1 Institute of Acoustics, Chinese Academy of Sciences Beijing 100190)

(2 Key Laboratory of Underwater Acoustic Environment, Chinese Academy of Sciences Beijing 100190)

(3 University of Chinese Academy of Sciences Beijing 100049)

Received Feb. 27, 2023

Revised May 19, 2023

Abstract A convolutional neural network method for direction of arrival estimation based on beam domain feature fusion of the horizontal line array is proposed to solve the problem of performance degradation of direction of arrival in marine environment. Firstly, a data computational simulation model is built for generating the dataset, and beam-domain data are obtained using both conventional beamforming and minimum variance distortionless response beam formers. Then, the convolutional neural network models are trained separately using the two beam-domain data. Both simulation and sea trial data processing results confirm that the direction of arrival estimation accuracy of the trained models outperforms that of conventional beamforming, minimum variance distortionless response and matched beam processing, especially in the end-fire direction where the performance improvement is obvious. Finally, the two beam-domains features are used to train the convolutional neural network by early-fusion. The sea trial data test results confirm that the direction of arrival estimation performance of the feature-fused convolutional neural network is further improved. The probability of reliable direction estimation within 5° error range is about 4% higher than that the single-feature convolutional neural network estimation result, and the root mean square error is reduced by about 0.2°.

Keywords Convolutional neural network, Horizontal line array, Direction of arrival estimation, Beam-domain, Feature fusion

* 国家自然科学基金项目 (12204506) 和中国科学院声学研究所自主部署项目 (MBDX202102) 资助

† 通讯作者: 王文博, wangwenbo@mail.ioa.ac.cn

引言

波束形成是实现目标波达方向 (DOA) 估计的关键技术^[1-2], 较为常用的波束形成方法包括常规波束形成 (CBF)^[3]、基于数据自适应的最小方差无失真响应 (MVDR)^[4]、匹配波束处理 (MBP)^[5]。CBF 与 MVDR 获取波束谱的谱峰位置确定来波方向, 端射方向性能下降明显。MBP 与声场特性相关, 对端射方向的估计偏差有修正作用, 但拷贝场模型失配会导致性能严重不足。因此, 传统 DOA 估计算法亟需进一步优化。

近年来, 人工智能技术已成为各领域关注焦点^[6-8], 在水声探测、定位等领域得到了广泛应用^[9-11]。2020 年, Elbir^[12] 设计了一种 CNN 与多重信号分类 (MUSIC) 算法结合的信号 DOA 估计框架, CNN 模型的输入是阵列接收到数据的协方差矩阵, 通过学习相应角度子区域的 MUSIC 谱进行 DOA 估计。2021 年, Liu 等^[13] 以接收阵列中水听器个数为目标设计多种 CNN 模型, 输入为实数与虚数协方差矩阵, 实现了目标方位的高效识别。2021 年, 曹怀刚^[14] 将 CNN 应用于单个矢量传感器目标 DOA 估计, 实现了多目标分辨。综上, 深度学习在目标定位研究中, 重点关注学习算法、数据集、特征选择和实际应用等。卷积神经网络 (CNN)^[15] 模型是基于共享卷积核的一种深度神经网络模型, 常规波束形成技术表征 DOA 信息具有一定的优势, 将 CNN 与常规 DOA 估计方法结合有望提高 DOA 估计性能, 而相比于网络本身, 数据集与特征选取往往影响更大。

基于深度学习的 DOA 估计模型多使用单一目标特征样本作为输入, 然后针对单一特征设计回归器或分类器, 目标的不同特征量对 DOA 特性有不同程度的描述, 这些特征量可能存在一致性和互补性。模式识别领域中, 特征融合技术已广泛应用, 水下目标识别领域特征融合策略也获得了很好的效果。2018 年, Zhang 等^[16] 提出了一种基于长短期记忆网络 (LSTM) 的多特征融合的目标识别方法, 融合了时域时间序列数据、频谱特征数据和 Mel 倒谱系数特征数据三种特征, 对水下目标辐射噪声识别的效果优于单一特征或者组合两种特征的分类效果。2020 年, Ke 等^[17] 应用二维特征融合技术进行水下目标识别, 显著提高了识别精度。上述水声目标识别任务, 通过对目标的不同特征进行融合处理, 实现了特征间的表征能力互补或者增强特征表达。将多种

特征进行联合表征, 目的是学习到更多关于目标的综合信息, 实现分类或者回归的学习任务^[18]。

本文首先探讨了传统 DOA 估计算法的特点及不足, 给出数值仿真示例分析; 然后结合 CNN, 引入特征融合技术, 利用水平阵实现了目标 DOA 估计, 提出了 CNN-CBF、CNN-MVDR 和基于 CBF 和 MVDR 两种特征融合的 CNN-II 三种估计算法。针对 CBF 和 MVDR 两种波束域特征展开了数值仿真分析, 在不同阵元间隔及不同信噪比下讨论 CNN 模型的 DOA 估计性能; 利用海试数据对本文算法进行了性能评价。

1 水平线列阵接收信号模型

首先建立单条水平线阵列接收信号模型, N 个水听器以间隔 d 均匀分布。假设单声源情况, 中心频率为 f 的信号 $s(t)$ 以波方向 θ 入射到线列阵上, 信号满足远场窄带假设条件, 信号包络在通过阵列期间不会发生改变, 设 $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_N(t)]^T$ 为各个阵元接收信号, 一般包含经信道传播后的声源信号和噪声信号, 则阵列接收信号模型可以表示为

$$\mathbf{x}(t) = \mathbf{a}(f, \theta)s(t) + \mathbf{n}(t), \quad (1)$$

其中, $\mathbf{n}(t)$ 为噪声矩阵, $\mathbf{a}(f, \theta)$ 为阵列流形向量矩阵:

$$\mathbf{a}(f, \theta) = [\exp(-jk p_1), \exp(-jk p_2), \dots, \exp(-jk p_N)]^T, \quad (2)$$

其中, p_i 为阵元位置, $i = 1, \dots, N$, $\mathbf{k}$ 为波数向量, $\mathbf{k} = (2\pi f/c)\mathbf{v}(\theta)$, $\mathbf{v}(\theta)$ 为声信号传播方向的单位向量, c 为信号的传播速度。

实际中假设 $\mathbf{x}$ 是遍历的, 数据的互谱密度矩阵 $\mathbf{R}$ 通常使用一段快拍样本进行估计, 多快拍估计下互谱密度矩阵可以表示为

$$\widehat{\mathbf{R}} = \frac{1}{M} \sum_{m=1}^M \mathbf{x}_f^H(m) \mathbf{x}_f(m), \quad (3)$$

其中, M 表示数据的频域快拍数, $\mathbf{x}_f(m)$ 为每个快拍下 N 个阵元接收声场信号的频域数值, $\mathbf{H}$ 表示共轭转置。

2 基于波束形成处理和匹配场处理的 DOA 估计

2.1 波束形成处理和匹配场处理理论

2.1.1 CBF 算法

CBF 的基本原理为: 对线列阵每个阵元所处位置的接收信号进行处理, 按照指定方向作一定的时

延补偿, 使期望方向上信号加强, 并抑制噪声和干扰。CBF 扫描波束谱为

$$\mathbf{P}_{\text{CBF}}(f, \theta) = \mathbf{w}_{\text{CBF}}^H(f, \theta) \widehat{\mathbf{R}} \mathbf{w}_{\text{CBF}}(f, \theta), \quad (4)$$

其中, $\mathbf{w}_{\text{CBF}}(f, \theta) = \mathbf{a}(f, \theta) / \sqrt{N}$ 为 CBF 的加权向量。

波束谱最大模值对应的角度 $\hat{\theta}_c$ 即为 DOA 估计角度:

$$\hat{\theta}_c = \arg \max_{\theta_c = \theta_1, \theta_2, \dots, \theta_{L1}} (|\mathbf{P}_{\text{CBF}}(f, \theta)|), \quad (5)$$

其中, $\theta_c = \theta_1, \theta_2, \dots, \theta_{L1}$ 为波束谱扫描的角度范围。

当目标来波方向为 θ_0 时, 衡量 DOA 分辨率的波束宽度为

$$\text{BW}_{0.5} \approx 0.886 \frac{\lambda}{Nd \cos(\theta_0)}, \quad (6)$$

其中, λ 为波长。

从式 (6) 可以看出, 对于多目标分辨, 若信号频率相同, 入射到阵列的 DOA 的间隔小于波束宽度, CBF 将不能准确分辨目标实际的 DOA, 这是 CBF 存在的固有瑞利准则局限性, 一般通过增加阵元间隔和增加阵元个数提高角度分辨率, 在实际应用中阵列孔径的增加有所限制。

2.1.2 MVDR 对角加载算法

MVDR 的基本原理为: 使得期望方向上的信号具有无失真输出的同时最小化噪声方差, 且在理想状态下旁瓣更窄, 主瓣更高, 但稳健性不足。基于对角加载的 MVDR 是指在样本互谱密度矩阵上添加对角矩阵项, 在模型失配情况下, 可以提高 MVDR 的输出性能。对角加载 MVDR 扫描波束谱为

$$\mathbf{P}_{\text{MVDR}}(f, \theta) = \widetilde{\mathbf{w}}_{\text{MVDR}}^H(f, \theta) \widehat{\mathbf{R}} \widetilde{\mathbf{w}}_{\text{MVDR}}(f, \theta), \quad (7)$$

其中

$$\widetilde{\mathbf{w}}_{\text{MVDR}}(f, \theta) = \frac{(\widehat{\mathbf{R}} + \rho \mathbf{I})^{-1} \mathbf{a}(f, \theta)}{\mathbf{a}^H(f, \theta) (\widehat{\mathbf{R}} + \rho \mathbf{I})^{-1} \mathbf{a}(f, \theta)}$$

为对角加载后的 MVDR 的加权向量, ρ 为对角加载系数, ρ 的选取一般没有严格的标准, $\mathbf{I}$ 为单位矩阵。

波束谱最大模值对应角度 $\hat{\theta}_M$ 即为 DOA 估计角度:

$$\hat{\theta}_M = \arg \max_{\theta_M = \theta_1, \theta_2, \dots, \theta_{L2}} (|\mathbf{P}_{\text{MVDR}}(f, \theta)|), \quad (8)$$

其中, $\theta_M = \theta_1, \theta_2, \dots, \theta_{L2}$ 为波束谱扫描的角度范围。

2.1.3 MBP 算法

MBP 进行目标 DOA 估计的基本原理为: 将阵列实际接收数据与所有可能的目标来波方向上产生的拷贝场数据的常规波束响应进行匹配。MBP 需要

已知声场先验知识, 通过参数不断调整寻求最优解或通过提高信号处理器的性能缓解环境失配问题, 但增加了计算量。MBP 的目标函数为

$$\mathbf{B}(s; \alpha|_{\theta=\theta_l}) = \frac{\left| \int \mathbf{B}_{\text{CBF}}^*(s; \alpha|_{\theta=\theta_l}) \mathbf{B}_{\text{CBF}}(\text{data}; \alpha|_{\theta=\theta_k}) d \cos \alpha \right|^2}{\left[\int |\mathbf{B}_{\text{CBF}}(s; \alpha|_{\theta=\theta_l})|^2 d \cos \alpha \right] \left[\int |\mathbf{B}_{\text{CBF}}(\text{data}; \alpha|_{\theta=\theta_k})|^2 d \cos \alpha \right]}, \quad (9)$$

其中, $\mathbf{B}_{\text{CBF}}(\text{data}; \alpha|_{\theta=\theta_k})$ 和 $\mathbf{B}_{\text{CBF}}(s; \alpha|_{\theta=\theta_l})$ 分别为测量场和拷贝场的波束响应, θ_k 为目标的来波方向, θ_l 为设置的匹配波束扫描方向, θ_l 遍历某个 DOA 可能的范围, α 是计算 CBF 波束响应时扫描的角度范围, 在本文 α 取值范围为 $1^\circ \sim 180^\circ$, 取值间隔为 1° 。

