

基于噪声自适应加权的GCC时延估计算法

甘坦*, 康绍绕, 熊林

江西理工大学理学院, 江西 赣州

收稿日期: 2026年2月27日; 录用日期: 2026年3月19日; 发布日期: 2026年3月30日

摘要

在强噪声环境下, 传统GCC时延估计方法性能会急剧下降, 主要原因是噪声会干扰信号的互相关函数, 造成旁瓣抬升和主瓣展宽。此时, 峰值所对应的时延会出现偏差。广义互相关(Generalized Cross-Correlation, GCC)通过引入加权函数来减轻噪声带来的影响, 例如, PHAT加权通过相位归一化减轻信号幅度对时延估计的影响, 但它对所有频率分量一视同仁, 不能有效区分高信噪比的有效信号频带和噪声占主导的低信噪比频带。为此, 本文提出了一种噪声自适应加权函数(Noise-Robust Adaptive Weighting, NRAW), 对GCC-PHAT的输出进行二次频域加权, 具体的权重根据局部信噪比来动态调节, 高信噪比分量保持PHAT特性, 低信噪比分量则被大幅抑制。该方法无需预设的噪声模型, 仅利用信号起始段的静音假设, 就能针对性地抑制噪声频带。实验结果显示, 在SNR = -5 dB的强噪声条件下, NRAW方法的平均绝对误差较GCC的7.14°降低到4.57°。同时, 在5°误差容限内, 方位角估计准确率从50%提高到68%, 这充分验证了该自适应加权策略在抑制噪声干扰和提升定位鲁棒性上的显著效果。

关键词

声源定位, 广义互相关, 时延估计, 自适应加权

A GCC Time Delay Estimation Algorithm Based on Noise-Adaptive Weighting

Tan Gan*, Shaorao Kang, Lin Xiong

School of Science, Jiangxi University of Science and Technology, Ganzhou Jiangxi

Received: February 27, 2026; accepted: March 19, 2026; published: March 30, 2026

Abstract

In strong noise environments, the performance of traditional GCC time delay estimation methods

*通讯作者。

degrades sharply, primarily because noise interferes with the signal's cross-correlation function, leading to sidelobe elevation and mainlobe broadening. As a result, deviations occur in the time delay corresponding to the peak value. The Generalized Cross-Correlation (GCC) mitigates the impact of noise by introducing weighting functions; for instance, PHAT weighting reduces the influence of signal amplitude on time delay estimation through phase normalization. However, it treats all frequency components equally and fails to effectively distinguish between effective signal bands with high Signal-to-Noise Ratio (SNR) and noise-dominated bands with low SNR. To address this, this paper proposes a Noise-Robust Adaptive Weighting (NRAW), which applies secondary frequency-domain weighting to the GCC-PHAT output. The specific weights are dynamically adjusted based on local SNR: high-SNR components retain PHAT characteristics, while low-SNR components are significantly suppressed. This method requires no preset noise models and achieves targeted suppression of noise frequency bands solely by leveraging the silence assumption in the initial segment of the signal. Experimental results show that, under strong noise conditions with SNR = -5 dB, the NRAW method reduces the mean absolute error from 7.14° for GCC to 4.57°. Meanwhile, within a 5° error tolerance, the azimuth estimation accuracy improves from 50% to 68%. This fully validates the significant effectiveness of the proposed adaptive weighting strategy in suppressing noise interference and enhancing localization robustness.

