

基于阵列的神经网络水声通信信号多参数联合估计算法

成乐¹, 刘悦², 胡正良², 朱宏娜¹, 罗斌¹

(1. 西南交通大学信息科学与技术学院, 四川 成都 610031; 2. 国防科技大学气象海洋学院, 湖南 长沙 410073)

摘要: 针对水声信道复杂多变且衰减严重等问题, 为提升非合作条件下水声通信信号的检测概率并扩大感知范围, 设计了一种新型基于阵列多通道时频谱输入的神经网络多参数联合估计算法。该算法通过引入载波频率标签分配策略, 将载波频率作为区分不同信号的关键物理特征, 有效避免了频带外信号和噪声的干扰; 利用端到端的多任务学习, 能够同时完成信号检测、调制模式识别, 以及对信号个数、载波频率、带宽和波达方向的联合估计, 从而避免了传统算法中需要先进行波束成形再进行检测识别的复杂流程。仿真实验结果表明, 在阵列阵元位置失配和信号被噪声掩蔽的情况下, 所提算法仍能实现准确的信号估计。进一步的湖上实验验证了所提算法的实用性和泛化能力。

关键词: 多参数联合估计; 波达方向估计; 调制模式识别; 阵列信号处理; 神经网络

中图分类号: TN911.7

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2025017

Array-based neural network algorithm for multi-parameter joint estimation of underwater acoustic communication signals

CHENG Le¹, LIU Yue², HU Zhengliang², ZHU Hongna¹, LUO Bin¹

1. School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China

2. College of Meteorology and Oceanology, National University of Defense Technology, Changsha 410073, China

Abstract: To address the challenges posed by complex and highly variable underwater acoustic channel (UWAC) with severe attenuation, a novel neural network algorithm was proposed for multi-parameter joint estimation. Multi-channel spectrograms derived from array signals were utilized by the algorithm to improve the detection probability of UWAC signals under non-cooperative conditions and extend the sensing range. A carrier frequency label assignment strategy was designed, in which carrier frequency served as the key physical feature to distinguish different signals, thereby effectively mitigating interference from out-of-band signals and noise. End-to-end multi-task learning was adopted to simultaneously perform signal detection, modulation recognition, and joint estimation of the number of signal sources, carrier frequency, bandwidth, and direction of arrival, eliminating the complex beamforming process typically required in traditional methods before detection and recognition. Simulation results confirm that accurate signal estimation is reliably achieved even in the presence of array element position mismatches and when signals are obscured by noise. Lake experiments further demonstrate the practicality and generalization capability of the proposed algorithm.

Keywords: multi-parameter joint estimation, direction of arrival estimation, modulation recognition, array signal processing, neural network

收稿日期: 2024-09-03; 修回日期: 2024-12-30

通信作者: 胡正良, zheng_liang_hu@163.com

基金项目: 国家重点研发计划基金资助项目 (No.2021YFC3101402, No.SQ2021YFF0500035)

Foundation Items: The National Key Research and Development Program of China (No.2021YFC3101402, No. SQ2021YFF0500035)

0 引言

近年来,世界各国纷纷投入大量人力物力加强水下信息网络的建设。为了建设水下信息网络、守卫领海安全,亟须加强对水声通信信号的监测。然而,相比于无线电信号^[1-2],水声信号在传播过程中衰减极为严重,典型的水声通信距离通常在 10 km 左右,探测范围受到了极大的限制。而我国海域广袤,为了提升探测效率、扩大监测空间,需要监测更低信噪比(SNR, signal-to-noise ratio)的信号。因此,人们可通过水听器阵列结合波束成形算法利用空间增益来提高信号 SNR,进而增强监测效率和空间感知能力。

然而,传统算法需要先完成波束成形,再进行检测、识别和参数估计,流程复杂且对先验参数依赖较大。特别是在水声环境中,针对非合作水声通信信号的监测,不仅缺乏关于信号数量、载波频率和带宽等的先验信息,还面临因柔性水听器阵列布放困难和海洋动态环境而导致的阵列阵元位置失配问题。这些因素导致常规波束成形(CBF, conventional beamforming)^[3]、最小方差无失真响应(MVDR, minimum variance distortionless response)^[4]和多重信号分类(MUSIC, multiple signal classification)^[5]等经典基于模型驱动的波束成形算法性能显著下降,严重影响信号监测效果。

为此,基于数据驱动的方法引起了学者的广泛关注,该方法通过学习大量数据中的特征模式,在缺乏先验信息的情况下具备更好的鲁棒性。在无线电频谱监测中,许多学者针对单通道信号设计了基于神经网络的多参数估计算法。文献[6-8]利用神经网络目标检测算法,对信号的载波频率、带宽和调制模式等参数进行估计,通过短时傅里叶变换(STFT, short time Fourier transform)将无线电信号转换到二维时频域,并在该域上完成了对信号目标的检测识别和时频坐标定位。然而,这些算法没有利用阵列增益或进行波达方向(DOA, direction of arrival)估计,限制了其在水声环境中的监测能力。

为了充分利用阵列增益,需先对阵列信号进行波束成形。然而,在缺乏先验信息以及阵元位置失配的条件下,传统的模型驱动波束成形算法由于预设的阵列模型与实际阵列模型不匹配,算法性能会严重下降,面临巨大挑战。而数据驱动

的神经网络算法可直接从阵列信号或其变换域中提取相关特征,学习到 DOA 等估计参数的映射关系,成为解决该问题的潜在途径。目前,针对固定频率单频信号的研究较多,如文献[9-10]分别以阵列信号原始波形和协方差矩阵为输入进行 DOA 估计,但这些方法仅考虑 SNR 和波达角为变量,输入特征维度较低,属于窄带 DOA 估计。对于宽带 DOA 估计,文献[11-12]采用小波时频谱结合卷积神经网络(CNN, convolutional neural network),或将宽带信号分为多个子带进行支持向量回归(SVR, support vector regression),然后融合结果进行 DOA 估计。然而,这些方法固定了信号个数和频段等参数,属于合作方 DOA 估计。文献[13-14]利用信号波束或频率方位谱结合 CNN 或目标检测算法进行了非合作方的宽带 DOA 估计,但依赖传统波束成形,面临阵元位置失配等模型失配问题。

在麦克风信号处理领域,类似的联合估计任务被称为声音事件检测与定位(SELD, sound event detection and localization)。SELD 旨在检测和识别室内声音事件并估计其波达角。根据标签分配策略,SELD 算法分为基于分类和基于轨迹的算法^[15]。基于分类标签分配的算法检测声音事件类别并估计其波达角,但无法检测多个相同类别的事件;基于轨迹标签分配的算法估计事件的 DOA 轨迹并对其进行分类,但存在标签分配不变性的问题。这些算法的输入特征主要包括时频谱或梅尔谱^[16-18]、原始波形^[19-20]、可控功率响应和相位变换(SRP-PHAT, steering response power with phase transform)^[21-22]、广义互相关和相位变换(GCC-PHAT, generalized cross correlation with phase transform)^[23-24]等,主要通过 CNN^[19-22]、循环神经网络(RNN, recurrent neural network)^[25-26]、Transformer 网络^[27-28]或它们的组合进行特征提取。其中,基于时频谱或原始波形的 SELD 算法避免了阵列参数相关的特征提取,不需要其他先验信息,因此不存在阵列模型失配的问题。但这些研究主要应用于阵列规模较小、定位精度要求不高、信噪比较高的室内环境。