根据 $\mathbf{B}(s; \alpha|_{\theta=\theta_l})$ 即可确定目标方向 $\hat{\theta}_{\text{MBP}}$:

$$\hat{\theta}_{\text{MBP}} = \arg \max_{\theta_{\text{MBP}} = \theta_1, \dots, \theta_l, \dots, \theta_{L2}} (|\mathbf{B}(s; \alpha|_{\theta=\theta_l})|), \quad (10)$$

其中, $\theta_{\text{MBP}} = \theta_1, \dots, \theta_l, \dots, \theta_{L2}$ 为匹配波束扫描的角度范围。

2.2 水平线列阵 DOA 估计数值仿真分析

在典型的浅海环境下, 接收阵为水平线列阵时, 对 CBF、MVDR 对角加载以及 MBP 三种算法的 DOA 估计性能进行数值仿真分析。

2.2.1 仿真条件

声源频率为 100 Hz, 声源深度为 20 m, 单条水平线列阵 (32 阵元, 阵元间隔 1.5 m) 深度为 50 m, 目标声源距离水平阵参考阵元的距离为 4.5 km。采用两层海底模型, 设置声场环境参数为: 水深为 100 m, 沉积层厚度为 6 m, 沉积层和基底层声速分别为 1600 m/s 和 1800 m/s, 沉积层和基底层密度分别为 1.7 g/cm^3 和 1.8 g/cm^3 , 沉积层和基底层衰减均为 $0.5 \text{ dB}/\lambda$, 海水声速为 1500 m/s, DOA 扫描范围为 $1^\circ \sim 180^\circ$, 扫描角度间隔为 1° 。利用声场仿真软件 KRAKENC 进行声场数据仿真。沉积层和基底层参数范围设定时参照文献 [19] 1.6 节中的表 1.3。

定义本节三个性能评估参数: DOA 估计准确度 Acc_5° (预测 DOA 与真实 DOA 的绝对误差小于等于 5° 的百分率)、波束谱主瓣宽度 $B_{-3 \text{ dB}} = |\theta_{-3}^+ - \theta_{-3}^-|$ (波束谱中的主瓣对应的波束谱强度下降到 -3 dB 时对应的两方向间的夹角, θ_{-3}^+ 表示主瓣波束谱下降到 -3 dB 时主瓣右侧对应角度, θ_{-3}^- 表示主瓣波束谱下降到 -3 dB 时主瓣左侧对应角度) 和水平接收阵列输出主旁瓣比 G_{HLA} (主瓣峰值与第一旁瓣峰值对应波束谱强度之差)。

2.2.2 仿真结果

使用 MBP 时, 拷贝场参数设置严格与测量场仿

真参数一致, DOA 扫描角度范围为 $1^\circ \sim 180^\circ$, 扫描角度间隔为 1° 。CBF、MVDR 对角加载 (对角加载系数为 0.01)、MBP 处理结果分别如图 1—图 3 所示。

分析非端射方向主瓣宽度和阵输出主旁瓣比。当来波方向为 60° , 90° , 120° , 150° 时, CBF、MVDR 和 MBP 处理结果的波束谱扫描图如图 4 所示。波方向为 60° , 90° , 120° , 150° 时的 $B_{-3\text{dB}}$ 和 G_{HLA} 数值分别列于表 1 和表 2。

从图 1—图 3 的方位历程图可明显看出, 对比三种方法的主瓣宽度, MVDR 有最窄的主瓣, CBF 和 MBP 的主瓣相当。MBP 方法的 Acc_{5° 为 100%, CBF 和 MVDR 的 Acc_{5° 均为 90.56%, CBF 和 MVDR 估计方法在端射方向失效, 出现较大 DOA 估计偏差, MBP 方法性能仍然保持精准, 这是因为端射方向波束输出的宽度增加导致角度分辨能力下降, 匹配波束在理想的模拟仿真下, 通过全声场处理可以有效

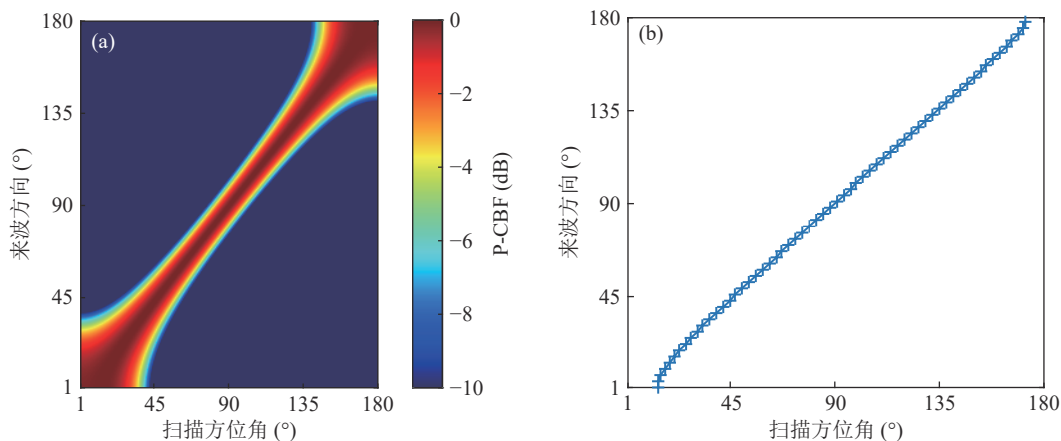


图 1 同一仿真条件下 CBF 处理结果 (a) 波束响应输出; (b) DOA 估计值

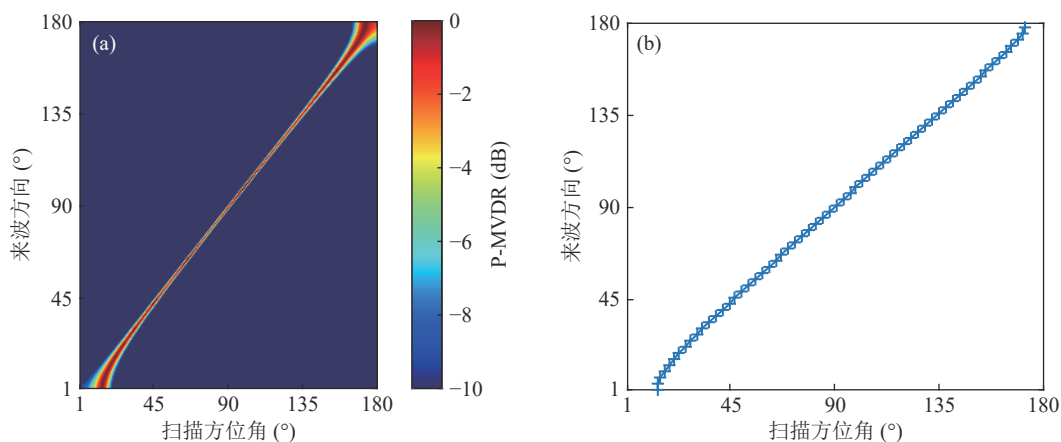


图 2 同一仿真条件下 MVDR 处理结果 (a) 波束响应输出; (b) DOA 估计值

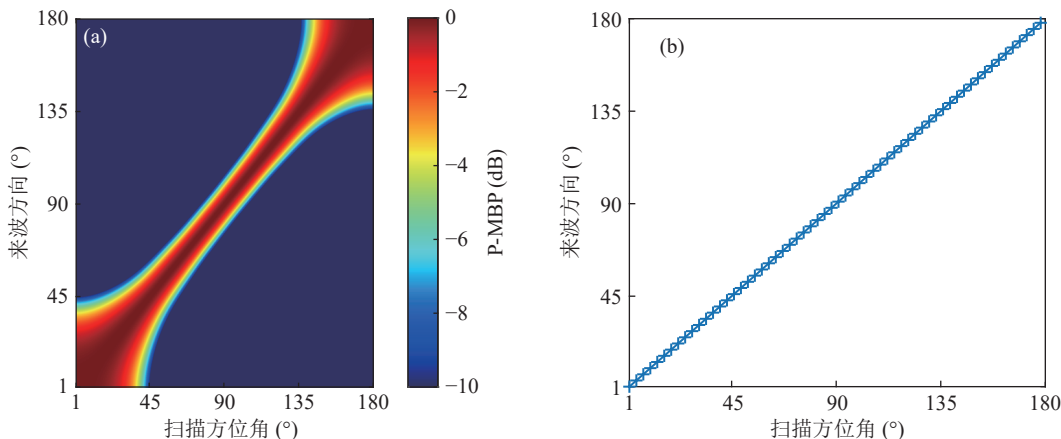


图 3 同一仿真条件下 MBP 处理结果 (a) 波束响应输出; (b) DOA 估计值

解决端射方向 DOA 估计问题。在本节设置仿真参数下,从表 1 可以看出 MBP 对应主瓣宽度稍大于 CBF, MBP 方法在保证 DOA 估计精度的同时损失了很小的阵输出主旁瓣比。

分析三种算法的原理, CBF 和 MVDR 在浅海水平阵下的应用与环境参数关联性不强,并不需要提前获得声场环境参数的先验知识,直接利用线性加权可以计算出波束谱,直接根据波束谱峰值对应角度确定估计的来波方向。MBP 技术从原理来看可以视为一个相关处理器,基于物理原理,在已知测量场相关信息后,可以准确进行 DOA 估计。具体实测数据下性能需要进一步分析。

3 基于深度学习算法的水声目标 DOA 估计

3.1 基于卷积神经网络的水声目标 DOA 估计

3.1.1 基于卷积神经网络的水声目标 DOA 估计流程

图 5 是本文将卷积神经网络方法应用于水声目标 DOA 估计的流程图。具体实现过程可以分为五

个环节: 声场数据仿真、波束形成处理、数据集预处理、网络模型训练、DOA 估计。

3.1.2 基于特征融合的水声目标 DOA 估计流程

特征融合方式可以分为基于特征级的前端融合、中间融合,以及基于决策级的后端融合。前端融合指在网络的样本输入层上做相关融合操作,最简单的就是将两个特征矩阵进行垂直连接;中间融合是指先将不同特征通过神经网络转变为高维特征表达,在网络模型的隐含层做融合操作,再去训练实现分类或回归;后端融合指在网络输出层融合,针对不同的特征分别训练网络模型,针对模型输出的预测结果,将不同的预测的结果通过某些融合机制进行融合。

本文进行简单的尝试,将前端特征融合应用到水下目标 DOA 估计中。3.1.1 节所述的 DOA 估计流程是指通过对阵列接收数据进行波束形成处理,再利用 CNN 进行目标 DOA 估计,输入到网络中的数据特征特征是单一波束域特征。为了学习到更多关于目标 DOA 的相关信息,通过不同的波束形成处理得到关于目标的两种波束域特征量,采用统一的数据