Keywords

Sound Source Localization, Generalized Cross-Correlation (GCC), Time Delay Estimation, Adaptive Weighting

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

声源定位是利用麦克风阵列捕捉到的空间语音信号及阵列几何结构来确定声源的方位,在智能音箱、会议系统[1]、助听设备及人机交互[2]领域得到广泛应用。基于到达时间差(Time Difference of Arrival, TDOA)的定位方法通过估算声源信号到达不同麦克风的时间差,并结合阵列几何信息,求解出声源空间位置。该方法原理简单,计算高效,成为当前应用最为广泛的定位技术。在多种时延估计算法中,广义互相关(Generalized Cross-Correlation, GCC)通过加权函数改善互相关方法的不足,其中相位变换加权的GCC-PHAT最为经典。然而,在实际室内环境中,背景噪声不可避免地干扰声波传播,导致互相关函数主瓣展宽,甚至还会出现虚假峰值,造成时延估计偏差或失效。

为提高GCC类方法在复杂声学环境下的准确性与鲁棒性,现有的研究从多个方面提出改进方案,主要分为三类:自适应加权函数设计、子带/子空间信号分离以及前置自适应滤波。

1) 自适应加权函数方面,李典航等[3]在PHAT中引入基于Sigmoid函数的加权因子,提高了定位的精度;徐彬等[4]使用改进的PHAT函数对二次互相关函数进行加权;李保伟等[5]则提出基于广义互相关的SCOT/PHAT联合加权时延估计算法,具有更强的抗噪能力。这类方法通过优化加权形式,改善了相关函数的峰值特性。

2) 子带或子空间处理方面,陈海宏等[6]提出了一种频率滑动广义互相关和一种低秩近似替代方案来处理子带广义互相关矩阵,从而获得更精确的到达时间差估计。

3) 前置滤波增强方面,梁家碧等[7]结合最小均方差(LMS)自适应滤波与二次互相关,并使用样条插值提升峰值定位精度;程方晓等[8]采用LMS自适应滤波算法弥补广义互相关方法的不足,提高混响环境

下的时延估计精度；简泽明等[9]引入了 LMS 自适应滤波器作为前端处理以获得高质量的信号，提高了当前广义二次相关算法的精度和鲁棒性。

上述工作表明，自适应加权与前置滤波机制能够有效提升 GCC 在噪声环境下的鲁棒性。然而，部分方法仍存在明显局限：例如冯斌等[10]虽在广义二次互相关中引入噪声加权因子 α 以增强抗噪性，但是加权系数 α 依赖于离线标定的固定值，导致泛化性能受限。

针对这些问题，本文提出一种噪声自适应加权函数(Noise-Robust Adaptive Weighting, NRAW)，在 GCC-PHAT 基础上引入基于局部信噪比的动态频域权重，实现对高 SNR 有效频带与低 SNR 噪声主导频带的差异化处理。该方法仅需利用信号起始段估计噪声功率谱，无需先验噪声模型或复杂环境参数标定，具有较强的实用性和易部署性。

2. 方法

2.1. 互相关算法

在一个存在混响和噪声的室内环境中，麦克风阵列第 m 个麦克风接收到的语音信号 $x_m(t)$ 可以表示为式(1)所示：

$$x_m(t) = h_m(t, \theta) * s(t) + n_m(t) \quad (1)$$

其中， $s(t)$ 为干净的语音信号， $n_m(t)$ 为噪声信号， $h_m(t, \theta)$ 为房间脉冲响应。

对于两个麦克风接收到的信号 $x_1(t)$ 和 $x_2(t)$ ，两个信号的互相关过程可以表示为式(2)所示：

$$R_{x_1x_2}(\tau) = E[x_1(t)x_2(t)] \quad (2)$$

对信号 $x_1(t)$ 和 $x_2(t)$ 进行傅里叶变换得 $X_1(\omega)$ 和 $X_2(\omega)$ ，则互相关函数又可以进一步表示为式(3)所示：

$$R_{x_1x_2}(\tau) = \int_{-\infty}^{+\infty} X_1(\omega) X_2^*(\omega) e^{-j\omega\tau} d\omega \quad (3)$$

函数峰值处所对应的横坐标即为求出来的时延估计值。

2.2. 广义互相关算法

上一节所介绍的互相关方法在理想条件下能够准确估计时延，但在实际声学场景中，由于信号与噪声、噪声与噪声之间存在相关性，互相关函数的峰值会变得模糊不清，导致时延峰值对应的时延出现误判，最终会影响声源定位的精度。