相比之下,非合作条件下的水声通信信号监测环境更加复杂,信号观测空间的维度更高。首先,为了扩大监测范围、提高检测效率,水听器阵列的阵元数量更多,输入信号的数据维度更高。其次,

水声通信信号的频段未知, 频段分布的不确定性更大。此外, 由于信号接收强度差异较大, 多个信号之间的强度组合会严重影响特征的稳定性。因此, 传统的模型驱动和数据驱动的方法在此场景下都面临巨大的挑战。

针对上述非合作条件下水声通信信号监测面临的信噪比低、先验信息缺乏和阵元位置失配等挑战, 本文提出了一种基于载波频率标签分配策略的多参数联合估计算法。该算法利用端到端的神经网络多任务学习技术, 以阵列信号的多通道时频谱的幅度谱和相邻通道间的相位互谱为输入, 无须先进行波束成形, 能直接完成对信号数量、载波频率、信号带宽、调制模式和 DOA 的联合估计。所提算法通过多通道时频谱和相位互谱的组合, 充分利用了阵列信号的空间和频谱信息, 增强了在低 SNR 条件下的检测和估计性能。由于输入特征的计算与阵列参数无关, 因此算法无须准确的先验信息, 对阵元位置失配具有鲁棒性。此外, 通过基于载波频率的标签分配策略, 所提算法能够聚焦信号频带内的特征, 有效抑制带外信号和噪声的干扰, 提升了泛化性能。

本文主要贡献如下。

1) 提出了一种非合作条件下的神经网络水声通信信号多参数联合估计算法。该算法以水听器阵列信号的多通道时频谱和相位互谱为输入, 无须先验信息, 能有效提升对低 SNR 信号的联合估计性能, 避免了传统算法先波束成形再参数估计的复杂流程。

2) 提出了基于载波频率的标签分配策略。针对多信号复杂场景, 在信号参数完全随机分布、观测空间维度极高的情况下, 基于信号载波频率进行标签分配, 利用有限的信号数据, 建立了鲁棒的多参数联合估计模型。

3) 通过仿真实验和湖上实验验证了所提算法的有效性和鲁棒性。仿真结果表明, 所提算法能够在复杂场景下高效完成对水声通信信号的多参数联合估计。与传统基于模型的 DOA 估计算法相比, 所提算法能够在阵列位置失配的情况下准确完成 DOA 估计; 相比于基于分类标签分配策略的 SELD 算法, 所提算法在多信号条件下具备更好的检测性能。湖上实验进一步证明了所提算法在实际应用场景下的有效性。

1 阵列信号模型

假设有 I 个水声通信信号, 其中第 i 个信号 $s^i(t)$ 的第 n 段发射数据的时频域可表示为

$$S^i(\omega, n) = \sum_{-\infty}^{+\infty} s^i(t) w(t - nT) e^{-j\omega t} \quad (1)$$

其中, ω 为信号的角频率, $w(t - nT)$ 为 STFT 的窗函数。水听器阵列 m 通道的接收信号可表示为

$$X_m(\omega, n) = \begin{bmatrix} X_1(\omega_{-\frac{L}{2}}, n) & \cdots & X_1(\omega_l, n) & \cdots & X_1(\omega_{\frac{L}{2}}, n) \\ \vdots & & \vdots & & \vdots \\ X_m(\omega_{-\frac{L}{2}}, n) & \cdots & X_m(\omega_l, n) & \cdots & X_m(\omega_{\frac{L}{2}}, n) \\ \vdots & & \vdots & & \vdots \\ X_M(\omega_{-\frac{L}{2}}, n) & \cdots & X_M(\omega_l, n) & \cdots & X_M(\omega_{\frac{L}{2}}, n) \end{bmatrix} \quad (2)$$

其中, L 为 STFT 的傅里叶变换长度, M 为阵列的基元数。此时, 阵列接收信号可表示为

$$X_m(\omega, n) = \sum_{i=1}^I S^i(\omega, n) e^{-j\omega \tau_m^i} + N_m(\omega) \quad (3)$$

其中, $N_m(\omega)$ 为阵列 m 基元的接收噪声, τ_m^i 为第 i 个发射信号到达阵列 m 基元相对于参考基元的时延。当阵列为线列阵时, 信号从阵列法向 θ^i 方向入射到第 m 基元

$$\tau_m^i(d_m, \theta^i) = \frac{-d_m \sin(\theta^i)}{c} \quad (4)$$

其中, d_m 为第 m 基元相对于参考基元的距离, c 为信号传播速度。此时, 第 m 基元在频点 ω 处的基阵响应为

$$p_m(\omega) = e^{\frac{j\omega d_m \sin(\theta^i)}{c}} \quad (5)$$

将式(5)代入式(3), 阵列接收信号可进一步表示为

$$X_m(\omega, n) = \sum_{i=1}^I S^i(\omega, n) e^{\frac{j\omega d_m \sin(\theta^i)}{c}} + N_m(\omega) \quad (6)$$

由式(6)可知, 阵列接收信号的时频谱取决于频率 ω 、时间 n 、基元位置 d_m 和信号波达方向 θ^i 。

2 水声通信信号联合估计模型

2.1 模型输入

本文选取阵列接收信号时频谱的幅度谱和相邻通道间的相位互谱作为神经网络模型的输入, 由于傅里叶变换的频域对称性质, 仅选取时频谱的正半

轴频谱。其中, 信号的幅度谱为

$$A_m(\omega, n) = 10\lg\left(\|X_m(\omega, n)\|^2\right) \quad (7)$$

第 2 到 M 通道的相位互谱为

$$P_m(\omega, n) = \angle[X_m(\omega, n)X_{m-1}^*(\omega, n)] \quad (8)$$

其中, $\angle(\cdot)$ 表示提取相位分量。则模型输入为

$$F_{\text{input}}(m, \omega, n) = \text{cat}[A_m(\omega, n); P_m(\omega, n)] \quad (9)$$

其中, $\text{cat}(\cdot)$ 表示在通道维度将所有阵元数据的幅度谱和相位互谱依次拼接。幅度谱和相位互谱包含了信号的频谱特征和波达方向特征, 因而神经网络模型可以从中提取目标信号特征, 并在这些特征与估计目标之间建立映射关系。假设信号和噪声不相关, 当 $I=1$, 即只存在一个信号时, 有

$$A_m(\omega, n) = 10\lg\left(\|S^1(\omega, n)\|^2 + \|N_m(\omega)\|^2\right) \quad (10)$$

$$P_m(\omega, n) = \angle\left[\|S^1(\omega, n)\|^2 e^{\frac{j\omega(d_m - d_{m-1})\sin(\theta^1)}{c}} + N_m(\omega)N_{m-1}^*(\omega)\right] \quad (11)$$

由式(10)和式(11)可知, 当频带内只含有一个信号时, 幅度谱和相位互谱包含了明显的目标特征, 并且特征只与信号和噪声的强度以及波达方向有关。所以针对单频或固定频点信号, 映射关系较为简单。当频带内存在多个信号时, 以 $I=2$ 为例, 有

$$A_m(\omega, n) = 10\lg\left(\left\|\sum_{i=1}^2 S^i(\omega, n) e^{\frac{j\omega(d_m - d_{m-1})\sin(\theta^i)}{c}}\right\|^2 + \|N_m(\omega)\|^2\right) \quad (12)$$

$$P_m(\omega, n) = \angle\left[\|S^1(\omega, n)\|^2 e^{\frac{j\omega(d_m - d_{m-1})\sin(\theta^1)}{c}} + \|S^2(\omega, n)\|^2 e^{\frac{j\omega(d_m - d_{m-1})\sin(\theta^2)}{c}} + S^1(\omega, n) [S^2(\omega, n)]^* e^{\frac{j\omega d_m \sin(\theta^1)}{c}} e^{-\frac{j\omega d_{m-1} \sin(\theta^2)}{c}} + [S^1(\omega, n)]^* S^2(\omega, n) e^{-\frac{j\omega d_m \sin(\theta^1)}{c}} e^{\frac{j\omega d_{m-1} \sin(\theta^2)}{c}} + N_m(\omega)N_{m-1}^*(\omega)\right] \quad (13)$$