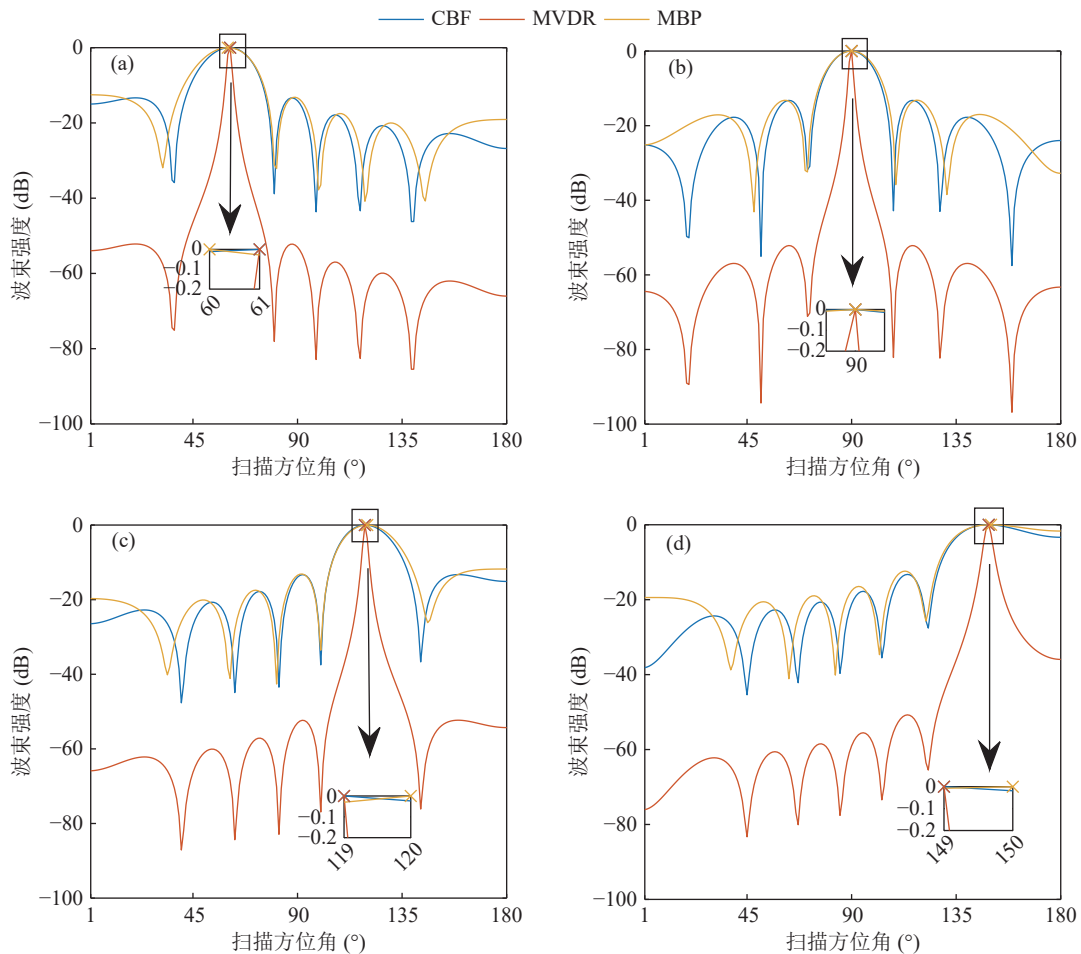


图 4 不同来波方向经 CBF、MVDR 和 MBP 处理后的扫描波束谱 (a) 60°; (b) 90°; (c) 120°; (d) 150°

表 1 波束谱主瓣宽度

入射角 (°)	CBF- $B_{-3\text{dB}}$ (°)	MVDR- $B_{-3\text{dB}}$ (°)	MBP- $B_{-3\text{dB}}$ (°)
60	18.37	1.60	20.11
90	15.87	1.39	16.81
120	18.29	1.35	19.47
150	37.44	2.79	38.17

表 2 水平阵阵输出主旁瓣比

入射角 (°)	CBF- G_{HLA} (dB)	MVDR- G_{HLA} (dB)	MBP- G_{HLA} (dB)
60	13.33	52.18	12.51
90	13.23	52.14	13.16
120	13.31	52.34	11.79
150	13.26	50.74	12.38

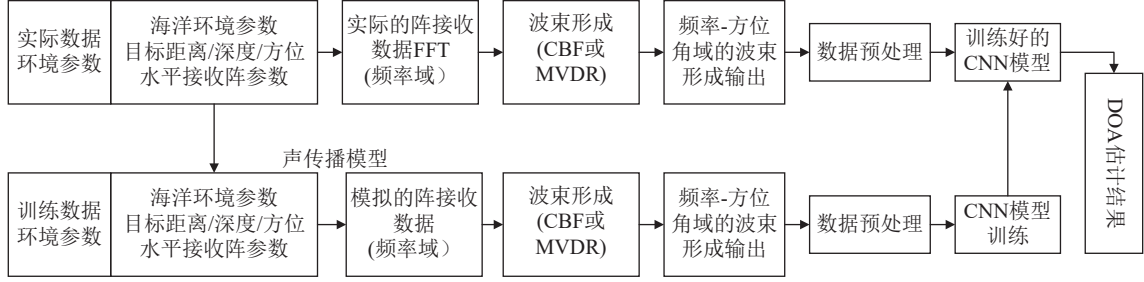


图 5 基于卷积神经网络的水声目标 DOA 估计流程

集预处理方法,将两种特征进行前端特征矩阵垂直连接融合,融合后的数据输入到网络中。多特征融合技术目标 DOA 估计具体流程如图 6 所示。

3.2 理论模型

针对水听器阵列获取的实测声场数据量不足以支撑网络模型训练的问题,本节建立了实际声场无源声呐阵列接收信号模型。

根据简正波理论^[9],水平均匀波导环境中线列阵各个阵元所接收到的声场数据是各阶简正波模态加权求和的结果:

$$P_H(r_i, z_r, f) = \frac{j}{\rho(z_s) \sqrt{8\pi} r_i} \sum_{m=1}^M Z_m(z_s) Z_m(z_r) \frac{e^{j(k_{rm} r - \pi/4)}}{\sqrt{k_{rm}}}, \quad (11)$$

$$\tilde{P}(r_i, z_r, f) = S(f) P_H(r_i, z_r, f), \quad (12)$$

其中, r_i 为接收距离, z_r 为接收深度, z_s 为声源深度, $\rho(z_s)$ 是声源深度上的介质密度, M 是有效简正波的阶数, $Z_m(z_s)$ 和 $Z_m(z_r)$ 是第 m 阶简正波相应的本征函数, k_{rm} 是本征函数对应的本征值, 即第 m 阶简正波水平波数。 $S(f)$ 为信号经傅里叶变换后在对应频段内的频谱, $P_H(r_i, z_r, f)$ 为各阵元感兴趣频段内利用声场计算软件计算出的若干离散频率的频域信道传输

函数, 频点与信号频点对应, 且与距离有关。

目标来波方位为 θ 时无源声呐阵列接收信号模型可以表示为

$$\mathbf{X}_S(\theta) = [\tilde{P}(r_1, z_r, f), \tilde{P}(r_2, z_r, f), \dots, \tilde{P}(r_N, z_r, f)]^T, \quad (13)$$

其中, $r_i, i = 1, 2, \dots, N$ 为目标到各阵元距离, z_r 为接收线列阵布放深度。

目标信号传播至接收线列阵处的声场为 $P_E(f, \theta, N_i)$, 其中 N_i 为阵元编号, $i = 1, 2, \dots, N$, $P_E(f, \theta, N_i)$ 与 DOA 信息有关, 与距离有关, 则有

$$P_E(f, \theta, N_i) = P_H(r_i, z_r, f), \quad (14)$$

$$r_i = \sqrt{r_0^2 + ((i-1)d)^2 - 2(i-1)r_0d \cos(180^\circ - \theta)}, \quad (15)$$

其中, r_0 为目标到接收线列阵参考阵元的距离, $r_i, i = 1, 2, \dots, N$ 为目标到各阵元的距离。式 (15) 的计算示意图如图 7 所示, N 个阵元以间隔 d 均匀分布。

通过扫描 DOA 可能存在的角度范围得到单频 CBF 波束谱为

$$\mathbf{P}_{\text{CBF}}(f, \theta) = \mathbf{w}_{\text{CBF}}^H(f, \theta) \mathbf{R} \mathbf{w}_{\text{CBF}}(f, \theta), \quad (16)$$

其中

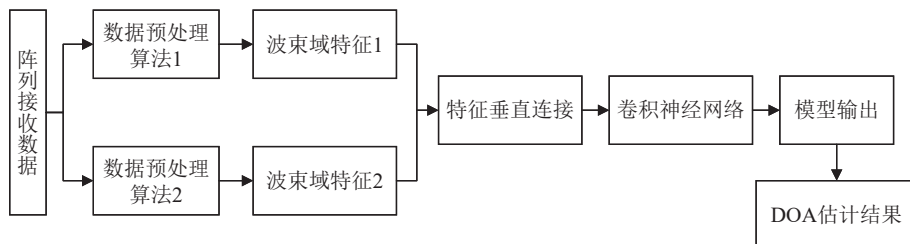


图 6 基于特征融合的 DOA 估计

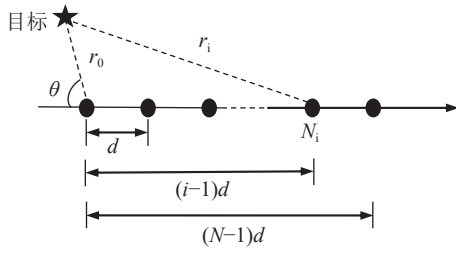


图 7 N 元水平线列阵示意图

$$\mathbf{R} = \mathbf{R}_x + \frac{\mathbf{R}_n}{10^{(R_{SN}/10)}} \mathbf{R}_x(1, 1),$$

$\mathbf{R}_x$ 为接收信号的互谱密度矩阵, $\mathbf{R}_n$ 为与 $\mathbf{R}_x$ 尺寸一样的单位噪声矩阵, R_{SN} 表示信噪比。

信噪比定义如下:

$$R_{SN} = 10 \log_{10} \frac{\sigma_s^2}{\sigma_n^2}, \quad (17)$$

其中, R_{SN} 定义为第一阵元接收信号与噪声方差的比值, $\sigma_s^2 = \mathbf{R}_x(1, 1)$ 为信号方差, σ_n^2 为零均值的高斯白噪声方差。

通过扫描 DOA 可能存在的角度范围得到单频对角加载 MVDR 波束谱为

$$\mathbf{P}_{MVDR}(f, \theta) = \mathbf{w}_{MVDR}^H(f, \theta) \mathbf{R} \mathbf{w}_{MVDR}(f, \theta), \quad (18)$$

其中, 不添加噪声时, $\mathbf{R} = \mathbf{R}_x + \rho \mathbf{I}$, ρ 为对角加载量。

根据上述建立的实际声场无源声呐阵列接收信号模型, 利用声场计算软件 KRAKENC 可以进行数据集的大量生成, 满足深度学习算法的训练数据量需求。

3.3 数据集预处理

本文实测数据处理方式与仿真数据归一化方式保持一致, 根据阵元域接收声场数据计算出互谱密度矩阵:

$$\mathbf{R}_x = \tilde{\mathbf{P}}(r, z, f) \tilde{\mathbf{P}}^H(r, z, f). \quad (19)$$

添加噪声后的互谱密度矩阵 $\mathbf{R}$ 归一化为

$$\mathbf{R}_{\text{norm}} = \frac{\mathbf{R}}{\sum_{i,j=1}^N \mathbf{R}(i=j)}, \quad (20)$$

统一将 CBF 和 MVDR 的波束谱写为 $\mathbf{P}_B(f, \theta)$:

$$\mathbf{P}_B(f, \theta) = \mathbf{w}^H(f, \theta) \mathbf{R}_{\text{norm}} \mathbf{w}(f, \theta). \quad (21)$$

波束谱取模值就可以表征 DOA 信息, 归一化并取模值后的波束谱 $\mathbf{P}_B(f, \theta)$ 的维度为 $N_f \times N_\theta$, 以此作为神经网络的输入, 其中 N_f 为频点个数, N_θ 为波束形成时扫描角度个数。每一个样本对应输入矩阵的形式为

$$\tilde{\mathbf{P}}_B(f, \theta_k) = \begin{bmatrix} |P_B(f_1, \theta_1)| & \cdots & |P_B(f_1, \theta_{N_\theta})| \\ \vdots & \ddots & \vdots \\ |P_B(f_{N_f}, \theta_1)| & \cdots & |P_B(f_{N_f}, \theta_{N_\theta})| \end{bmatrix}. \quad (22)$$