为此，广义互相关时延估计方法通过在频域引入加权函数 $\Psi(\omega)$ ，对输入信号的不同频率分量进行增强或衰减，处理之后再行互相关运算，可以提升时延估计的性能。具体如式(4)所示：

$$R_{x_1x_2}(\tau) = \mathcal{F}^{-1}\{\Psi(\omega) \cdot X_1(\omega) X_2^*(\omega)\} \quad (4)$$

经典的几类加权函数如表 1 所示。

Table 1. Summary of classical weighting functions and their characteristics

表 1. 经典的加权函数以及特点总结

方法	加权函数	特点
GCC	1	无加权，对噪声敏感
PHAT	$1/ G_{x_1x_2} $	保留相位信息，将幅度归一化
ROTH	$1/G_{x_1x_1}$ 或 $1/G_{x_2x_2}$	对互功率谱应用倒数权重

续表

SCOT	$1/\sqrt{G_{x_1x_1} \cdot G_{x_2x_2}}$	功率相近时主瓣展宽
HB	$ G_{x_1x_2} /G_{x_1x_1} \cdot G_{x_2x_2}$	强干扰下方差增大、峰值模糊
WP	$ G_{x_1x_2} ^2/G_{x_1x_1} \cdot G_{x_2x_2}$	对宽带噪声抑制能力有限

虽然这几类加权函数各有优势,但是有一个共同的特性:它们均为固定权重,无法自适应环境变化。为此,本文提出了一种根据噪声自适应加权函数NRAW,可以根据实时信噪比或噪声特性动态调整权重,从而在不同噪声环境下自动优化互相关函数的峰值锐度。

2.3. 噪声自适应加权的广义互相关算法

两麦克风接收信号为 $x_1(t)$ 与 $x_2(t)$, NRAW 的实现流程如下:

1) 噪声功率谱估计

在噪声估计初始阶段,本文先假设信号起始段为纯噪声,其噪声功率谱 $P_n(\omega)$ 定义为:

$$P_n(\omega) = \frac{1}{2} \left(|X_{1,n}(\omega)|^2 + |X_{2,n}(\omega)|^2 \right) + \varepsilon \quad (5)$$

其中, $X_{i,n}(\omega)$ 为第 i 通道噪声信号进行傅里叶变换(FFT)后得到的, $\varepsilon = 1 \times 10^{-12}$ 为防止数值溢出的小常数。

2) 信号功率谱计算

对 $x_1(t)$ 与 $x_2(t)$ 信号进行 FFT, 得 $X_1(\omega)$ 和 $X_2(\omega)$, 信号平均功率谱为:

$$P_s(\omega) = \frac{1}{2} \left(|X_1(\omega)|^2 + |X_2(\omega)|^2 \right) + \varepsilon \quad (6)$$

3) 计算先验信噪比

在线性域定义先验信噪比为:

$$\xi(\omega) = P_s(\omega) / P_n(\omega) \quad (7)$$

将信噪比转化为分贝:

$$\text{SNR}(\omega) = 10 \log_{10} (\xi(\omega) + \varepsilon) \quad (8)$$

为避免极端 SNR 导致权重不稳定, 本文将其限制在区间 $[\text{SNR}_{\min}, \text{SNR}_{\max}]$, 取值为 $[-10, 20]$ dB。随后通过 S 型函数映射为权重:

$$W(\omega) = W_{\min} + \frac{W_{\max} - W_{\min}}{1 + \exp(-k(\text{SNR}(\omega) - \text{SNR}_{\text{mid}}))} \quad (9)$$

其中, $W_{\min} = 0.1$ 、 $W_{\max} = 0.95$ 为权重的上下限。 $\text{SNR}_{\text{mid}} = (\text{SNR}_{\min} + \text{SNR}_{\max}) / 2$, $k = 0.4$ 控制过渡区陡峭度。