由式(12)和式(13)可以看出, 当一个频带内存在多个信号时, 信号的频谱特征和相位互谱特征不是简单的线性叠加, 而是与信号个数、信号强度、噪声强度、信号初相、波达方向等多种因素有关, 形成了一个复杂的非线性关系。因此, 对于多个水声通信信号检测的复杂场景, 不再能如文献[9]或文献[10]所述穷举不同信号场景。文献[18]提出了期望模型, 从只存在一个信号的频点进行特征提取。基于此, 本文提出了基于载波频率的标签分配策略, 从不同信号单独所占频点进行特征提取并进行联合估计。

2.2 主干网络

根据通信信号调制技术, 信号的幅度谱特征具备时频移不变性。这意味着, 同一信号在不同时间或被搬移至不同频段时, 其幅度谱特征不会发生改变。而 CNN 具有平移不变性和参数共享的特性, 因此, 在二维时频平面上, CNN 特别适用于提取幅度谱信号的时频不变特征。然而, 由式(11)可知, 通道间的相位互谱包含的 DOA 特征不仅与波达方向相关, 还与信号频率相关; 也就是说, 在二维时频平面的频率维度上, DOA 特征不具备平移不变性。因此本文不使用二维 CNN 直接提取信号时空频特征, 而是采用一维 CNN 分阶段提取信号时空域特征和频域特征。

水声通信信号多参数联合估计模型及其主干网络如图 1 所示。输入特征 F_{input} 的维度为 $(2M-1) \times W_t \times H_f$, 分别表示通道维度、时间维度和频率维度的大小。主干网络首先对通道维度和时间维度进行 K 次卷积池化操作 $\text{Conv}_t(\cdot)$, 以对时间维度的信息进行完全聚合, 对各频点的 DOA 通道特征分别进行非线性映射, 得到维度为 $\text{ch} \times 1 \times H_f$ 的输出特征。然后再通过 Q 次卷积池化操作 $\text{Conv}_f(\cdot)$ 聚合邻近频点信息, 提取信号频带特征, 经过 Q 次 2 倍下采样后得到输出特征 F_{output} , 其维度为 $\text{ch} \times 1 \times \frac{H_f}{2^Q}$ 。整个主干网络的特征提取过程描述如下

$$F_{\text{output}} = \text{Conv}_f\left(\text{Conv}_t\left(F_{\text{input}}\right)\right) \quad (14)$$

主干网络通过先聚合时间特征并对 DOA 信息进行非线性映射, 再聚合邻近频点频率特征的方式, 避免了直接聚合不具备频移不变性的方位特征; 经过实验, 这种方式能有效提升模型的泛化能力。

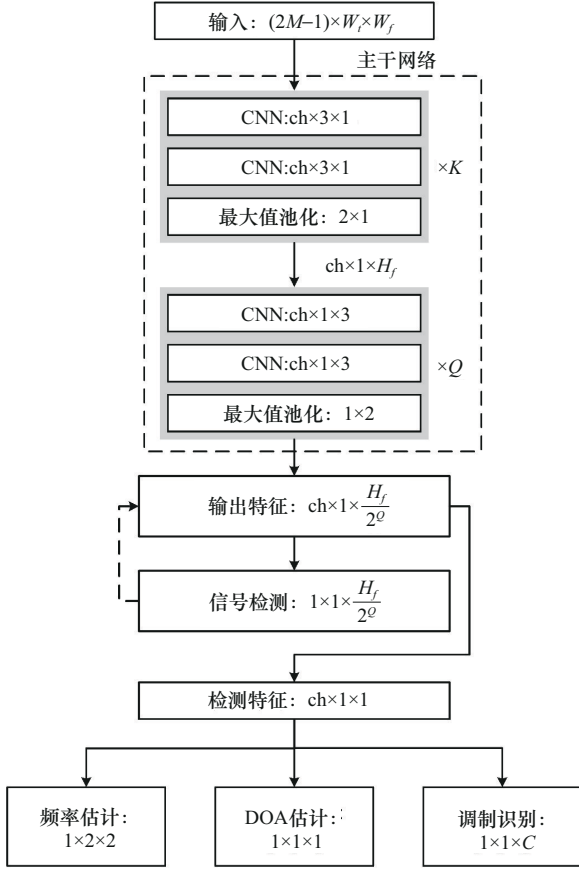


图1 水声通信信号多参数联合估计模型及其主干网络

2.3 基于载波频率的标签分配策略

相比于音频信号，水声通信信号的分布频段更广，且信号间的强度差异更大。为了避免不同频段信号之间相互干扰，本文提出了基于载波频率的标签分配策略。相较于 SELD 算法中采用信号类别或波达角作为特征，所提策略利用载波频率作为区分不同信号的关键物理特征，通过频谱划分将信号分离到各自的子频带，从而有效减少不同信号间的交叉干扰。此时，由于 CNN 的局部和多通道特征提取能力，载波频率范围内的信号特征直接关联于信号的类别、DOA 和带宽等联合特征，便于后续的多参数估计。

如图 2 所示，模型首先对每个子频带进行信号存在性检测，随后根据检测结果提取对应子频带特征进行下一步联合估计。经过主干网络后，输出特征的维度为 $ch \times 1 \times \frac{H_f}{2^Q}$ ；将该特征通过一个信号检测子网络，可得到维度为 $1 \times 1 \times \frac{H_f}{2^Q}$ 的子频带检测结果。图 2 中假设 $\frac{H_f}{2^Q} = 8$ ，即整个频带被等间隔划分

为 8 个子频带，每个频带的检测结果表示对应频带内是否存在载波频率在该范围内的信号。假设采样率为 f_s ，则对应子频带范围表示为

$$\left((i-1) \frac{2^{Q-1} f_s}{H_f}, i \frac{2^{Q-1} f_s}{H_f} \right), i = 1, 2, \dots, \frac{H_f}{2^Q} \quad (15)$$

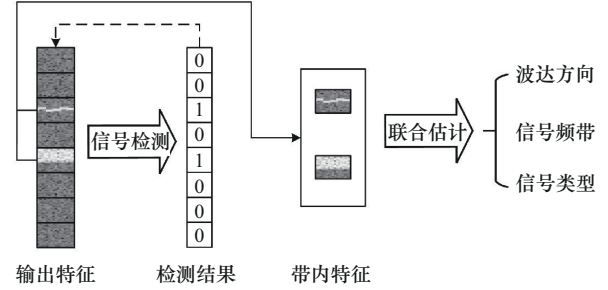


图2 基于载波频率的标签分配联合估计过程

如果在对应频带内检测到信号，则提取出该频带内的特征，称为带内特征，维度为 $ch \times 1 \times 1$ 。带内特征仅包含当前载波频率信号的特征，因此能够有效避免频带外的信号和噪声干扰。在获取信号带内特征后，将其输入信号联合估计子网络中，分别进行波达方向、信号载波频率和带宽以及信号类型的预测。本文所提算法是一个二阶段的检测估计框架，在信号检测阶段对每个子频带进行信号存在性检测，确定哪些子频带内存在目标信号。然后，在参数估计阶段对带内特征进行多参数联合估计。通过这种基于载波频率的标签分配策略，算法能够有效聚焦于目标信号的频谱特征和方位信息，抑制带外信号和噪声的干扰，从而显著提升在复杂水声环境中的泛化性能和检测估计能力。