将 CBF 和 MVDR 两种波束域特征进行特征融合, 根据 3.1.2 节所述流程具体实现如下: 对于每一个样本, 分别进行 CBF、MVDR 波束形成处理, $\tilde{\mathbf{P}}_C(f, \theta)$ 为 CBF 处理结果, $\tilde{\mathbf{P}}_M(f, \theta)$ 为 MVDR 处理结果, 将两个矩阵进行垂直拼接, 得到每个输入样本特征为

$$\mathbf{P}_{\text{Input}} = \begin{bmatrix} \tilde{\mathbf{P}}_C(f, \theta) \\ \tilde{\mathbf{P}}_M(f, \theta) \end{bmatrix} = \begin{bmatrix} |P_C(f_1, \theta_1)| & \cdots & |P_C(f_1, \theta_{N_\theta})| \\ \vdots & \ddots & \vdots \\ |P_C(f_{N_f}, \theta_1)| & \cdots & |P_C(f_{N_f}, \theta_{N_\theta})| \\ |P_M(f_1, \theta_1)| & \cdots & |P_M(f_1, \theta_{N_\theta})| \\ \vdots & \ddots & \vdots \\ |P_M(f_{N_f}, \theta_1)| & \cdots & |P_M(f_{N_f}, \theta_{N_\theta})| \end{bmatrix}. \quad (23)$$

目标的 DOA 通常具有空间相关性, 而分类问题的输出中不同的类别所对应的空间 DOA 并无关联; 回归问题的输出则是用于预测连续的数值且网络的输出可以回归到多个值。结合分类输出中不同神经元位置代替目标的 DOA 值和回归输出中通过输入预测输出数值的特点, 本文将目标的 DOA 估计看作分类与向量回归的联合问题^[20]。先对目标的 DOA 可能存在的角度范围进行分类, 再对目标的 DOA 在每个类别上的概率进行回归, 根据 DOA 估计概率分布确定 DOA 数值。这种情况下采用高斯编码将 DOA 在角度域表示进行预处理, 得到每个样本 DOA 对应的正态分布。

设目标 DOA 为 θ_k , $k = 1, 2, \dots, K$, 对扫描角度范围以间隔 $\Delta\theta$ 等间隔划分为 N 份, 每一个输入矩阵 $\mathbf{P}_B(f, \theta)$, 对应一个标签向量为 $\mathbf{t}_k$:

$$\mathbf{t}_k = [t_{k1}, t_{k2}, \dots, t_{kN}], \quad (24)$$

$$t_{kj} \propto \exp\left[-(\theta_{kj} - \theta_{kr})^2 / (2\sigma^2)\right], \quad (25)$$

其中, k 为样本的标签序数, $\theta_{kj} = j\Delta\theta$ 为第 j 个划分角度, j 为各划分角度的序数, θ_{kr} 为第 k 个样本对应的真实 DOA, 对应到测试样本则是根据 GPS 记录计算出的真实 DOA 值, σ 为 DOA 正态分布的标准差, 表征真实 DOA 与预测出的 DOA 之间的偏差, 与波束宽度有关, σ 越大代表了对 DOA 估计误差容忍度越大, 但 σ 不是无限大, 超过一定界限 DOA 估计精度会下降, 在实际中为可调参数。

本文利用 CNN 网络进行 DOA 估计时, 扫描角度范围为 $1^\circ \sim 180^\circ$, 取值间隔为 1° , 输出层节点对应

的角度值为

$$\theta = [1^\circ, 2^\circ, \dots, 180^\circ]. \quad (26)$$

当模型训练完成时,将测试集样本输入至训练好的模型中进行测试。每一个样本分别与输出层输出的 1×180 的向量相对应,将输出向量经过处理后转换为有关各测试样本所对应的角度向量 $\mathbf{t}_k$,估计的 DOA 值 θ_{ek} 就是 $\mathbf{t}_k$ 最大值位置所代表角度值,即

$$\theta_{ek} = \arg \max(\mathbf{t}_k). \quad (27)$$

3.4 深度学习算法

3.4.1 网络模型设计及训练过程

图 8 为本文所采用 CNN 模型网络结构,其中 Batch 表示批量样本数目。所用模型包含 6 层卷积核尺寸为 3×3 的卷积层,包含输出层在内的 3 层全连接层和 2 层池化层。每层卷积核个数设置值依次为 32, 32, 32, 32, 128, 128, 分别在第 1 层卷积层和第 6 层卷积层后加入 1 层核为 2×2 的池化层,设置的 3 层全连接层神经元数量分别为 512, 512, 180。使用 Dropout 正则化方法,第 6 层卷积层之后将 Dropout 设置为 0.7。卷积层所采用的激活函数都设定为 ReLU 激活函数,损失函数选用最小均方误差函数。优化算法选用 Adam,初始学习率设置为 0.0001, CNN 的训练过程包含 500 个 epoch,批量样本数目为 32。

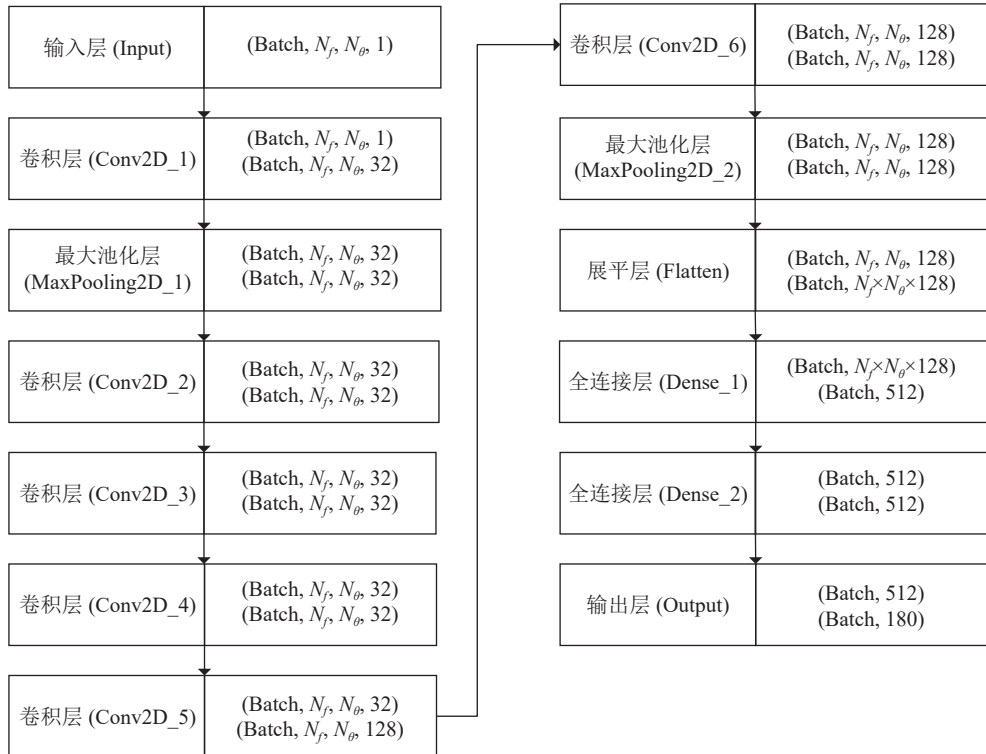


图 8 CNN 网络模型结构

CNN 训练过程分为前向传输预测阶段和后向传输更新权重阶段。首先,预处理后的数据集经各层网络变换传输到输出层进行预测输出,网络训练结束后的预测阶段也是如此;然后,根据样本数据的对应标签与模型的预测输出进行最小均方误差计算;随后计算每个权重的梯度,利用 Adam 优化器更新权重。

3.4.2 网络模型性能评价

为了定量评估不同 DOA 估计算法下的定位性能,定义可靠测向概率 (PCDE) 为 DOA 估计结果的绝对误差小于 5° 的概率,可以反映实际各种算法预测出的 DOA 结果的可靠性;均方根误差 (RMSE) 用于衡量预测出的 DOA 与真实 DOA 之间的偏差,即

$$\text{PCDE-}5^\circ = \frac{1}{N_{\text{test}}} \sum_{k=1}^{N_{\text{test}}} \eta(k), \quad (28)$$

$$\eta(k) = \begin{cases} 1, & |\theta_{ek} - \theta_{rk}| \leq 5^\circ, \\ 0, & \text{其他}, \end{cases} \quad (29)$$

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{test}}} \sum_{k=1}^{N_{\text{test}}} |\theta_{ek} - \theta_{rk}|^2}, \quad (30)$$

其中, N_{test} 是待估计样本总数, $\eta(k) = 1$ 时表示第 k 个样本的 DOA 被有效估计出。 θ_{rk} 是第 k 个样本根据 GPS 计算出的真实 DOA, θ_{ek} 是第 k 个样本根据 DOA 估计算法预测出的 DOA。

4 基于 CNN 的 CBF 和 MVDR 特征域仿真分析

4.1 仿真环境

采用 Pekeris 型浅海波导, 设置声场参数: 海底的声速、密度和衰减分别为 1600 m/s, 1.5 g/cm³, 0.2 dB/λ, 声速剖面呈等声速梯度并设置为 1500 m/s, 32 阵元接收阵, 目标的 DOA 取值范围为 30°~150°。来波信号半波长条件约束为阵元间隔 $d \leq c/2f$, 其中 c 为接收阵所在深度上的声速, f 为信号频率, 信号频率取值范围为 100~200 Hz, 则 $d \leq 3.75$ m。训练集和测试集参数设置分别如表 3 和表 4 所示。训练集样本数据量为 $121 \times 5 \times 8 \times 6 = 29040$, 输入 CNN 网络的特征维度为 11×180 , 不同信噪比对应下每组测试集样本数据量为 $121 \times 1 \times 8 \times 6 = 5808$, 输入 CNN 网络的特征维度为 11×180 。

4.2 仿真结果

设置声源深度 20 m, 声源到参考阵元距离为 4 km, 水平线列阵接收深度为 50 m, 水深为 60 m。当 $f = 150$ Hz, 阵元间隔 $d = 3$ m, 来波方向取值范围为 30°~150°, 取值间隔为 1°时, CBF 和 MVDR 输出特征见图 9; 改变阵元间隔为 $d = 15$ m, 结果见图 10。当 f 取值范围为 100~200 Hz, 取值间隔为 1 Hz, 来波方向为 30°时, 阵元间隔 $d = 3$ m, 频率-波束谱结构见图 11; 改变阵元间隔为 $d = 15$ m, 结果见图 12。最高能量点在图中均用黑色圆圈标出。

如图 9 和图 11 所示, 在 $d = 3$ m 时, CBF 和 MVDR 估计结果最高能量输出点指示 DOA 估计结果, 主瓣强度最高, 几乎无旁瓣的干扰, MVDR 拥有窄的主瓣宽度。当 $d = 15$ m 时, 如图 10 和图 12 所示, 二者波束谱不仅包含 DOA 方位输出信息, 还包含多个强度很高的栅瓣信息, 这将造成很大的干扰, 波束图指示的最高能量点发生改变, 不是完全处于来波方向误差范围内, CBF 和 MVDR 不能正确估计 DOA。

波长的变化会引起波束谱能量分布发生变化。分别在满足半波长条件 ($d = 3$ m) 和不满足半波长条件 ($d = 15$ m) 下进行训练和测试, 训练集中均未加入噪声, MVDR 训练集采用对角加载方式生成 (对角加载量为 0.01), 设置对应波长条件下 CBF 和 MVDR 测试集信噪比取值范围为 -20~20 dB, 取值间隔为 5 dB, 在不同信噪比对应测试集下, 测试不同阵元间隔下模型的性能, 比较 RMSE 和 PCDE-5°。

