# 使用三麦克风阵列精确实时估计二维到达方向

1st AntonKovalyov德克萨斯大学达拉斯分校电气与计算机工程美国理查森, anton.kovalyov@utdallas.edu

2nd KashyapPatel德克萨斯大学达拉斯分校电气与计算机工程美国理查森, patelkashyap@utdallas.edu

3rd IssaPanahi德克萨斯大学达拉斯分校电气与计算机工程美国理查森，i mp015000@utdallas.edu

arXiv:2305.05630v1 [eess.AS] 9 May 2023

摘要——**本文提出了一种使用非线性三麦克风阵列实时估计一个或多个声源二维到达方向（2D-DOA）的方法。该方法利用帧级到达时间差（TDOA）测量值估计二维到达方向。与传统方法通过理论模型根据到达时间差推断位置参数不同，我们提出了一种基于监督学习、更为实用的方法。所提出的模型采用最近邻搜索（NNS），并应用于球面斐波那契晶格；该晶格由在实际环境中学习得到的到达时间差到二维到达方向的映射构成。此外，本文还介绍了滤波和聚类后处理器，以提高声源检测和定位鲁棒性。**

索引词—**到达方向、到达时间差、阵列、最近邻、聚类**

## I. 引言

基于到达时间差（TDOA）的声源定位（SSL）是文献中广泛采用的一种方法。当声源位于阵列的远场，或麦克风数量少于四个时，基于 TDOA 的三维（3D）声源定位无法实现，因此会改为估计到达方向（DOA）。实际应用包括波束形成[1], 盲源分离 [2], 以及为存在空间听力损失（SHL）的人提供可视化的到达方向指示器 [3]–[5]。在这项工作中，我们尤其关注后者。

Anshuman等人 [3] 提出了一款智能手机应用程序，为患有重度听力损失的用户提供语音源的可视化方位角到达方向指示器。使用智能手机实现这一目的尤其便利，因为智能手机的普及率很高。然而，到达方向仅使用两个麦克风进行估计，从而产生所谓的前后 模糊，这是线性阵列中常见的问题。如今，许多智能手机都配备了至少三个麦克风组成的阵列。当阵列为非线性时，我们不仅可以避免前后模糊，还能够同时估计声源的方位角和仰角，这称为二维到达方向估计。Tokg¨oz等人 [4], [5]提出了对 [3] 中工作的改进方案，适用于由三个麦克风组成的$L$形阵列。然而，这些方法仅限于检测和定位单个语音源。此外，对于实测到达时间差与理论到达时间差之间的偏差，尚未提出实际可行的补偿方案；这种偏差在实际系统中很常见，其原因包括阵列校准错误以及麦克风相位响应的变化

本研究得到美国国立卫生研究院（NIH）下属国家耳聋与其他沟通障碍研究所（NIDCD）的资助，资助编号为5R01DC015430-05。本文内容仅由作者负责，并不一定代表 NIH 的官方观点。

对于测量得到的到达时间差与理论到达时间差之间的差异，这种差异在实际系统中很常见，原因包括阵列校准错误以及麦克风相位响应各异。

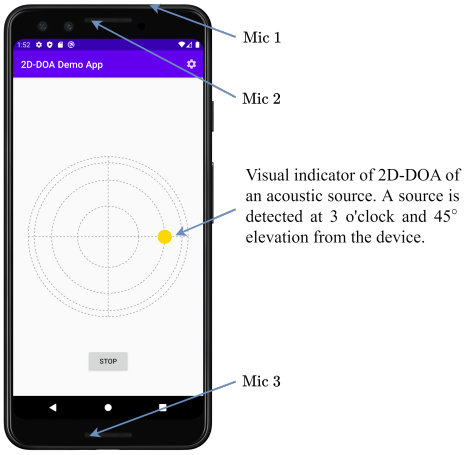



图1. 将所提出的方法部署在智能手机应用程序上，为患有重度听力损失的人提供可视化的二维到达方向指示器（声源位置在设备上方半球上的正投影）。

受上述观察结果启发，我们提出了一种使用非线性三麦克风阵列对多个声源进行实时二维到达方向估计的实用方法。该方法不采用理论模型，而是在实际环境中学习TDOA到二维到达方向映射，并通过对包含所学习映射的球面斐波那契晶格执行最近邻搜索（NNS）来进行推理。此外，我们还引入了用于可靠检测和定位一个或多个声源的滤波与聚类后处理模块。如图1所示，所提出的方法已在智能手机上实现。