2.4 多参数联合估计输出和损失函数

经过主干网络后，得到的输出特征 F_{output} 维度为 $ch \times 1 \times \frac{H_f}{2^Q}$ 。将输出特征通过一个输出通道维度为 1 的 CNN 层（信号检测子网络），得到信号检测输出 F_{detc} ，其维度为 $1 \times 1 \times \frac{H_f}{2^Q}$ 。设子带是否存在信号的标签为 \bar{F}_{detc} ，则信号检测损失函数为

$$\ell_{detc} = \text{MSE}(\text{sigmoid}(F_{detc}), \bar{F}_{detc}) \quad (16)$$

其中， $\text{MSE}(\cdot)$ 表示均方误差损失函数， $\text{sigmoid}(\cdot)$ 表示 sigmoid 激活函数。根据检测结果，得到对应的带内特征后，可通过 3 个全连接子网络分别完成对信号 DOA、频带以及调制模式的估计。假设 DOA 估计的目标空间 $\theta_o \in [-90^\circ, 90^\circ]$ ，观测空间是

接收阵列的时空响应信号；设观测空间经过模型和 DOA 估计子网络的输出为 F_{DOA} ，当频带内的波达方向真值为 $\bar{\theta}$ 时，DOA 估计损失函数为

$$\ell_{\text{DOA}} = \text{MSE} \left[\tanh(F_{\text{DOA}}), \frac{\bar{\theta}}{\max(\theta_o)} \right] \quad (17)$$

其中， $\tanh(\cdot)$ 表示 \tanh 激活函数。信号的频带估计被定义为真实高/低频频点到子带起始频点（如式(15)所示）差值的归一化绝对值 \bar{F}_{fre} 。带内特征经过一个频带估计子网络后，得到频带估计 F_{fre} ，其维度为 $1 \times 1 \times 2$ ，分别表示对低频频点和高频频点的估计，损失函数为

$$\ell_{\text{fre}} = \text{MSE} \left[\text{sigmoid}(F_{\text{fre}}), \bar{F}_{\text{fre}} \right] \quad (18)$$

带内特征经过一个调制识别子网络完成对信号调制模式的识别，其输出为 F_{modu} ，维度为 $1 \times 1 \times C$ ，其中 C 为调制类别的种类。设其真实调制模式标签为 \bar{F}_{modu} ，则信号调制识别估计损失函数为

$$\ell_{\text{modu}} = \text{CE} \left[\text{softmax}(F_{\text{modu}}), \bar{F}_{\text{modu}} \right] \quad (19)$$

其中， $\text{CE}(\cdot)$ 表示交叉熵损失函数， $\text{softmax}(\cdot)$ 为 softmax 激活函数。训练过程中对信号检测、DOA 估计、频带估计和调制识别进行联合优化，提取信号联合特征，其总损失函数为

$$\mathcal{L} = \ell_{\text{detc}} + \ell_{\text{DOA}} + \ell_{\text{fre}} + \beta \ell_{\text{modu}} \quad (20)$$

其中， β 为损失函数权重系数，本文所有实验中设置 $\beta = 0.1$ 。

3 仿真实验

为验证本文所提算法的有效性，本节首先介绍仿真信号的生成过程和算法实现的细节，然后在仿真数据上验证算法对水声通信信号的多参数联合估计能力，最后测试阵列位置误差对算法模型的影响。

3.1 仿真数据生成

假设有一条 32 基元的水听器阵列，采样率为 64 kHz，在没有位置误差时，阵元间距均为 0.075 m。则当声速为 1 500 m/s 时，阵列的中心频率为 10 kHz。本文针对水声通信中常用的 2FSK、4FSK、8FSK、PSK、LFM、OFDM 调制方式进行调制模式识别。在信号生成过程中，结合实际水声通信的常用参数，仿真实验中水声通信信号的具体参数设置如表 1 所示，设置水声信道为高斯白噪声信道，所有参数均在预定区间内服从均匀随机分布，以确保信号的多样性和随机性。为了生成不同 SNR 的信号样本，假设环境噪声能量保持恒定；根据 SNR 生成对

应能量的信号，最后将噪声和所有信号叠加，得到一个接收样本。由于当信号频率远大于阵列中心频率时会产生严重的栅瓣，影响信号检测性能，因此在仿真实验中设置信号载波频率为 3~10 kHz、信号带宽为 1~5 kHz，一次接收的信号中最大 SNR 差距小于 20 dB。此外，为了便于与 SELD 算法进行对比，在信号生成过程中确保信号频带不会完全重叠，并且一次不会接收同一调制的信号。

表 1 水声通信信号的具体参数设置

参数名称	参数值
DOA 范围	$-65^\circ \sim 65^\circ$
载波频率/kHz	3~10
信号带宽/kHz	2~5
SNR/dB	-10~30
最大 SNR 差距/dB	20
信号源个数/个	1~3

3.2 对比方法和实验细节

本文使用水声定位中常用的 MVDR 和 MUSIC 算法作为信号检测和 DOA 估计的对比算法，这些算法通过峰值搜索进行信号检测；使用基于分类标签分配的 SELD 算法作为多参数联合估计的对比算法，该算法的输入特征、主干网络等参数均与本文算法相同。

在所有仿真实验中，综合计算复杂度和频率分辨率，设置 STFT 的傅里叶变换长度为 512，处理快拍数为 32；仅取时频谱频率轴的正半轴，因此输入特征的维度为 $63 \times 32 \times 256$ 。设置水声通信信号联合估计模型中 $K=5$ ， $Q=4$ ， $\text{ch}=128$ 。模型使用随机梯度下降（SGD, stochastic gradient descent）法进行优化，以学习率 10^{-3} 、批大小（batch size）16 训练 120 轮，在确保模型收敛的情况下选取测试性能最佳的模型进行对比测试。所有的实验均在 NVIDIA RTX 4090 显卡和 Pytorch 2.1 深度学习框架下进行。

在测试过程中，当信号估计的 DOA 和真实 DOA 的绝对误差小于 6° 时，视为一次成功检测。算法检测性能的评估标准为 F1 分数（F1 Score）。本文使用绝对平均误差作为 DOA 估计和频带估计的性能评估标准，使用识别准确率评估调制模式识别的性能。频带估计误差定义为高频频点与低频频点的估计值和实际值之间绝对误差的平均值。DOA 估计、频带估计和调制模式识别仅在信号成

功检测时进行评估。

3.3 信号多参数联合估计结果分析

实验中的训练和测试信号数据的参数设置均如表 1 所示, 随机生成了 10 000 个信号样本用于训练和 10 000 个信号样本用于测试。本文所提算法和 SELD 算法的水声通信信号多参数联合估计结果如图 3 所示, 其中, 调制模式 A~F 分别表示 2FSK、4FSK、8FSK、PSK、LFM 和 OFDM。实验分别对算法的检测性能、DOA 估计性能、调制模式识别