表 3 训练集声学参数设置范围

仿真参数	训练集参数范围	参数个数
声源频率 (Hz)	100:10:200	11
DOA (°)	30:1:150	121
声源拖曳深度 (m)	10:5:30	5
声源到参考阵元距离 (km)	2.5:0.5:6	8
水深 (m)	50:10:100	6
水平阵接收深度 (m)	40:10:90 (与水深一一对应)	1

表 4 测试集声学参数设置范围

仿真参数	测试集参数范围	参数个数
声源频率 (Hz)	100:10:200	11
DOA (°)	30:1:150	121
声源拖曳深度 (m)	24	1
声源到参考阵元距离 (km)	2.5:0.5:6	8
水深 (m)	50:10:100	6
水平阵接收深度 (m)	40:10:90 (与水深一一对应)	1
信噪比 (dB)	-20:5:20	9

注: 表 3 和表 4 中参数范围数据格式为: 最小取值: 取值间隔: 最大取值。例如, 声源频率参数范围为 100:10:200, 表示声源频率最小取 100 Hz, 最大取 200 Hz, 取值间隔为 10 Hz; 40:10:90 (与水深一一对应) 表示水平阵接收深度与水深差值 10 m, 例如, 当水深为 50 m 时, 水平阵接收深度为 40 m。

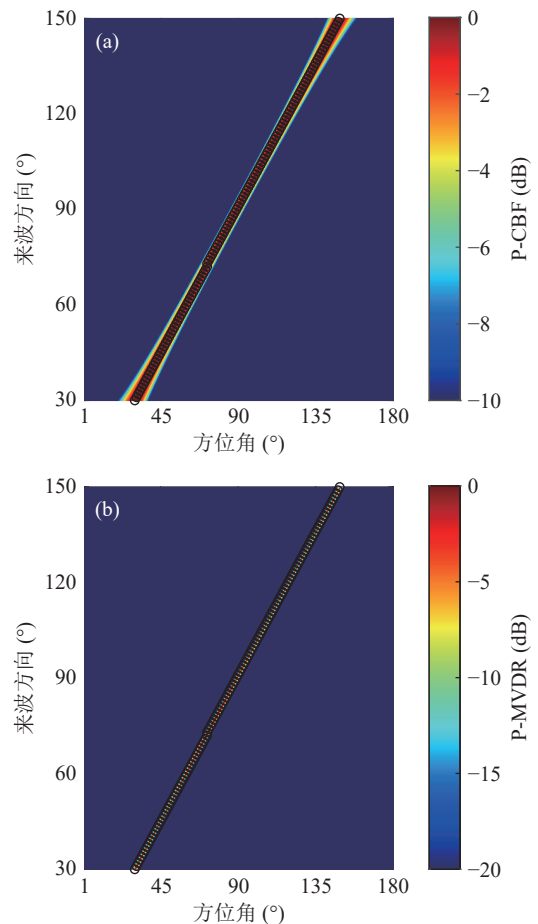


图 9 $d = 3$ m 时来波方向取值范围为 30°~150°, 取值间隔为 1°时的方位扫描图 (a) CBF; (b) MVDR

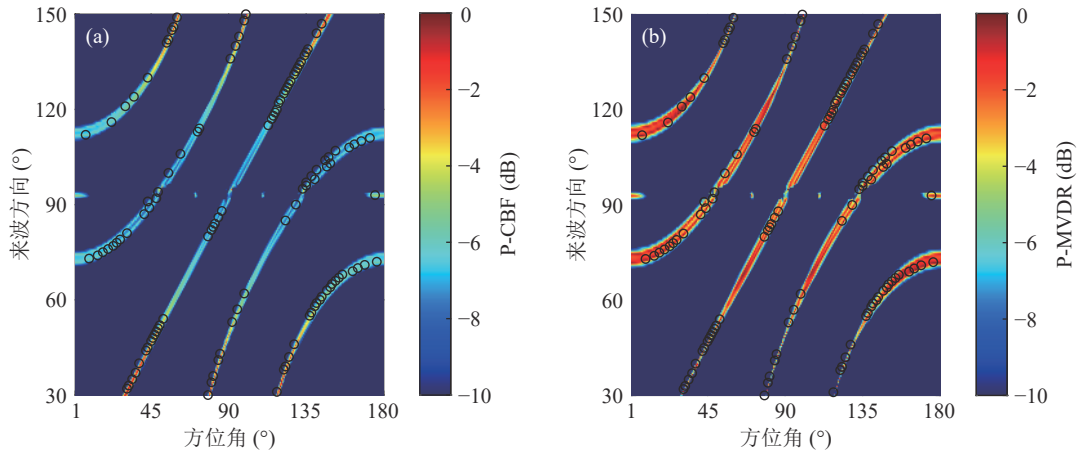


图 10 $d = 15$ m 时来波方向取值范围为 $30^\circ \sim 150^\circ$, 取值间隔为 1° 时的方位扫描图 (a) CBF; (b) MVDR

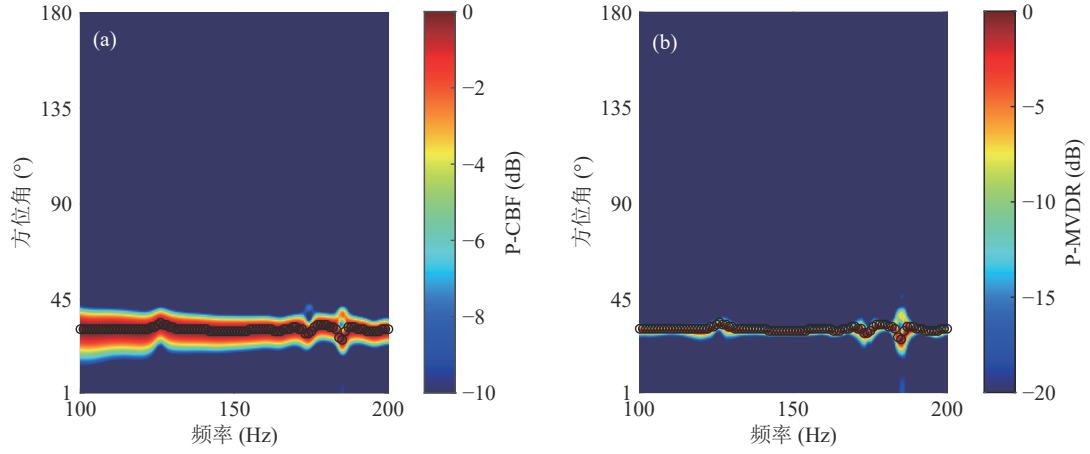


图 11 $d = 3$ m 时来波方向 30° 的频率-波束谱输出 (a) CBF; (b) MVDR

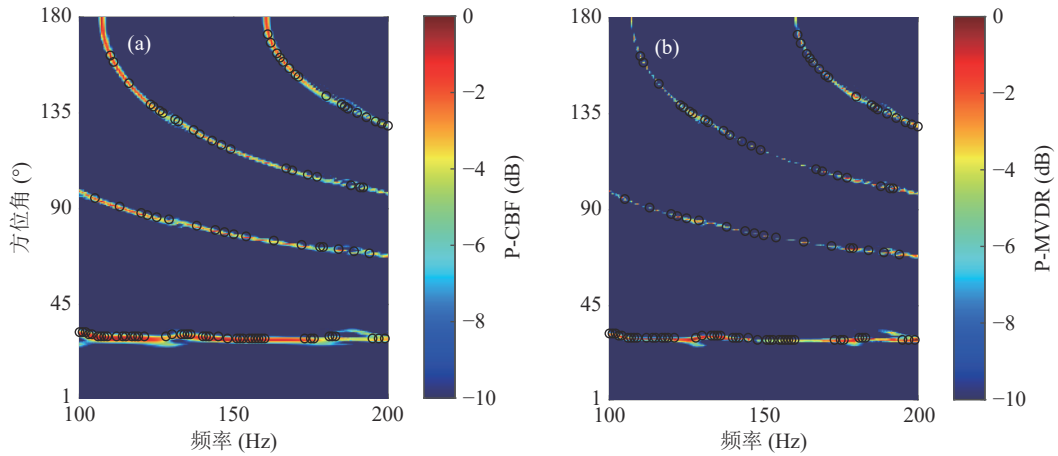


图 12 $d = 15$ m 时来波方向 30° 的频率-波束谱输出 (a) CBF; (b) MVDR

网络模型结构和训练参数与 3.4.1 节保持一致, 迭代的最大次数为 500, 信噪比取值范围为 $-20 \sim 20$ dB 时, 取值间隔为 5 dB, 定义以 CBF 和 MVDR 波束谱作为特征输入时模型为 CNN-CBF 和 CNN-MVDR, 则在不同阵元间隔条件下 CNN-CBF、CNN-MVDR 模型预测结果的 RMSE 和 PCDE-5° 随信噪比

变化如图 13 所示。

由图 13 可见, 当 $d = 3$ m 时, CNN-CBF 和 CNN-MVDR 性能表现差别不大, CNN-MVDR 估计结果在信噪比小于 -5 dB 时略优于 CNN-CBF。在信噪比大于等于 -10 dB 时, 两种估计方法的 PCDE-5° 均逼近 100% 并趋于稳定, RMSE 下降到 2° 以下趋于稳定,

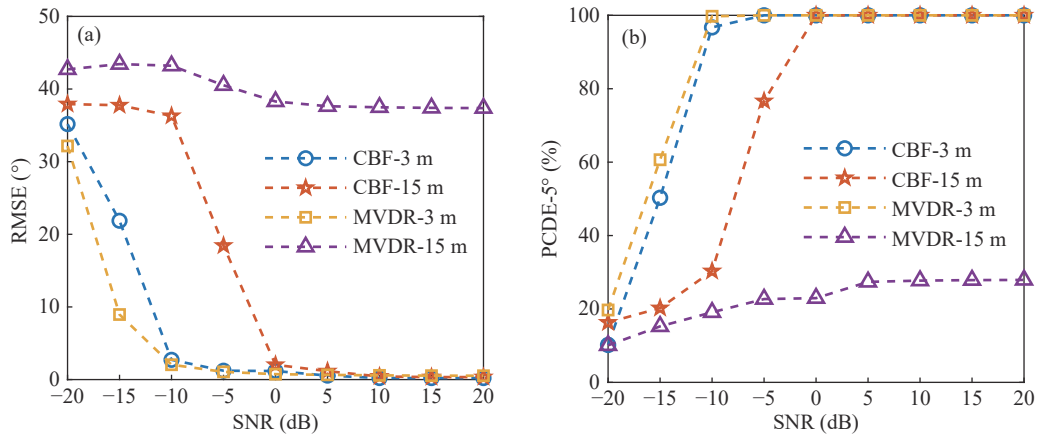


图 13 CNN 模型的 DOA 测试结果性能评价 (a) RMSE 随信噪比变化; (b) PCDE-5°随信噪比变化

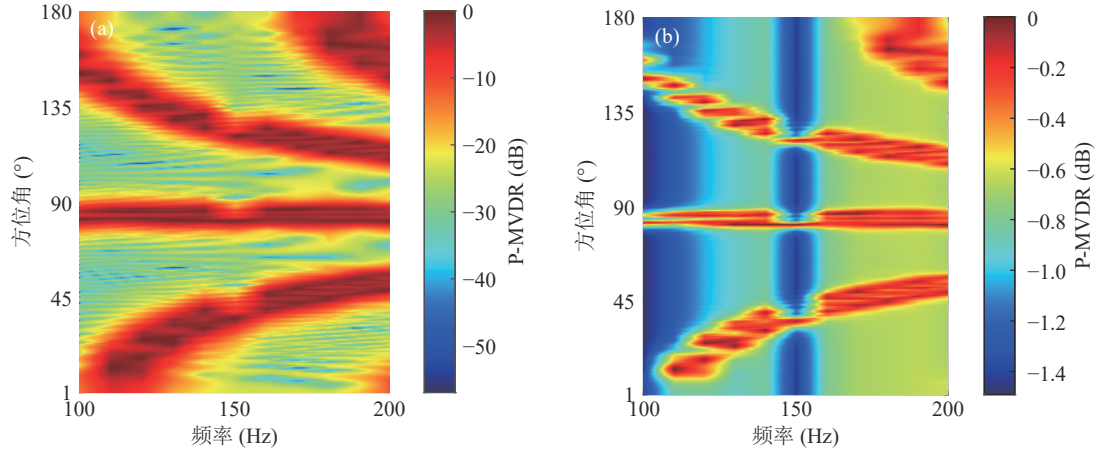


图 14 输入 CNN 网络的训练样本与测试样本特征比较 (来波方向为 90°) (a) 训练数据的频率-波束谱输出特征; (b) 测试数据的频率-波束谱输出特征