## II. 问题表述

考虑一个处于嘈杂混响环境中的三麦克风阵列，其中有 $C$ 个声源。令$\mathbf{m}_i$ 和$\mathbf{s}_c$ 分别表示第 $i$个麦克风和第 $c$个声源的三维位置，其中 $i \in \{1,2,3\}$ 且 $c \in \{1,2,\dots,C\}$。第 $i$个麦克风采集到的信号建模为

$$\mathbf{y}_i = \mathbf{v}_i + \sum_{c=1}^{C} \mathbf{x}_{ic} \,, \tag{1}$$

其中$\mathbf{v}_i$ 是非相干噪声，$\mathbf{x}_{ic}$ 是第 $c$个声源的混响信号。假设角度间隔足够大，该问题被表述为对 $C$ 个声源进行实时检测，并估计它们相对于麦克风阵列的方位角和仰角。

## III. 方法

让我们将所有$\mathbf{y}_i$ 分割为长度为$L$的 $K$ 个重叠帧。所提出的方法以因果方式在帧级别运行。处理流程由四个阶段组成：(1) 到达时间差估计；(2) 将测量得到的到达时间差映射到二维到达方向；(3) 滤除不可靠的估计值；以及 (4) 聚类。

### *A.* 到达时间差估计

令$\mathbf{y}_{ik}$为$\mathbf{y}_i$的第 $k$个片段，其中 $k = 1, 2, \ldots, K$。令 $V$ 为这样的帧索引子集：在每个时刻恰好存在一个主导声源。假设 $V$ 非空，并且包含混合信号中每个声源所对应的帧索引；只要片段长度不过长，这些假设就是合理的，尤其是对于语音等稀疏信号而言。令$\mathbf{s}(v)$表示帧索引为 $v \in V$时主导声源的三维位置。当直达路径信号源自$\mathbf{s}(v)$，并在$\mathbf{m}_i$ 与$\mathbf{m}_j$之间接收时，其以米为单位的到达时间差，对于满足 $j \neq i$的 $j \in \{1, 2, 3\}$ ，可表示为

$$r_{ij}(v) = ||\mathbf{s}(v) - \mathbf{m}_i|| - ||\mathbf{s}(v) - \mathbf{m}_j|| \,. \tag{2}$$

假设已知信号传播速度，可以通过$\mathbf{y}_{iv}$ 与$\mathbf{y}_{jv}$ [6]之间某个加权互相关函数的峰值，求得 $r_{ij}(v)$的估计值。这里我们采用修正互功率谱相位 [7] (mCPSP)，并结合二次插值 [8] (QI)，以提高估计分辨率。每隔 $k$帧估计一次到达时间差。不可靠的估计值会在滤波阶段被剔除。

### *B.* 将到达时间差准确映射为二维到达方向

根据上一节的公式，我们考虑单个声源的帧级定位。因此，为简洁起见，除非出现歧义，否则省略声源和帧索引的标记。令$\mathbf{s}$、 $r_{ij}$和$\hat{r}_{ij}$ 分别表示三维声源位置、以米为单位的真实到达时间差及其相应的估计值。为更好地理解问题几何结构，我们先考虑一种将到达时间差映射到二维到达方向的简单闭式解（CF）。为减少自由度，将$\mathbf{m}_1$固定在原点，即$[0, 0, 0]^T$；将$\mathbf{m}_2$固定在 $[b, 0, 0]^T$；并将$\mathbf{m}_3$固定在 $[c_x, c_y, 0]^T$。令 $r$表示声源到原点的距离。当 $r$较大，即声源处于远场时，其实际值对到达时间差的影响可以忽略。因此，我们通过将 $r$ 固定为某个较大的值，进一步减少一个自由度。图2展示了问题几何结构。感兴趣的参数是 $\theta$ 和 $\phi$，分别表示声源的方位角和仰角。为简洁起见，我们重新参数化该问题：在以原点为中心、远场半径为 $r$的球面上求声源的三维位置$\mathbf{s}$ ，其中 $\theta$ 和 $\phi$ 可根据图2中的关系由$\mathbf{s}$推得。令 $s_x$、 $s_y$和 $s_z$ 表示相应的$\mathbf{s}$坐标。将线性无关的到达时间差 $r_{12}$ 和 $r_{13}$ 映射到$\mathbf{s}$的闭式解为