性能和频带估计性能进行了分析, 并探讨了各个预测结果与信号 SNR、调制模式、信号个数、载波频率和信号带宽之间的关系。

如图 3(a)所示, 信号的检测 F1 分数随着信噪比的增加而提高, 并在信噪比大于 -5 dB 后检测性能趋于稳定。在各个 SNR 下, 本文所提算法的检测 F1 分数均比 SELD 算法高约 0.04。可以看到, 相对于 MFSK 信号, 本文所提算法和 SELD 算法对 PSK、LFM 和 OFDM 这类频带连续信号的检测效果

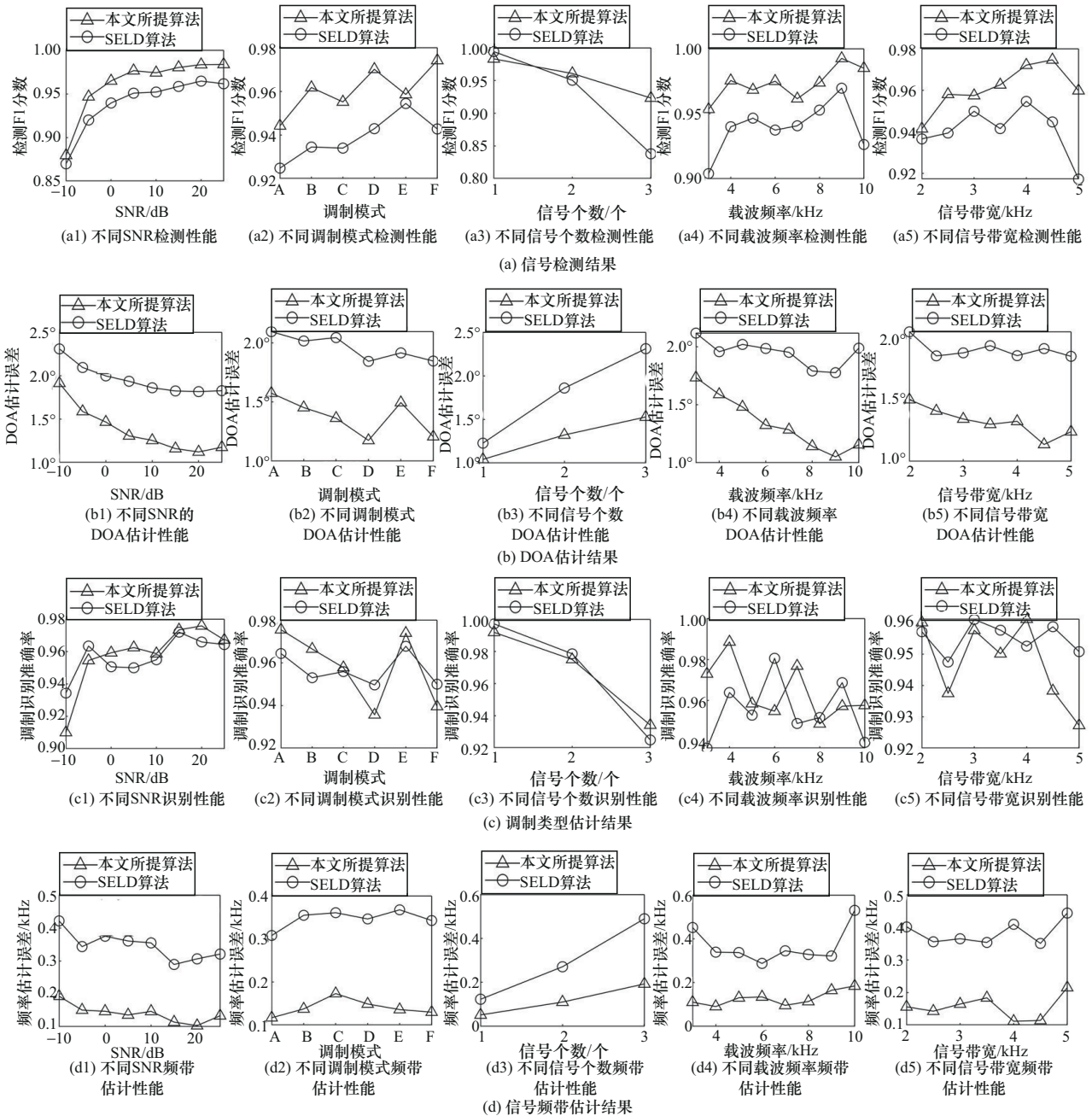


图 3 水声通信信号多参数联合估计结果

更好。值得注意的是，当信号在单通道时频谱上完全被掩蔽（小于 0 dB）时，本文所提算法与 SELD 算法仍能够实现有效的信号检测，相比于基于单通道时频谱的算法有很大提升，在信噪比为-10 dB 的条件下，检测 F1 分数仍可分别达到 0.88 和 0.87。随着信号个数的增加，算法的检测性能出现了明显下降。当信号个数从 2 增加到 3 时，本文所提算法性能下降明显小于 SELD 算法，本文所提算法的检测分数下降了 0.037，而 SELD 算法下降了 0.117。这是因为本文所提算法利用信号的频带特征进行决策，不同频带的信号特征之间干扰较小。信号的检测性能随着载波频率的增加而提升，这是由于信号频率越接近阵列中心频率，阵列增益效果越明显。检测效果与信号带宽的相关性不大，但在信号带宽为 2 kHz 和 5 kHz 时，信号的检测性能较差，这是因为该带宽的训练数据较少，信号的带宽满足均匀随机分布，因此在该带宽处信号数据相对较少。

由图 3(b) 可以看到，本文所提算法和 SELD 算法的 DOA 估计误差都随 SNR 的增大而减小，并且随信号个数的增多而增大；本文所提算法在测试集上的平均 DOA 估计误差约为 1.37°，而 SELD 算法的估计误差约为 1.97°。与检测性能相似，针对 MFSK 信号的 DOA 估计性能较差，但针对 LFM 信号的估计性能也较差，这是因为 LFM 信号的频率随时间变化，频带特征的提取更加困难。可以看到，DOA 估计误差随载波频率和信号带宽的增大有明显的减小趋势，这是由于频率越高越靠近阵列的中心频率，信号的波达方向特征越明显；带宽越宽则信号的频带增益更大，因而性能更好。

由图 3(c) 可以看出，本文所提算法的识别性能和 SELD 算法相当，且均与信噪比呈正相关，与信号个数呈负相关，两者的识别准确率都约为 0.96。2 种算法对 PSK 和 OFDM 信号的识别性能较差，这是因为 PSK 和 OFDM 属于宽带连续信号，其时频特征容易发生混淆，而 MFSK 和 LFM 信号在时频

谱上具有较高的区分度。值得注意的是，识别准确率和信号带宽相关，因为实验中 PSK 等宽带连续信号的带宽设置较宽，而 MFSK 的带宽较窄。

由图 3(d) 可以看出，2 种算法的频带估计能力也随 SNR 的增高而增强，随信号个数的增多而下降，且其估计能力与调制模式的相关度不大。本文所提算法的频带估计误差约为 0.141 kHz，而 SELD 算法的频带估计误差约为 0.347 kHz。它们的估计性能随载波频率和带宽的升高略有下降，这是由于频带越宽，对神经网络感受野的要求越高，受到邻近信号的干扰也就越多，从而导致性能有所下降。