都能准确估计出 DOA。两种方法对于信噪比都具有很强的宽容性。

当 $d = 15$ m 时, 在设定的信噪比范围内, CNN-MVDR 模型预测结果的 PCDE-5°均在 30% 以下, RMSE 最小也在 35°以上, 可以视为估计方法失效。当信噪比大于等于 0 dB 时, CNN-CBF 模型预测结果的 PCDE-5°达到 100% 趋于稳定, RMSE 也下降到 2°以下并趋于平稳, 说明 CNN-CBF 方法对于阵元间隔具有一定的宽容性。当阵元间隔不满足半波长条件时, CNN 仍能够从 CBF 特征量上学习到关于真实 DOA 的信息, 栅瓣的存在一定程度上降低了 CNN-CBF 对于信噪比的鲁棒性。

分析 $d = 15$ m 时 CNN-MVDR 失效的原因, 以及训练集和带噪声的测试集的特征分布。仿真时, 声场环境参数一致, 来波方向为 90°, 训练样本及测试样本的频率-波束谱输出特征结构如图 14 所示。在训练集无噪声情况下进行了对角加载 (加载量为 0.01); 测试集在进行仿真时, 在互谱密度矩阵上加入噪声 (信噪比为 5 dB)。

由图 14 可见, 两子图特征分布存在一定差异。MVDR 加权向量与数据的互谱密度矩阵相关, 带噪声条件下互谱密度矩阵的误差会增加, 训练样本的互谱密度矩阵进行了对角加载, 对角加载量的选取本身没有严格的标准, 存在特征差异分布不一致的问题, 从而导致测试结果的误差。

针对上述问题, 可以根据具体使用情景使用带噪声训练集训练模型。在 MVDR 训练集样本数据中加入随机的高斯白噪声, 随机信噪比的取值范围为 0~20 dB, 测试集信噪比取值范围为 -20~20 dB, 取值间隔为 5 dB。CNN-MVDR 模型测试结果的 RMSE 和 PCDE-5°随信噪比变化如图 15 所示。

由图 15 可见, 训练集中加入一定的信噪比后, 网络模型在测试集上的预测结果显著提高且具有规律性。信噪比大于等于 -20 dB 的估计结果均可靠。这也验证了上述在对角加载下的 CNN 算法失效的原因, 特征分布存在很大的差异性会带来测试误差。阵元间隔失配下 CBF 和 MVDR 两种特征量均可以有效地被用于训练和测试, 并且在训练集中加

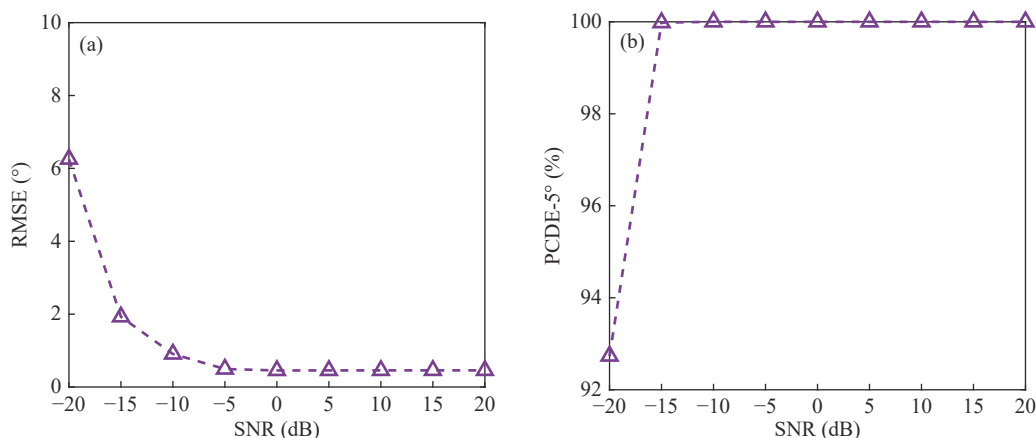


图 15 CNN-MVDR 的 DOA 测试结果性能评价 (a) RMSE 随信噪比变化; (b) PCDE-5°随信噪比变化

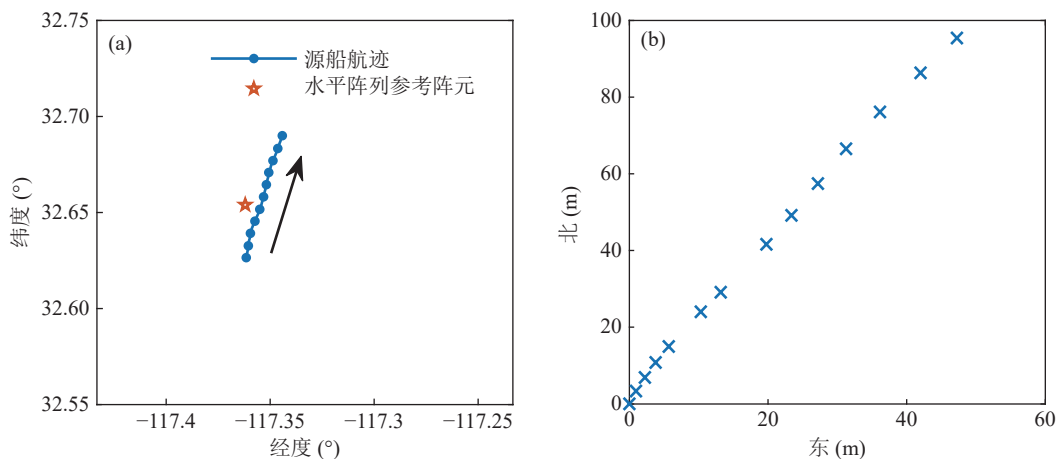


图 16 源船及接收阵信息 (a) 源船经纬度变化 (箭头表示行进方向); (b) 子阵 1 各阵元位置

入一定噪声, 网络模型的性能有明显的提升。

5 实验数据处理

5.1 实验介绍

S5 测线是指 SWellEx-96 实验期间沿等深线拖曳的声源的实验轨迹, S5 测线实验不存在干扰源, 源船速度为 5 kn, 深源拖曳深度为 54 m。

水平阵 HLAN 接收数据共持续 50 min。源船的经纬度坐标变化如图 16 (a) 所示, 接收阵列位于 213 m 的海底, 孔径大小为 240 m, 阵元个数为 32 (有效个数 27, 损坏 5 个), 阵型从形状来看呈弓形, 以前 14 个阵元作为子阵 1, 子阵 1 各阵元位置如图 16 (b)。声源到参考阵元 (水平阵 HLAN 第一个阵元为 (0, 0) 位置, GPS 已知) 的距离及 DOA 随时间变化如图 17 所示, GPS 记录为每 1 分钟记录一次, 采用插值方法计算整体真实距离及 DOA 数值。

S5 测线的环境模型 (距离无关的浅海波导) 如图 18 (a) 所示, 水深为 213 m, 三层海底模型, 声源位于 54 m。沉积层厚度为 23.5 m, 沉积层顶部和底部

的声速分别为 1572.368 m/s 和 1593.016 m/s, 沉积层的密度和衰减分别为 1.76 g/cm^3 和 $0.2 \text{ dB/(km}\cdot\text{Hz)}$ 。泥岩层顶部和底部的声速分别为 1881 m/s 和 3245 m/s, 泥岩层的密度和衰减分别为 2.06 g/cm^3 和 $0.06 \text{ dB/(km}\cdot\text{Hz)}$ 。基底的声速、密度和衰减分别为 5200 m/s, 2.06 g/cm^3 , $0.06 \text{ dB/(km}\cdot\text{Hz)}$ 。声速剖面如图 18 (b) 所示, 为 CTD 记录的真实数据。

水听器采样率为 3276.8 Hz, 声源信号为单频脉冲信号, 高电平信号对应频率为 49 Hz, 64 Hz, 79 Hz。首先, 将水平阵数据转化为整数倍采样, 采样率扩大 10 倍为 32768 Hz, 以 1.25 s 脉冲宽度进行采样, 共得到 2400 个样本; 然后, 对 2400 个样本进行单快拍 CBF 处理。源船方位历程图如图 19 所示, 其中圈出的三部分是源船拖曳时的停止发射深源信号段, 去除这三段停止发射信号段样本, 共得到 2090 个有效样本。

5.2 训练数据集仿真

本节训练数据仿真参照图 18 (a) 环境模型。采用 3.2 节无源声呐阵列接收信号模型, 利用声场仿真

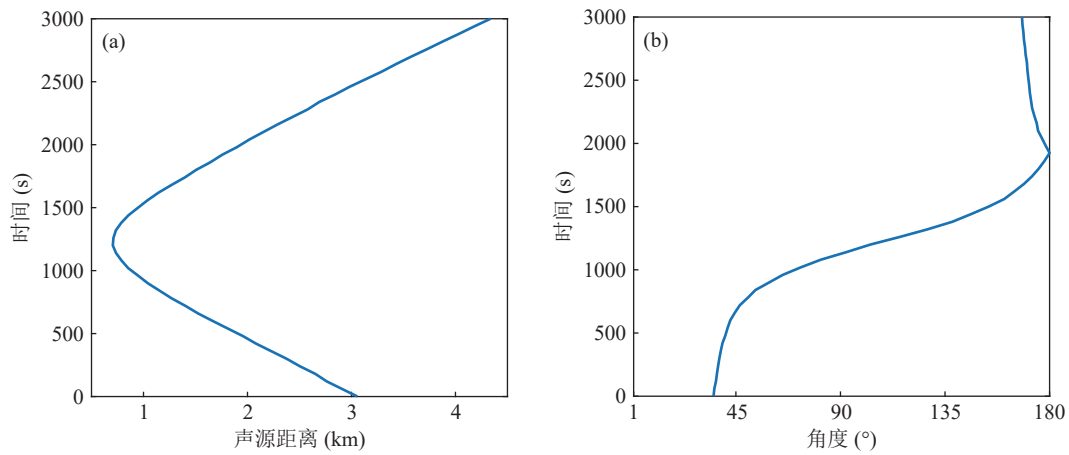


图 17 声源到参考阵元的距离及 DOA 随时间变化情况 (a) 距离随时间变化; (b) DOA 随时间变化

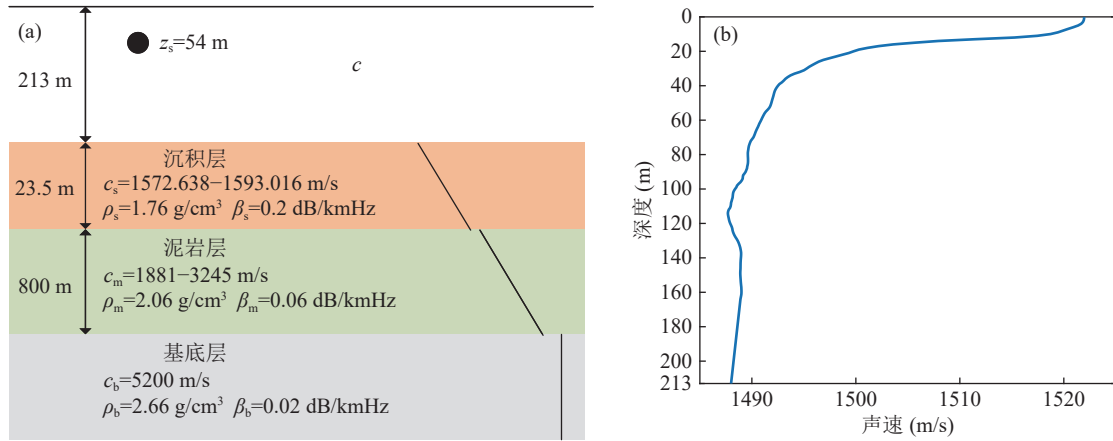


图 18 环境模型及水文参数 (a) 环境模型; (b) 声速剖面

软件 KRAKENC 计算声压。 $S(f)$ 线谱幅值设为 1。参考实验相关参数的训练集相关声学参数范围设置如表 5 所示, 声速剖面参考实测声速剖面如图 18 (b) 所示。