$$\begin{aligned} s_x &= \frac{b^2 + 2r_{12}r - r_{12}^2}{2b} \\ s_y &= \frac{c_x^2 + c_y^2 - r_{13}^2 + 2r_{13}r - 2c_x s_x}{2c_y} \\ s_z &= \pm\sqrt{r^2 - s_x^2 - s_y^2} \,. \end{aligned} \tag{3}$$

我们注意到， $b$ 和 $c_y$ 不能为0，这意味着需要采用非线性阵列来估计二维到达方向。我们还注意到，$s_z$ 可以为负，也可以为正。在本文中，我们令 $s_z$ 为正。这样，这个问题便可视为对阵列上方半球中的声源进行定位，这等价于令 $\theta \in [-\pi, \pi]$ 和 $\phi \in [0, \pi/2]$。

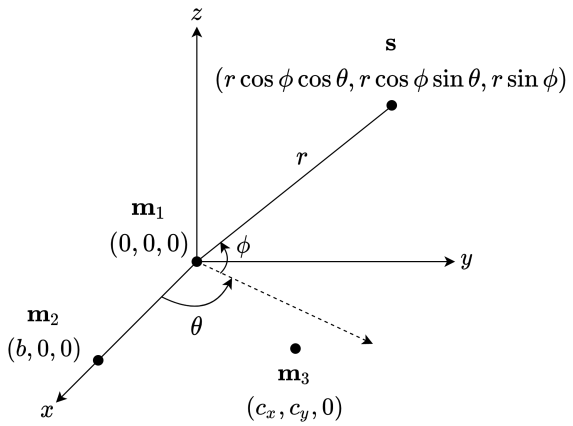



图2. 使用三个麦克风进行二维到达方向估计的问题几何结构。

如果测得的到达时间差因混响和噪声以外的原因而不符合理论值，例如阵列校准错误以及麦克风相位响应变化，则 CFin (3) 可能不够准确。为提高性能，可以直接在现场学习 TDOA到二维到达方向映射。因此，我们提出以下监督学习方法。将搜索空间离散化为

$$S = \{\boldsymbol{\gamma}^{(1)}, \boldsymbol{\gamma}^{(2)}, \ldots, \boldsymbol{\gamma}^{(N)}\} \,, \tag{4}$$

其中， $\boldsymbol{\gamma}^{(n)} = [\theta^{(n)}, \phi^{(n)}]^T$ 对于 $n = 1, 2、\ldots\ldots、N$，将搜索空间中的第 $n$ 个二维到达方向元组分组。同样，令

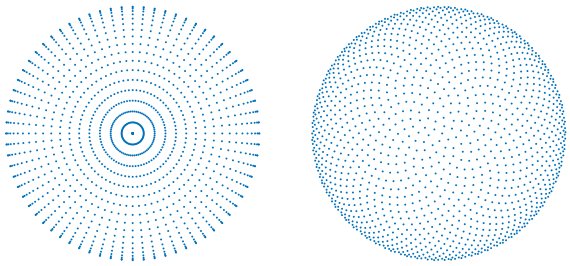
$$Q = \{\mathbf{q}^{(1)}, \mathbf{q}^{(2)}, \ldots, \mathbf{q}^{(N)}\} \,, \tag{5}$$

其中 $\mathbf{q}^{(n)} = [r_{12}^{(n)}, r_{13}^{(n)}, r_{23}^{(n)}]^T$ 归类对应的到达时间差映射。为增强稳健性，考虑所有组合。 $Q$ 中的到达时间差映射以监督方式离线收集。在推理过程中，NNS 的应用如下

$$\hat{\boldsymbol{\gamma}} = \underset{\boldsymbol{\gamma}^{(n)}}{\arg\min} ||\mathbf{q}^{(n)} - \hat{\mathbf{q}}||^2 \,, \tag{6}$$