本文所提算法和 SELD 算法的信源个数估计混淆矩阵如图 4 所示。当只存在一个信号时，本文所提算法和 SELD 算法都有很好的预测性能。值得注意的是，对于本文所提算法而言，当预测信号个数大于真实信号个数时，预测到了多个标签分配的载波频率对应一个信号，此时它们的预测结果都是正确的。当信号个数为 1 和 2 时，本文所提算法和 SELD 算法的预测性能相当；当信号个数为 3 时，本文所提算法的预测准确率比 SELD 算法高了 14%。可见，基于载波频率的标签分配策略更适用于多信号的联合估计。

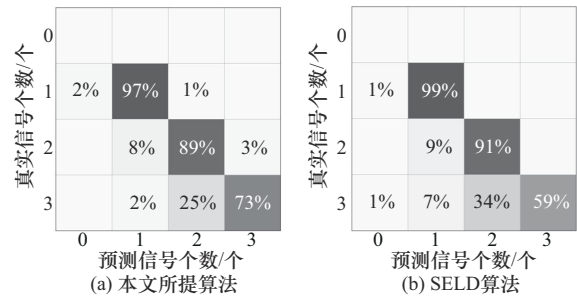


图 4 信源个数估计混淆矩阵

本节将本文所提算法和 SELD 算法与基于模型的 MVDR 和 MUSIC 算法进行了信号检测和 DOA 估计性能对比，结果如表 2 所示。由于 MVDR 和

表 2 单个水声通信信号多参数联合预测结果

参数	检测精确率	检测召回率	检测 F1 分数	DOA 估计误差	识别准确率	频带估计误差/kHz
本文算法	0.987	0.980	0.984	1.050°	0.990	0.050
SELD	0.991	0.991	0.991	1.240°	0.990	0.127
MVDR	0.999	0.999	0.999	0.090°	—	—
MUSIC	0.999	0.999	0.999	0.026°	—	—

MUSIC算法使用峰值搜索算法进行信号检测,因此在低信噪比、多信号或出现栅瓣干扰的情况下,会出现许多伪峰,导致性能显著下降。因此,为了降低峰值搜索算法对MVDR和MUSIC性能的影响,本文仅选择单信号情况进行对比。从表2可以看出,由于不存在模型失配且噪声为理想的高斯白噪声, MVDR算法和MUSIC算法信号的检测和DOA估计性能都要优于本文所提算法和SELD算法。不过,在完全非合作条件下,基于数据驱动的本文所提算法与SELD算法也展现出优异的联合估计性能。接下来,将进一步评估在阵列位置失配的情况下,本文所提算法与MVDR、MUSIC算法在信号检测和DOA估计性能上的差异。

3.4 不同阵列位置误差下的DOA估计分析

柔性水听器阵列布放到水下后由于施工难度和水文因素通常会产生一定弯曲,此时阵元位置会出现误差,从而对水声通信信号检测和DOA估计产生影响。如图5所示,假设一条均匀线阵间距为 d ,算法模型通常假设其布放方式如图5中虚线基元所示,即均匀直线分布,因此当阵列弯曲后就会发生阵元位置失配。本文假设第 $M-1$ 基元和第 M 基元的连线与 X 轴之间的弯曲角为 η_M 。

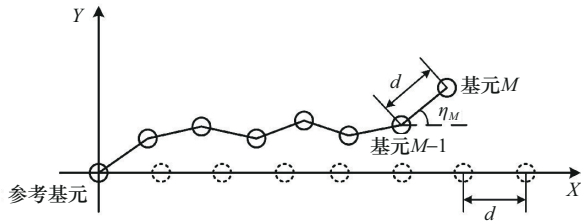


图5 阵元位置失配示意

当 $\eta_M=0$ 时,阵列为均匀线阵;当 $\eta_M \neq 0$ 时,阵元位置 (x, y) 可表示为

$$x_m = \sum_{i=1}^m d \cos \eta_i, y_m = \sum_{i=1}^m d \sin \eta_i \quad (21)$$

假设 η_M 在角度弯曲区间 $[-\eta_{\max}, \eta_{\max}]$ 内为均匀随机分布,其余阵列参数设置与信号参数设置同3.1节。表3展示了在不同角度弯曲区间内本文所提算法、MVDR和MUSIC算法在单信号上的检测和DOA估计结果,以及本文所提算法在多信号上的检测和DOA估计结果。随着 η_{\max} 的增大,阵列的弯曲程度更大,阵型失配也越严重。

从表3可以看出,本文所提算法在单信号和多信号情况下,检测性能和DOA估计性能均不受阵列位置失配的影响。单信号情况下,阵列的检测F1分数始终约为0.98,DOA估计误差约为 1.1° ;多信号情况下,阵列的检测F1分数约为0.94,DOA估计误差约为 1.4° ,其性能与没有阵列位置误差时基本一致。然而,MVDR和MUSIC算法的性能受阵列位置误差影响严重。当 $\eta_{\max} < 30^\circ$ 时,模型失配不严重,MVDR和MUSIC算法的检测估计性能优于本文所提算法。但当阵列位置误差继续增大后,模型驱动的算法性能急剧下降。当 $\eta_{\max} > 30^\circ$ 后,基于数据驱动的本文所提算法在检测性能和DOA估计性能上都要优于模型驱动算法。由表3可知,本文所提算法基于信号数据驱动,估计性能对阵列位置误差鲁棒。

4 湖上实验

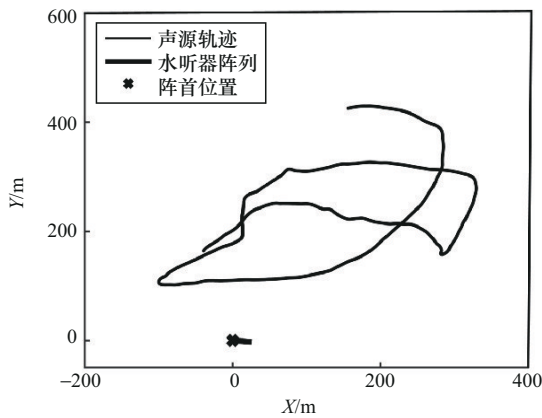
为了进一步验证本文所提算法的实用性和泛化能力,本节通过湖上实验进行测试。湖上实验使用了一条23基元的光纤水听器阵列^[29],采样率为64 kHz,阵元间距为0.075 m。实验过程中,使用的信号发射水声换能器工作频带为7~14 kHz,因此设置水声通信信号的载波频率为9~12 kHz。水声换能器一次只发射一个信号,阵列接收信号的信噪比约为20 dB。

表3 不同角度弯曲区间内的检测和DOA估计结果

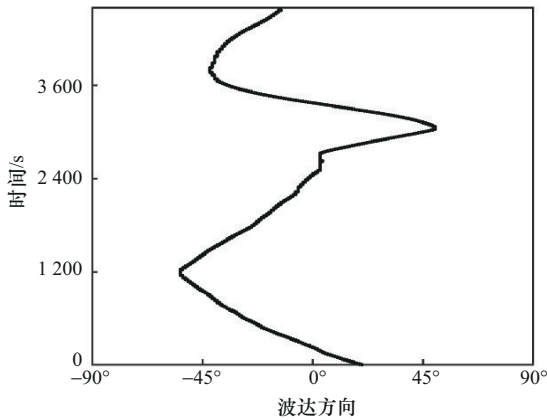
算法	[0°, 0°]		[-10°, 10°]		[-20°, 20°]		[-30°, 30°]		[-40°, 40°]		[-50°, 50°]	
	检测 F1分数	DOA 估计误差	检测 F1分数	DOA 估计误差	检测 F1分数	DOA 估计误差	检测 F1分数	DOA 估计误差	检测 F1分数	DOA 估计误差	检测 F1分数	DOA 估计误差
本文所提算法	0.98	1.05°	0.98	1.21°	0.98	1.18°	0.98	1.19°	0.98	1.16°	0.98	1.16°
MVDR	0.99	0.09°	1.00	0.25°	0.99	1.03°	0.98	2.12°	0.78	2.70°	0.60	2.84°
MUSIC	0.99	0.03°	0.99	0.24°	0.99	1.04°	0.99	2.20°	0.78	2.75°	0.60	2.88°
本文所提算法 (多信号)	0.95	1.37°	0.94	1.49°	0.93	1.47°	0.94	1.49°	0.93	1.45°	0.93	1.47°