表 5 针对实验数据相关参数的训练集声学参数设置范围

参数	训练集参数范围	参数个数
声源频率 (Hz)	49, 64, 79	3
DOA (°)	1:1:180	180
声源拖曳深度 (m)	50:1:60	11
声源到参考阵元距离 (km)	2.5:0.5:6	8
水深 (m)	180:20:220	3
水平阵接收深度 (m)	170:20:210 (与水深一一对应)	1

训练集个数为 $180 \times 11 \times 8 \times 3 = 47520$, 输入 CNN 网络的特征维度为 3×180 。对仿真的数据集, 利用 3.3 节中选择的预处理方式进行数据及标签的预处理。训练集样本中未添加噪声, 将仿真训练集数据按照 7:3 进行划分训练集和验证集, 网络模型采用图 8 所示结构。在训练和测试时, 分别对单频或综合多个频率对应数据进行训练和测试。

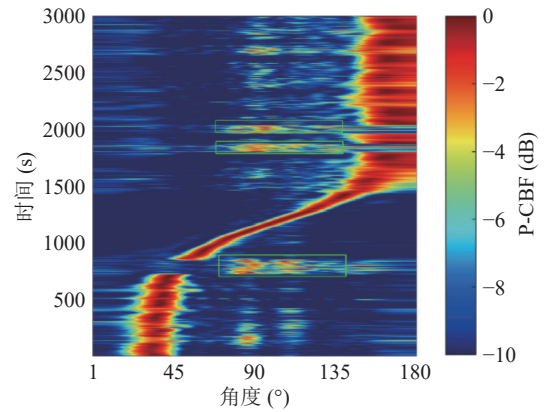


图 19 源船方位历程图

5.3 实验结果及分析

将 CNN-CBF 和 CNN-MVDR 方法应用到 5.1 节实际测线实验中, 进行目标的 DOA 估计, 测试样本数为 2090。对测试样本数据进行单快拍 CBF 处理和 MVDR 对角加载处理, 得到的频率-波束输出特征分别作为 CNN 网络输入特征进行测试。训练集样本标签标定时设定的标准差 $\sigma = 20$, 在训练过程中针对不同频点, 网络模型的相关参数设置都是相

同的,在训练过程中未出现误差反增情况,训练模型有效,未出现过拟合现象,验证集误差最小时保存最佳模型。

利用 3.4.2 节定义的 RMSE 和 PCDE-5° 进行各算法性能评价。将 CNN、MBP、CBF 和 10 个快拍下对角加载 MVDR 的 DOA 预测结果的相关性能评价标准列于表 6,其中包括单频 49 Hz、64 Hz、79 Hz 训练的测试结果的性能评价,以及三个频点 {49 Hz、64 Hz、79 Hz} 综合训练的测试结果的性能评价,柱状图对比如图 20 所示。

分析表 6 统计结果。在综合频点 {49 Hz、64 Hz、79 Hz} 条件下,五种方法的 DOA 估计性能均比单频条件下有很大提升。传统波束形成方法及 MBP 方法的 PCDE-5° 不超过 65%,CNN 方法相比其提高了 10% 以上;CNN 方法的 RMSE 也有所降低。

CBF 作为常规的处理方法,在综合频点下 DOA 估计结果的 RMSE 为 8.76°,PCDE-5° 为 64.21%;给出的对角加载 MVDR 在 10 个快拍条件下的性能表现严重下降, RMSE 为 28.53°, 偏离真实值的离散点增多, PCDE-5° 下降为 52.92%。MVDR 在多快拍情况下,估计性能并没有得到明显的提升,这可能是因为源船在拖曳过程中的 DOA 随时间变化太大,各个快拍相关性表现均不同,互谱密度矩阵估计误差增大,多快拍下的估计效果并不理想。

MBP 和 CNN 方法在数据处理阶段都应用了波束形成技术,详细对比 MBP 与 CNN 的性能评价结果。MBP 在单频 49 Hz、64 Hz、79 Hz 和所有频点 {49 Hz、64 Hz、79 Hz} 四种条件下, RMSE 测向误差分别为 22.78°, 37.53°, 25.72°, 11.78°, PCDE-5° 分别为 40.57%, 39.86%, 45.84%, 55.74%。CNN-CBF 相比 MBP 在上述四种条件下的 RMSE 测向误差分别降低了 5.92°, 23.84°, 4.47°, 4.52°, CNN-CBF 相比 MBP 在上述四种条件下的 PCDE-5° 分别提高了 3.02%, 11.58%, 12.72%, 19.91%。CNN-MVDR 估计方法的 DOA 估计性能与 CNN-CBF 相似,其 PCDE-5° 也都

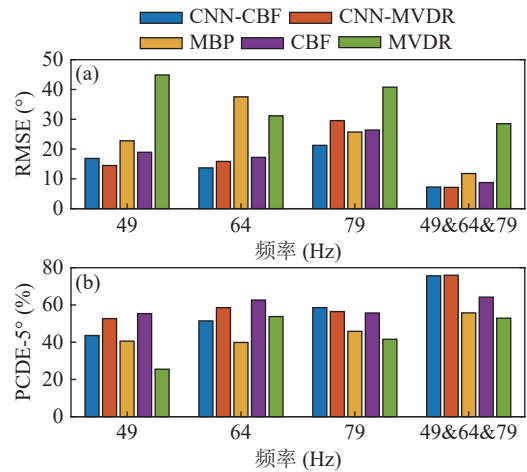


图 20 单频和多频条件下 DOA 测试结果性能评价 (a) 不同频率下的 RMSE; (b) 不同频率下的 PCDE-5°

高于 MBP 且 RMSE 明显降低。

综合频点 {49 Hz、64 Hz、79 Hz} 下 CBF、10 个快拍下对角加载 MVDR、MBP、CNN-CBF 和 CNN-MVDR 的估计结果与真实角度的散点图如图 21 所示。CNN-CBF 和 CNN-MVDR 总体拟合效果优于 CBF、MVDR 和 MBP 三种方法。五种方法估计结果特点总结来看:在中间时间段的 DOA 估计性能优于开始一段时间和最后一时间段,中间时间段的 DOA 估计结果与真实 DOA 拟合程度更高。从 CNN-CBF 和 CNN-MVDR 两种方法在端射方向估计结果来看,估计值更加聚集于真实值附近,端射方向的估计性能获得很大提升。

采用高斯输出层时,根据式 (25) 可知,针对不同实际情况,角度划分间隔 θ_{kj} 和标准差 σ 要自行设置。 θ_{kj} 决定了网络输出神经元个数, σ 表示预测 DOA 与真实 DOA 间允许的最大误差。为了研究不同标准差设定值对 DOA 估计性能的影响,通过设置不同的 σ 值,其他训练参数保持一致,在相同的训练数据集和训练迭代次数下比较测试结果的 RMSE 和 PCDE-5° 的变化,如图 22 所示。RMSE 随着 σ 的增

表 6 单频和多频条件下 DOA 测试结果性能评价

性能评价	频率 (Hz)	CNN-CBF	CNN-MVDR	MBP	CBF	MVDR
RMSE (°)	49	16.86	14.48	22.78	18.93	44.87
	64	13.69	15.86	37.53	17.21	31.17
	79	21.25	29.55	25.72	26.39	40.78
	{49, 64, 79}	7.26	7.14	11.78	8.76	28.53
PCDE-5° (%)	49	43.59	52.68	40.57	55.36	25.50
	64	51.44	58.56	39.86	62.63	53.78
	79	58.56	56.46	45.84	55.69	41.58
	{49, 64, 79}	75.65	75.96	55.74	64.21	52.92

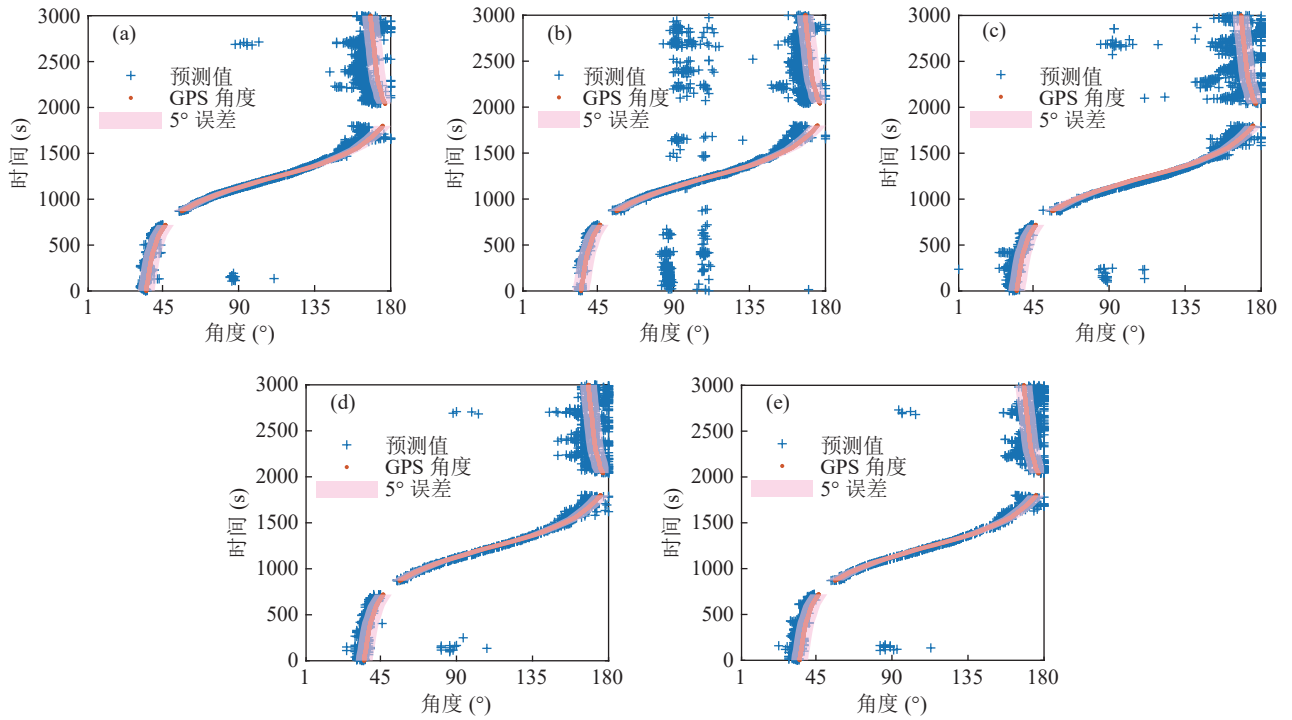


图 21 不同估计方法散点图对比 (a) CBF 的 DOA 估计结果; (b) 10 个快拍 MVDR 的 DOA 估计结果; (c) MBP 的 DOA 估计结果; (d) CNN-CBF 的 DOA 估计结果; (e) CNN-MVDR 的 DOA 估计结果