原始互功率谱为 $G(\omega) = X_1(\omega) X_2^*(\omega)$ 。NRAW 的最终加权形式为:

$$\Psi_{\text{NRAW}}(\omega) = \frac{W(\omega)}{|G(\omega)| + \varepsilon} \quad (10)$$

互相关函数为:

$$R_{\text{NRAW}}(\tau) = \mathcal{F}^{-1} \{ \Psi_{\text{NRAW}}(\omega) \cdot G(\omega) \} \quad (11)$$

TDOA 估计值为:

$$\hat{\tau} = \arg \max_{\tau} |R_{\text{NRAW}}(\tau)| \quad (12)$$

计算出时延之后, 后续用 Chan 氏算法[11]进一步计算出声源的具体方位角。

2.4. 参数敏感性分析

本文所提 NRAW 加权函数包含四个关键参数: Sigmoid 陡峭度 k 、最小权重 W_{\min} 、最大权重 W_{\max} 及过渡中点 SNR_{mid} 。为验证各个参数鲁棒性, 在 $\text{SNR} = -5 \text{ dB}$ 条件下进行了敏感性实验。为了验证单一目标参数的敏感性, 固定其他参数, 仅变化目标参数进行实验。具体的参数范围基于预实验确定, 总共有 192 个参数的组合, 对所有的参数组合进行声源定位实验, 比较的指标为定位的平均绝对误差 MAE。实验得出参数的最佳值如表 2 所示。

Table 2. Parameter sensitivity analysis results

表 2. 参数敏感性实验结果

参数	取值范围	最佳值
k	[0.2, 0.4, 0.6, 0.8]	0.4
W_{\min}	[0.05, 0.1, 0.15]	0.1
W_{\max}	[0.85, 0.9, 0.95, 1]	0.95
SNR_{mid}	[-5, 0, 5, 10]	5

实验结果表明, $k = 0.4$, $W_{\min} = 0.1$, $W_{\max} = 0.95$, $\text{SNR}_{\text{mid}} = 5 \text{ dB}$, 定位 MAE 最低为 4.57° 。此时定位效果最好。

2.5. 与维纳滤波的理论关联

本文所提出的 NRAW 权重可看作是维纳滤波的非线性推广。设先验信噪比 $\xi(\omega) = P_s(\omega)/P_n(\omega)$, 维纳滤波权重为 $H_{\text{WF}} = \xi/(1+\xi)$ 。而 NRAW 通过 Sigmoid 映射实现, 如式(13)所示:

$$W(\omega) \approx \frac{\xi(\omega)^\beta}{1 + \xi(\omega)^\beta} \quad (13)$$

其中, $\beta = 10k/\ln 10 \approx 4.34k$ 。当 $k = 0.4$ 时, $\beta \approx 1.74 > 1$, 表明 NRAW 对低 SNR 频段的抑制强度高于标准维纳滤波, 同时通过 $W_{\min} > 0$ 可以避免完全衰减导致的相位失真, 更适合 TDOA 任务。

3. 实验与讨论

3.1. 数据构建

本文实验中所用的多通道麦克风阵列音频数据通过以下流程生成: 首先, 利用 gpuRIR 工具箱[12]仿真生成房间冲激响应(Room Impulse Responses, RIRs), 将 RIRs 与干净语音信号进行卷积, 得到纯净的多通道语音信号。

不含噪声的语音数据来源于 LibriSpeech 语料库[13], 从中随机选取 50 段的语音样本, 确保说话人、内容和语速的多样性。噪声信号则从 NOISEX-92 数据库[14]中随机选取, 以模拟真实场景中的多样化背景干扰。随后在时域将卷积后的多通道语音信号与噪声信号按指定信噪比进行叠加, 生成带噪麦克风阵列信号。房间尺寸设为 $6 \times 6 \times 3 \text{ m}^3$, 混响时间设置为 0.2 s, 即小混响情况, 以隔离混响因素对噪声的干