实验中共采集了 12 200 个样本，其中 50% 用于训练，50% 用于测试，其余信号参数和模型的设置与仿真实验相同。

实验过程中，光纤水听器阵列悬挂在水下 5 m 处，接收水声通信信号。水声换能器以 4 节的速度进行拖曳，约在水下 2 m，间歇性发射水声通信信号。拖曳过程中，装有 GPS 的实验船在湖面按照既定航线航行，水听器阵列的布放位置和声源船的 GPS 航迹示意如图 6(a) 所示，测试过程中的航行 GPS 方位历程如图 6(b) 所示。



(a) 水听器阵列的布放位置和声源船的 GPS 航迹示意



(b) 航行 GPS 方位历程

图 6 湖上实验布局示意

本文所提算法在测试集上的联合估计结果如表 4 所示。由表 4 可知，即便在真实的水声通信条件下，本文所提算法仍能进行有效的联合估计。然而，由于水声通信信道中严重的时间和频率衰落以及其他环境干扰，算法的调制模式识别性能略有下降。水声通信信号调制识别结果的归一化混淆矩阵如图 7 所示。当信号环境变得复杂后，频带特征差异减小，导致在 PSK 和 OFDM 这类宽带连续信号

的识别上发生了较严重的混淆。湖上实验环境相对于仿真实验信道更加复杂，在不同时刻存在不同程度的环境干扰，难以通过峰值搜索法对信号进行有效检测。因此，通过图 8 展示测试过程中第 2 590 ~ 3 590 s 数据的归一化方位历程。实验船的 GPS 方位历程如图 8(a) 所示，CBF、MVDR、MUSIC 和本文所提算法的估计结果分别如图 8(b)~图 8(e) 所示，其中，亮度越高表示相对能量越高。GPS 持续记录而信号间歇发射，所以信号的波束历程估计结果都存在断续。从 CBF 和 MUSIC 的估计结果中可以明显看出，在 2 700~3 200 s 的 $-75^{\circ}\sim 0^{\circ}$ 范围内存在一个干扰，这一干扰是一艘航行的渔船，而非水声通信信号。然而，如图 8(e) 所示，本文所提算法能够对很好地对渔船干扰进行抑制，这是因为它检测对象始终为水声通信信号，能够根据频谱特征抑制非水声通信信号目标的干扰。此外，根据方位历程估计结果，本文所提算法受信号能量影响小且主瓣窄，其估计轨迹相对于其他算法更加清晰。可见，本文所提算法即便在真实水声通信环境中，也能够基于水声通信信号数据驱动有效完成信号多参数联合估计。

表 4 本文所提算法在测试集上的联合估计结果

参数	结果
检测精确率	0.993
检测召回率	0.987
检测 F1 分数	0.990
DOA 估计误差	0.986°
识别准确率	0.970
频带估计误差/kHz	0.042

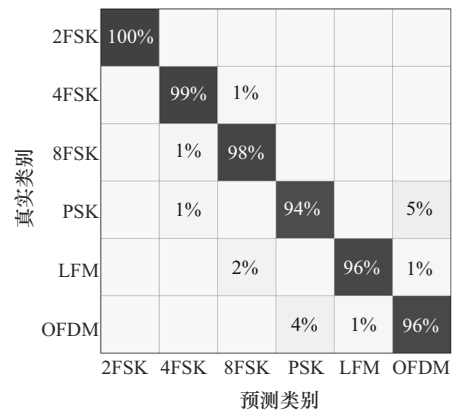


图 7 水声通信信号调制识别结果的归一化混淆矩阵

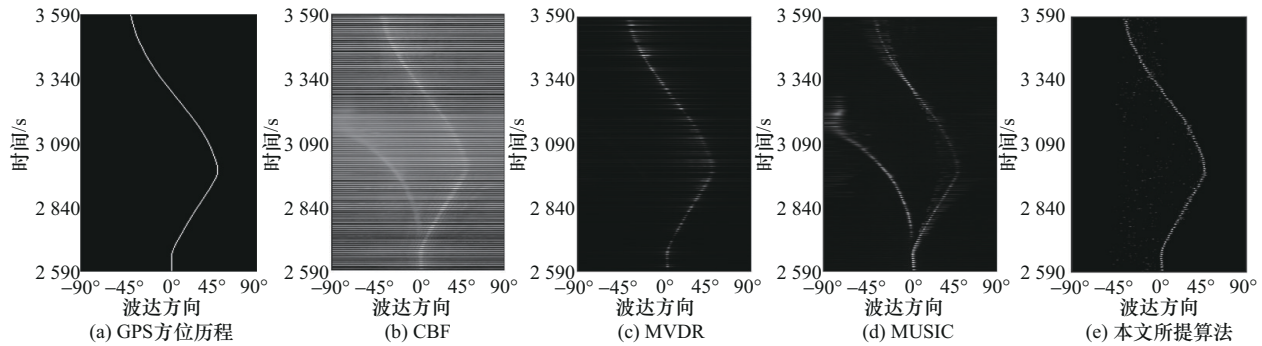


图8 湖上实验水声通信信号方位历程和估计结果

5 结束语

针对非合作条件下的水声通信信号监测问题,本文利用数据驱动的神经网络技术,以水听器阵列信号的多通道时频谱为输入,实现了水声通信信号的多参数联合估计。本文所提算法能够同时完成信号检测、信号个数估计、载波频率和带宽估计、调制模式识别以及DOA估计,有效克服了传统算法需先完成波束成形再进行检测识别的复杂过程,并充分利用了阵列增益,提升了监测效率和空间感知能力。针对水声通信信号多通道时频谱的分布特点,本文提出了基于载波频率的标签分配策略和分阶段提取信号时空域与频域特征的主干网络,显著提升了算法在多信号监测场景中的检测与估计性能以及泛化能力。相较于SELD算法,在信号个数为3的情况下,检测F1分数提高了0.085,DOA估计误差降低了 0.79° 。

此外,针对水声阵列阵元位置失配严重的情况,本文所提算法仍能准确完成多参数联合估计,扩展了算法在复杂水声环境下的应用场景。在水听器阵列角度随机弯曲区间为 $[-50^\circ, 50^\circ]$ 时,相比于传统算法,本文所提算法检测F1分数提高了0.38,DOA估计误差降低了 1.68° 。湖上实验进一步表明,本文所提算法在实际水声环境中具备很强的实用性和泛化性,且能有效地抑制船舶噪声等其他干扰。本文所提算法不仅适用于水声通信信号处理,还能够应用于其他关注频带信息的阵列信号处理场景,如声纳系统、无线电天线阵列和麦克风阵列等。