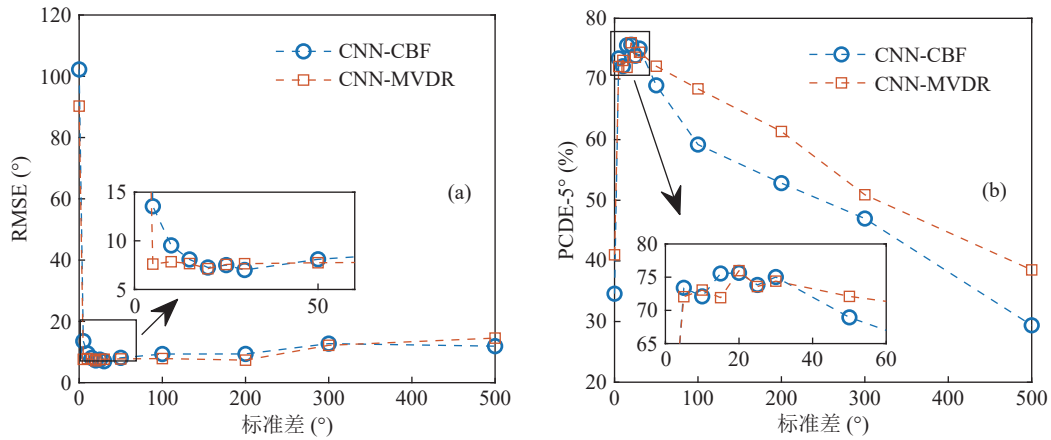


图 22 不同标准差下 DOA 估计结果性能评价 (a) RMSE 随信噪比变化; (b) PCDE-5°随信噪比变化

大不断减小至最佳预测位置后趋于稳定后开始回涨, RMSE 更能反映估计中出现的离散点的偏离程度, 说明偏离真实值的点数增多。随着 σ 的不断增大, PCDE-5° 不断增大后趋于稳定。当 σ 值超过一定界限时, 从原理上看预测对于真实 DOA 附近角度的宽容性变大, 可以容忍更大误差范围的 DOA 进入预测区间, 导致预测结果偏差增大, PCDE-5° 开始下降。

在基于单特征的 CNN 算法基础上, 进行特征融合时, 输入 CNN 特征维度为 6×180 , 训练数据集参数设置与 5.2 节保持一致, 训练集个数为 $180 \times 11 \times 8 \times 3 = 47520$ 。利用 3.3 节数据预处理方式, 训练集和验证集按照 7:3 的比例进行划分, 网络模型依旧

采用图 8 结构, 相关超参数设置不变, 将这种特征融合方法命名为 CNN-II。

对 5.1 节中 S5 测线实验样本数据进行测试, 基于 CNN-II 的目标 DOA 估计结果与真实角度的散点图如图 23 所示。CNN-II 预测结果的 PCDE-5° 提升到了 79.78%, RMSE 降低到 6.95° , 相比较单特征估计方法 CNN-CBF 和 CNN-MVDR, CNN-II 的 PCDE-5° 分别提高了 4.13% 和 3.82%, RMSE 分别降低了 0.31° 和 0.19° 。证明特征融合方法对于 DOA 估计性能提升是有效的。

通过对比, 融合两种特征的深度学习算法在本文 DOA 估计任务中的 PCDE-5° 最高且 RMSE 最低,

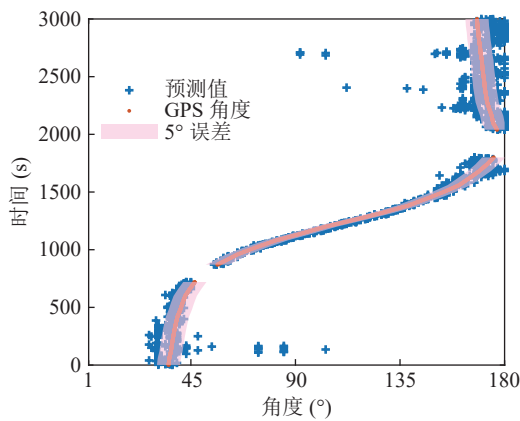


图 23 CNN-II 的 DOA 估计结果

融合后效果优于单独使用 CBF 和 MVDR 特征的训练测试结果,且垂直连接的操作简单易实现。分析其原因可能是:CBF 特征反映了扫描 DOA 范围内的波束谱的整体差异性,旁瓣信息更加丰富,主瓣更宽,特征结构上信息量更多;而 MVDR 特征在目标的真实 DOA 上信息更加集中,单一的特征对于相近的不同 DOA 的表征能力存在局限性。将 CBF 和 MVDR 两种特征按照通道数方向垂直连接起来,可以综合利用不同特征表征出的 DOA 对应波束谱的敏感度信息,CNN 模型可以同时学习到两个特征包含的目标方向的信息,从而获得更完整、更准确的目标方向信息,进一步影响 DOA 估计的性能。因此,融合两种特征的 CNN 估计方法可以获得更好的性能。

5.4 算法处理速度评估

本文所用数据处理操作系统为 Windows 11。GPU 配置为 NVIDIA GeForce RTX 3080, CPU 配置为酷睿 i7-12700KF。采用统一计算架构为 CUDA 11.7 和 TensorFlow 2.7 深度学习库,集成开发环境为 Anaconda 22.9.0,使用 Spyder 集成开发环境,编程语言采用 Python (Ver. 3.8)。使用 MATLAB 软件对线阵列接收信号进行波束形成处理。

CNN、MBP、CBF 和 10 个快拍下对角加载 MVDR 的处理速度如表 7 所示。处理速度评估结果是指每种算法对总的测试样本处理的平均时间。CNN 方法的处理时间包含了数据样本的特征提取与测试的总时间。CBF、MVDR 和 MBP 的 GPU 版的处理速度为空。

实验数据处理时 CNN 模型的训练是利用 GPU 完成的。使用 GPU 的训练时间:CNN-CBF 模型训练时间为 3742.60 s, CNN-MVDR 模型训练时间为 3846.65 s, CNN-II 模型训练时间为 5907.33 s,特征融合方法增加了输入特征的维度,训练模型速度会

表 7 DOA 估计方法处理时间

DOA 估计方法	CPU 处理速度 (ms)	GPU 处理速度 (ms)
CNN-CBF	2.20	1.42
CNN-MVDR	12.04	11.26
CNN-II	14.10	12.58
CBF	1.30	—
MVDR	11.11	—
MBP	3450	—

降低。

由表 7 可见,传统的波束形成方法速度最快,但其性能表现并不是最优的;匹配场方法处理速度最慢,在实际应用中优势不大;CNN 方法可以在已知预训练模型时,快速高效地应用于真实环境中未知数据的实时处理,在今后还可以考虑改用其他语言将数据矩阵处理模块与 CNN 测试模型进行联合部署,提高处理速度。

6 结论

水平线列阵 DOA 估计的数值仿真分析表明,仅依赖于数据的波束形成 DOA 估计方法在端射方向上会产生大的角度偏差,而 MBP 技术能够较好地解决理想条件下端射估计准确度不高的问题。然而 MBP 在实际应用中则需要更精确的声场模型参数,由于真实海洋环境中的不确定性,使得真实环境相关知识的获取困难,这可能会导致拷贝场模型失配,MBP 的性能也会大幅度降低,在实测海试数据下性能不如常规波束形成。

本文提出了一种利用卷积神经网络模型进行 DOA 估计的方法,该方法将水平线列阵接收数据的波束域特征输入到网络模型中。通过数值仿真验证,以单一波束域特征数据作为 CNN 模型的输入时,模型对信噪比具有一定的鲁棒性,对阵元间隔具有一定的容忍性。在阵元间隔不能满足半波长情况下,波束谱存在栅瓣的影响,CNN 仍能学习到输入数据的相关 DOA 信息来估计目标 DOA。海试数据处理结果表明,采用 CNN 的估计方法性能比传统的 MBP 方法以及波束形成方法有较大提升,端射方向估计也有较大改善。简单实现的多特征融合的 DOA 估计方法在海试数据下 5°误差内可靠测向概率最高且均方根误差最低,比单一特征训练测试表现好,这也说明了对于目标的多个特征可考虑合适的融合处理以取得更高的估计精度。

参 考 文 献

1 王超, 笄良龙, 韩梅, 等. 单矢量水听器稀疏近似最小方差方位

- 估计算法. *声学学报*, 2021; **46**(6): 1050–1058
- 2 谭鹏, 胡博, 张友文, 等. 交叉验证多路径匹配追踪水声矢量阵稀疏方位估计. *声学学报*, 2022; **47**(5): 557–567
- 3 Krim H, Viberg M. Two decades of array signal processing research: The parametric approach. *IEEE. Signal. Proc. Mag.*, 1996; **13**(4): 67–94
- 4 Cox H, Zeskind R, Owen M. Robust adaptive beamforming. *IEEE Trans. Acoust. Speech Signal Process.*, 1987; **35**(10): 1365–1376
- 5 Yang T C, Yates T. Matched-beam processing: Application to a horizontal line array in shallow water. *J. Acoust. Soc. Am.*, 1998; **104**(3): 1316–1330
- 6 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015; **521**(7553): 436–444
- 7 Schmidhuber J. Deep learning in neural networks: An overview. *Neural Networks*, 2015; **61**: 85–117
- 8 Deng L, Yu D. Deep learning: Methods and applications. *Found. Trends Signal Process.*, 2014; **7**(3-4): 197–387
- 9 Niu H, Reeves E, Gerstoft P. Source localization in an ocean waveguide using supervised machine learning. *J. Acoust. Soc. Am.*, 2017; **142**(3): 1176–1188
- 10 Niu H, Ozanich E, Gerstoft P. Ship localization in Santa Barbara Channel using machine learning classifiers. *J. Acoust. Soc. Am.*, 2017; **142**(5): EL455–EL460
- 11 Steinberg B Z, Beran M J, Chin S H, *et al.* A neural network approach to source localization. *J. Acoust. Soc. Am.*, 1991; **90**(4): 2081–2090
- 12 Elbir A M. Deep MUSIC: Multiple signal classification via deep learning. *IEEE Sens. Lett.*, 2020; **4**(4): 1–4
- 13 Liu Y, Chen H, Wang B. DOA estimation based on CNN for underwater acoustic array. *Appl. Acoust.*, 2021; **172**: 107594
- 14 曹怀刚. 单矢量水听器波达方向估计的深度学习模型构建与应用. 博士学位论文, 北京: 中国科学院大学, 2021: 44–53
- 15 Lawrence S, Giles C L, Tsoi A C, *et al.* Face recognition: A convolutional neural-network approach. *IEEE Trans. Neural Networks*, 1997; **8**(1): 98–113
- 16 Zhang S, Xing S. Intelligent recognition of underwater acoustic target noise by multi-feature fusion. 11th International Symposium on Computational Intelligence and Design, IEEE, Hangzhou, China, 2018, **1**: 212–215
- 17 Ke X, Yuan F, Cheng E. Integrated optimization of underwater acoustic ship-radiated noise recognition based on two-dimensional feature fusion. *Appl. Acoust.*, 2020; **159**: 107057
- 18 Baltrušaitis T, Ahuja C, Morency L P. Multimodal machine learning: A survey and taxonomy. *IEEE. Trans. Pattern Anal. Mach. Intell.*, 2018; **41**(2): 423–443
- 19 Jensen F B, Kuperman W A, Porter M B, *et al.* Computational ocean acoustics. 2nd Edition. New York: Springer, 2011: 269
- 20 Cao H, Wang W, Su L, *et al.* Deep transfer learning for underwater direction of arrival using one vector sensor. *J. Acoust. Soc. Am.*, 2021; **149**(3): 1699–1711