扰，聚焦于噪声对定位性能的影响。SNR 设置为 -5 dB 和 5 dB 两种情况。

在 $\text{SNR} = -5$ dB 时，归一化后的语音信号如下图 1 所示。

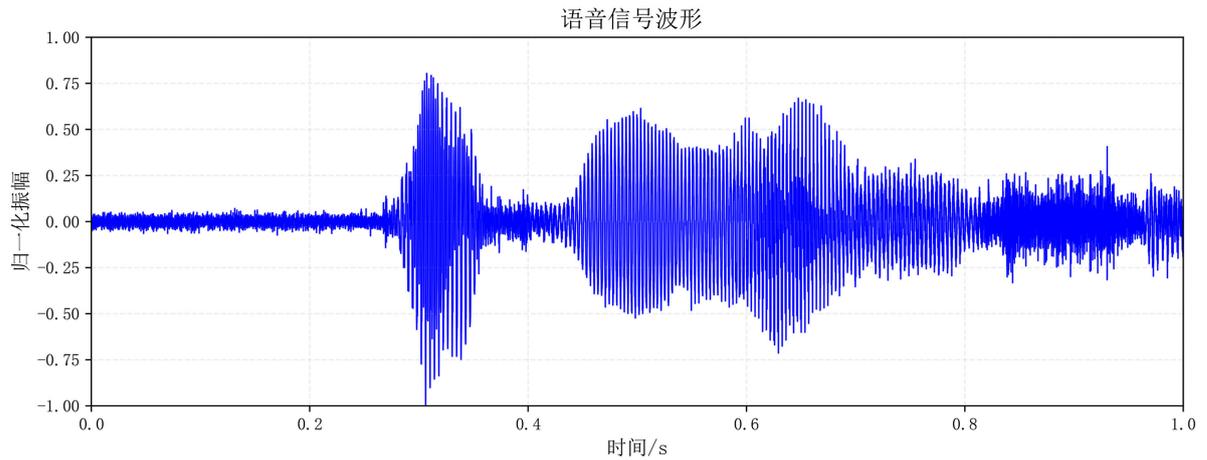


Figure 1. Normalized speech signal waveform

图 1. 归一化语音信号波形图

实验中信号采样率为 16 kHz，麦克风阵列为 4 元均匀线性阵列，阵元间距设置为 0.08 m，如图 2 所示。声源方位角在 $[0^\circ, 180^\circ]$ 范围内均匀随机分布，共生成 50 个测试样本， 50 个样本中，声源分布在预设房间的随机位置。

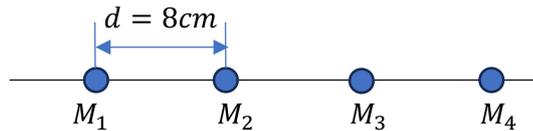


Figure 2. 4-element uniform linear microphone array

图 2. 4 元均匀线性麦克风阵列

3.2. 评价指标

为系统衡量不同加权算法在各类声学环境下的定位性能，本实验选取平均绝对误差(Mean Absolute Error, MAE)和定位准确率(Accuracy, ACC)作为核心评价指标。其中，MAE 定义为所有测试样本中估计方位角与真实方位角之间绝对差值的算术平均，如式(14)所示：

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (14)$$

其中， N 为测试样本总数， y_i 为第 i 个样本的估计方位角， \hat{y}_i 为第 i 个样本对应的真实方位角。MAE 数值越小表明定位精度越高。

ACC 表示在指定角度容差阈值内，方位角估计误差处于允许范围内的、包含活跃声源的帧所占的比例，从而反映模型准确定位活跃声源的成功率。ACC 以百分比形式表示，其计算公式如式(15)所示：

$$\text{ACC} = \frac{N_r}{N_c} \times 100\% \quad (15)$$

其中， N_r 表示满足角度容差条件的帧数量， N_c 表示数据集中总的帧数。

3.3. 结果与讨论

为系统评估所提噪声自适应加权方法 NRAW 在不同信噪比条件下的定位性能, 本文在混响时间 $T60=0.2$ s 的室内环境下, 分别于 $SNR = -5$ dB 与 $SNR = 5$ dB 两种不同的条件下对多种经典 GCC 加权函数进行对比实验。评价指标包括 MAE 与方位角估计误差在 5° 容差内的 ACC, 实验的结果如表 3 与表 4 所示。