参考文献:

- [1] KULIN M, KAZAZ T, MOERMAN I, et al. End-to-end learning from spectrum data: a deep learning approach for wireless signal identification in spectrum monitoring applications[J]. IEEE Access, 2018, 6: 18484-18501.
- [2] KAPOOR S, SINGH G. Non-cooperative spectrum sensing: a hybrid model approach[C]//Proceedings of the 2011 International Conference on Devices and Communications (ICDeCom). Piscataway: IEEE Press, 2011: 1-5.
- [3] BARTLETT M S. Properties of sufficiency and statistical tests[J]. Proceedings of the Royal Society Series A-Mathematical and Physical Sciences, 1937, 160(901): 268-282.
- [4] CAPON J. High-resolution frequency-wavenumber spectrum analysis[J]. Proceedings of the IEEE, 1969, 57(8): 1408-1418.
- [5] SCHMIDT R. Multiple emitter location and signal parameter estimation[J]. IEEE Transactions on Antennas and Propagation, 1986, 34(3): 276-280.
- [6] 周鑫, 何晓新, 郑昌文. 基于图像深度学习的无线电信号识别[J]. 通信学报, 2019, 40(7): 114-125.
- [7] ZHOU X, HE X X, ZHENG C W. Radio signal recognition based on image deep learning[J]. Journal on Communications, 2019, 40(7): 114-125.
- [8] LI W H, WANG K R, YOU L, et al. A new deep learning framework for HF signal detection in wideband spectrogram[J]. IEEE Signal Processing Letters, 2022, 29: 1342-1346.
- [9] CHENG L, ZHU H N, HU Z L, et al. A sequence-to-sequence model for online signal detection and format recognition[J]. IEEE Signal Processing Letters, 2024, 31: 994-998.
- [10] ZHENG S L, YANG Z, SHEN W G, et al. Deep learning-based DOA estimation[J]. IEEE Transactions on Cognitive Communications and Networking, 2024, 10(3): 819-835.
- [11] PAPAGEORGIU G K, SELLATHURAI M, ELDAR Y C. Deep networks for direction-of-arrival estimation in low SNR[J]. IEEE Transactions on Signal Processing, 2021, 69: 3714-3729.
- [12] WANG J J, QUAN T Q, JIAO L L, et al. DOA estimation of underwater acoustic array signal based on wavelet transform with double branch convolutional neural network[J]. IEEE Transactions on Vehicular Technology, 2023, 72(5): 5962-5972.
- [13] WU L L, HUANG Z T. Coherent SVR learning for wideband direction-of-arrival estimation[J]. IEEE Signal Processing Letters, 2019, 26(4): 642-646.
- [14] NIE W H, ZHANG X W, XU J, et al. Adaptive direction-of-arrival estimation using deep neural network in marine acoustic environment[J]. IEEE Sensors Journal, 2023, 23(13): 15093-15105.
- [15] CHENG L, LIU Y, ZHANG B B, et al. Direction of arrival joint prediction of underwater acoustic communication signals using faster R-

- CNN and frequency - azimuth spectrum[J]. Remote Sensing, 2024, 16 (14): 2563.
- [15] NGUYEN T N T, WATCHARASUPAT K N, NGUYEN N K, et al. SALSA: spatial cue-augmented log-spectrogram features for polyphonic sound event localization and detection[J]. IEEE/ACM Transactions on Audio, Speech and Language Processing, 2022, 30: 1749-1762.
- [16] ADAVANNE S, POLITIS A, NIKUNEN J, et al. Sound event localization and detection of overlapping sources using convolutional recurrent neural networks[J]. IEEE Journal of Selected Topics in Signal Processing, 2019, 13(1): 34-48.
- [17] SCHYMURA C, BÖNNINGHOFF B, OCHIAI T, et al. PILOT: introducing transformers for probabilistic sound event localization[C]//Proceedings of the 2021 Conference of the International Speech Communication Association. Piscataway: IEEE Press, 2021: 2117-2121.
- [18] CHAKRABARTY S, HABETS E A P. Multi-speaker DOA estimation using deep convolutional networks trained with noise signals[J]. IEEE Journal of Selected Topics in Signal Processing, 2019, 13(1): 8-21.
- [19] SUNDAR H, WANG W R, SUN M, et al. Raw waveform based end-to-end deep convolutional network for spatial localization of multiple acoustic sources[C]//Proceedings of the ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2020: 4642-4646.
- [20] HE Y, TRIGONI N, MARKHAM A. SoundDet: polyphonic moving sound event detection and localization from raw waveform[C]//Proceedings of the 2021 International Conference on Machine Learning (ICML). Piscataway: IEEE Press, 2021: 4160-4170.
- [21] CHAKRABARTY S, HABETS E A P. Broadband DOA estimation using convolutional neural networks trained with noise signals[C]//Proceedings of the 2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA). Piscataway: IEEE Press, 2017: 136-140.
- [22] DIAZ-GUERRA D, MIGUEL A, BELTRAN J R. Robust sound source tracking using SRP-PHAT and 3D convolutional neural networks[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2020, 29: 300-311.
- [23] COMANDUCCI L, BORRA F, BESTAGINI P, et al. Source localization using distributed microphones in reverberant environments based on deep learning and ray space transform[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2020, 28: 2238-2251.
- [24] VERA-DIAZ J M, PIZARRO D, MACIAS-GUARASA J. Towards domain independence in CNN-based acoustic localization using deep cross correlations[C]//Proceedings of the 2020 28th European Signal Processing Conference (EUSIPCO). Piscataway: IEEE Press, 2021: 226-230.
- [25] CONG J Y, WANG X P, HUANG M X, et al. Robust DOA estimation method for MIMO radar via deep neural networks[J]. IEEE Sensors Journal, 2021, 21(6): 7498-7507.
- [26] MERKOFER J P, REVACH G, SHLEZINGER N, et al. DA-MUSIC: data-driven DoA estimation via deep augmented MUSIC algorithm[J]. IEEE Transactions on Vehicular Technology, 2024, 73(2): 2771-2785.
- [27] GUO Y, ZHANG Z, HUANG Y Z. Dual class token vision transformer for direction of arrival estimation in low SNR[J]. IEEE Signal Processing Letters, 2023, 31: 76-80.
- [28] OCHIAI T, DELCROIX M, NAKATANI T, et al. Mask-based neural beamforming for moving speakers with self-attention-based tracking[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2023, 31: 835-848.
- [29] LIU Y, CHENG L, ZOU Y C, et al. Thin fiber-optic hydrophone towed array for autonomous underwater vehicle[J]. IEEE Sensors Journal, 2024, 24(9): 15125-15132.

[作者简介]



成乐 (1996-), 男, 湖南怀化人, 西南交通大学博士生, 主要研究方向为通信信号检测、阵列信号处理、神经网络等。



刘悦 (1991-), 男, 湖北荆门人, 国防科技大学博士生, 主要研究方向为光纤传感、阵列信号处理等。



胡正良 (1975-), 男, 湖南长沙人, 国防科技大学教授、硕士生导师, 主要研究方向为光纤传感、声信号处理等。



朱宏娜 (1979-), 女, 山东烟台人, 西南交通大学教授、博士生导师, 主要研究方向为光纤通信、机器学习及智能信息处理等。

罗斌 (1968-), 男, 四川成都人, 西南交通大学教授、博士生导师, 主要研究方向为光纤通信、光电子技术等。