Table 3. Comparative localization experiments with seven methods ($SNR = -5$ dB)

表 3. 七种方法定位对比实验($SNR = -5$ dB)

方法	MAE [$^\circ$]	ACC [%] (5°)
GCC	7.14	50
PHAT	4.75	66
ROTH	7.72	54
SCOT	4.75	66
HB	4.75	66
WP	7.14	50
NRAW	4.57	68

表 3 展示了在低信噪比环境($SNR = -5$ dB)下, 七种 GCC 类定位方法的性能对比。实验结果表明, 在强噪声干扰条件下, 传统方法如 GCC 与 WP 的 MAE 均为 7.14° , 定位误差明显下降; 而 PHAT、SCOT 和 HB 三者表现相近, MAE 均为 4.75° , 在 5° 的误差容限下, ACC 达 66%, 显示出对噪声较强的鲁棒性。本文所提出的 NRAW 在低信噪比场景下取得了最优的定位性能, 从具体的评价指标来看, MAE 为 4.57° , ACC 为 68%, 优于表格中所列的其他所有对比方法, 验证了其自适应加权机制在低信噪比场景下的有效性。

为了充分地对比七种加权函数的性能, 本文增加了一组对比试验, 在 $SNR = 5$ dB 时, 重新对比了七种方法的定位性能, 实验还是采用 MAE 和 ACC 作为评价的指标, 具体的实验结果如表 4 所示。

Table 4. Comparative localization experiments with seven methods ($SNR = 5$ dB)

表 4. 七种方法定位对比实验($SNR = 5$ dB)

方法	MAE [$^\circ$]	ACC [%] (5°)
GCC	6.68	50
PHAT	4.64	72
ROTH	4.92	68
SCOT	4.64	72
HB	4.64	72
WP	6.68	50
NRAW	4.46	74

从表 4 可以看出, 信噪比升高之后, 信号的质量提升, 表格中所列出的所有方法的定位性能均有不同程度的提升。其中, PHAT、SCOT 与 HB 定位性能相近, 在 5° 的误差容限下, 定位的 ACC 从 66% 提升至 72%, WP 方法因对宽带噪声抑制能力有限, MAE 还是较高, 定位的 ACC 为 50%。值得注意的是, 所提 NRAW 方法保持优势, 定位的 MAE 最小, 定位的 ACC 最高, 这进一步说明其在低信噪比场景和中低信噪比下定位性能较其他六种方法更佳。

为了更直观地看到各种时延估计方法的定位性能差异, 本文在 $\text{SNR} = -5 \text{ dB}$ 时, 进行了七种方法声源定位的可视化实验, 具体的结果如下图 3 所示。

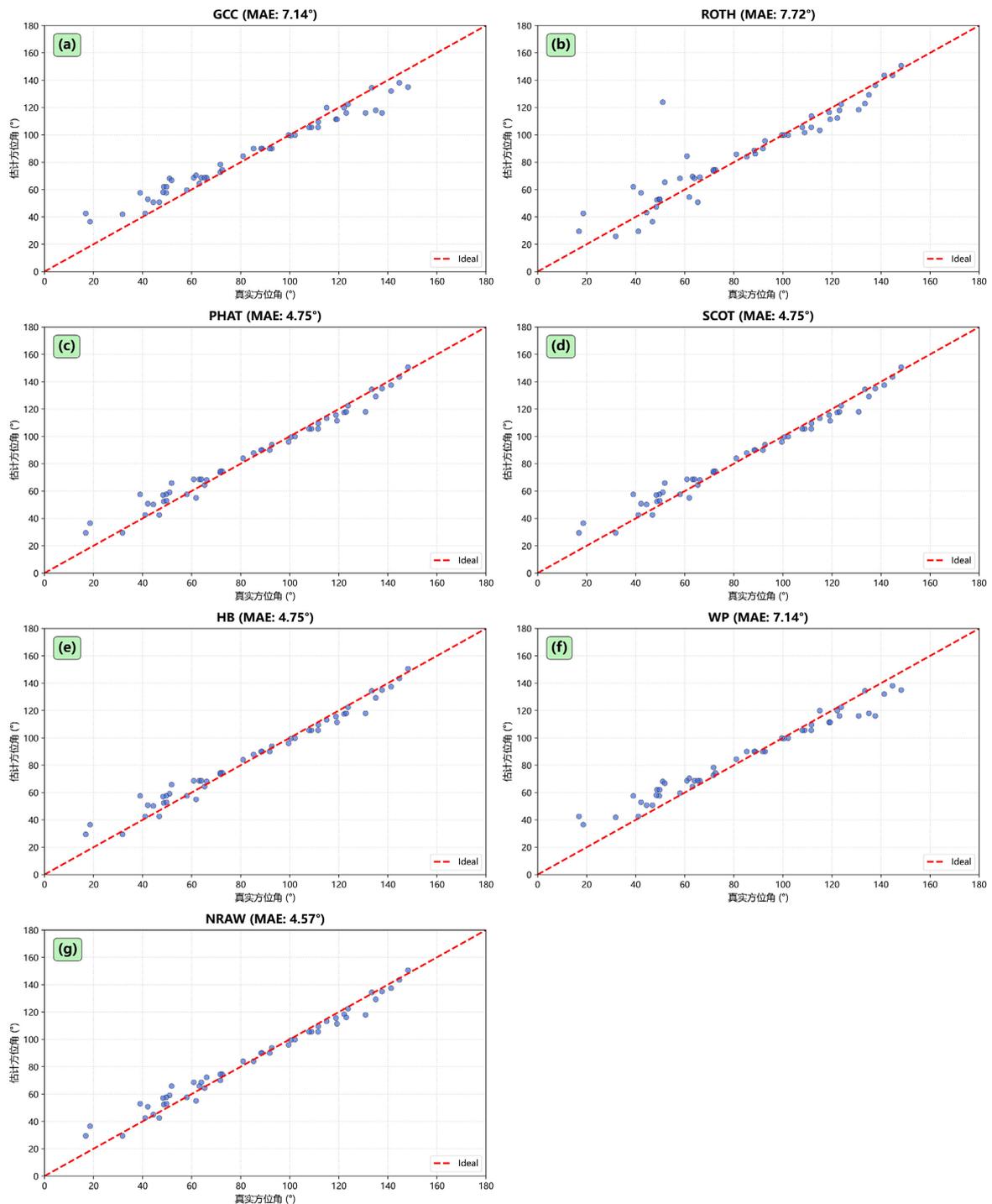


Figure 3. Comparison of scatter plots of DOA estimation performance for seven methods under $\text{SNR} = -5 \text{ dB}$

图 3. 七种方法在 $\text{SNR} = -5 \text{ dB}$ 下的 DOA 估计散点图对比

图 3 展示了所提方法 NRAW 与 GCC 以及其他五种经典加权函数的 DOA 估计散点分布。总共有七

个子图, 每一个子图对应一种方法, 图中已经做好了详细的标注。所有子图中, 横轴都为真实的方位角, 纵轴为算法估计出来的方位角。从图中可见, GCC 与 WP 方法的估计点分布相对离散, 尤其在中高角度区域存在明显偏差, ROTH 方法局部波动较大, 导致其 MAE 更高。相比之下, PHAT、SCOT 和 HB 三者表现相近, 数据点聚集在对角线附近, 反映其在噪声环境下仍能保持较好的一致性。所提出的 NRAW 方法在所有对比方法中表现最优, 其估计结果最为集中, 点云分布最接近理想直线, 说明该方法采用的自适应加权策略在低信噪比条件下有效提升了定位的稳定性和准确性。

4. 结语

针对低信噪比下声源定位性能退化的问题, 本文提出基于噪声自适应加权的时延估计算法, 通过引入局部信噪比驱动的频域动态加权机制, 有效缓解了 GCC-PHAT 在强噪声环境下因频带混叠导致的时延估计失真问题。实验验证表明, NRAW 不仅在 $\text{SNR} = -5 \text{ dB}$ 的严苛噪声场景中显著优于 GCC、GCC-PHAT 等基准方法, 且在 $\text{SNR} = 5 \text{ dB}$ 条件下仍保持性能优势, 展现出良好的环境适应性。

然而, 当前方法也存在不足之处。本方法仅针对噪声进行自适应加权, 在高混响环境下的鲁棒性有限。后续的工作将进一步探索深度学习的方法框架, 进一步提升算法在复杂真实声学环境中的适用性。

参考文献

- [1] Dai, X.M., Lou, W.Z., Liu, P., *et al.* (2014) Speaker Tracking Based on Microphone Cross Array in the Smart Conference System. 2014 *IEEE International Conference on Consumer Electronics—China*. Shenzhen, 9-13 April 2014, 1-4. <https://doi.org/10.1109/ICCE-China.2014.7029860>
- [2] 柯显信, 张文朕, 杨阳, 等. 仿人机器人多传感器定位系统[J]. 浙江大学学报(工学版), 2018, 52(7): 1247-1252.
- [3] 李典航, 王申营, 程骏, 等. 结合 Sigmoid 函数的自适应时延估计算法研究[J]. 自动化仪表, 2025, 46(4): 31-35+41.
- [4] 徐彬, 马鹏飞, 张恒远, 等. 声源定位系统的广义二次互相关算法改进与测试[J]. 现代电子技术, 2025, 48(13): 133-137.
- [5] 李保伟, 张兴敢. 基于广义互相关改进的麦克风阵列声源定位方法[J]. 南京大学学报(自然科学), 2020, 56(6): 917-922.
- [6] 陈海宏, 易永利, 韩钰, 等. 基于频率滑动广义互相关的电力噪声源定位算法[J]. 声学技术, 2024, 43(4): 550-556.
- [7] 梁家碧, 邵剑, 李群, 等. 一种用于声源定位的改进的广义互相关时延估计算法[J]. 光通信技术, 2024, 48(6): 28-33.
- [8] 程方晓, 刘璐, 姚清华, 等. 基于改进时延估计的声源定位算法[J]. 吉林大学学报(理学版), 2018, 56(3): 681-687.
- [9] 简泽明, 彭阳, 高泽平, 等. 基于改进二次相关算法的声源定位仿真研究[J]. 压电与声光, 2021, 43(2): 244-247, 293.
- [10] 冯斌, 赵军峰, 郭强. 自适应广义互相关的声阵列炸点定位[J]. 探测与控制学报, 2023, 45(5): 60-65.
- [11] Li, Z.P. and Li, Z.C. (2025) Reference Coordinate Based Chan Algorithm for UWB Personnel Localization in Underground Coal Mines. *Scientific Reports*, **15**, Article No. 17922. <https://doi.org/10.1038/s41598-025-03007-6>
- [12] Diaz-Guerra, D., Miguel, A. and Beltran, J.R. (2021) gpuRIR: A Python Library for Room Impulse Response Simulation with GPU Acceleration. *Multimedia Tools and Applications*, **80**, 11757-11775. <https://doi.org/10.1007/s11042-020-09905-3>
- [13] Panayotov, V., Chen, G., Povey, D., *et al.* (2015) Librispeech: An ASR Corpus Based on Public Domain Audio Books. 2015 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, 19-24 April 2015, 5206-5210. <https://doi.org/10.1109/ICASSP.2015.7178964>
- [14] Varga, A. and Steenenkel, H.J.M. (1993) 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 Communication*, **12**, 247-251. [https://doi.org/10.1016/0167-6393\(93\)90095-3](https://doi.org/10.1016/0167-6393(93)90095-3)