其中 $\hat{}\mathbf{q} = [\hat{}r_{12}, \hat{}r_{13}, \hat{}r_{23}]^T$。映射存储在 $k$ 维树结构中 [9], [10], ，使得 NNS 的期望时间复杂度为对数级。

为了进一步提高效率， $S$ 中的候选解应尽可能均匀地分布在整个半球上。



(a) (b)

图 3. 在半球面上放置 $N = 1297$ 个点的两种不同方法。经纬度格点 (a) 和斐波那契格点 (b)。以极点为中心的正投影。

一种简单的方法是按某个角间距 $\delta = 180°/u$ 对 $\theta$ 和 $\phi$ 进行离散化，其中 $u$ 为正整数，由此得到一种称为经纬度格点[11]的点分布。这种格点可以视作一组点，它们位于经线与纬线构成的网格交点上（图 3a）。点的总数为 $N = u^2 + 1$，即经线的数量（$2u$）乘以纬线的数量（$u/2$），再加上一个极点。然而，在这种格点中，点会集中在极点附近，导致明显的各向异性。为此，我们采用球面斐波那契点集算法 [12]，以获得更均匀的点分布（见图 3b）。这种格点通过以下方式生成

$$\begin{aligned} \theta^{(n)} &= \frac{2\pi(n-1)}{\Phi} \\ \phi^{(n)} &= \frac{\pi}{2} - \cos^{-1}\left(1 - \frac{2n-1}{2N}\right), \end{aligned} \tag{7}$$

其中，$\Phi = (1+\sqrt{5})/2$ 为黄金比例。

### C. 滤波

我们提出了一种简单的三步滤波器，用于拒绝不可靠的到达时间差测量集合，即 (6) 中的 $\hat{\mathbf{q}}$。滤波的目的是滤除在帧索引 $k \notin V$ 处计算出的所有测量值。每个连续步骤都包括验证测量值是否满足特定的可靠性条件。前两个条件仅分别应用于各个到达时间差。如果任何估计值不满足给定条件，则拒绝整个集合。第一个条件通过确保峰值的互相关值高于正阈值 $T_R$，来验证估计延迟处存在声学活动，具体如下

$$R(\ell_{\max}) > T_R\,, \tag{8}$$

其中，$R(\ell)$ 是由修正互功率谱相位在时延索引 $\ell$ 处给出的加权互相关函数，$\ell_{\max}$ 是峰值所在的时延。在第二步中，我们通过计算置信度 $\beta \in [0,1]$，来量化互相关峰值相较于其他候选解的显著程度，其定义为

$$\begin{aligned} \beta &= 1 - \frac{\eta}{R(\ell_{\max})} \\ \eta &= \frac{1}{|\mathcal{L}|}\sum_{\ell\in\mathcal{L}} \max\{0, R(\ell)\}\,, \end{aligned} \tag{9}$$

其中，$\mathcal{L}$ 是除 $\ell_{\max}$ 之外所有合理时延的集合。接着，我们验证

$$\beta > T_\beta\,, \tag{10}$$

其中，$T_\beta$ 为某个阈值。最后，第三步通过验证 (6) 中 NNS 估计器的误差低于阈值 $T_q$，确保 $\hat{\mathbf{q}}$ 中的到达时间差之间具有一致性，具体如下

$$||\mathbf{q}(\hat{\boldsymbol{\gamma}}) - \hat{\mathbf{q}}||^2 < T_q\,, \tag{11}$$

其中，NNS 找到的与 $\hat{\mathbf{q}}$ 最接近的匹配项是$\mathbf{q}(\hat{\boldsymbol{\gamma}})$。

### D. 聚类

聚类用于将帧级二维到达方向估计值分配给正确的声源，并以一种能够提升声源检测和定位可靠性的方式合并累积测量值。为此，我们提出了时效与频次感知指数滤波聚类（RFEFC）。RFEFC 的设计与实时指数滤波聚类（RTEFC） [13]中使用的指数滤波概念密切相关。在 RTEFC 中，指数滤波用于更新与给定位置测量值距离最小的簇。除了支持实时帧级处理、内存和计算开销低之外，RTEFC 还具有良好的跟踪能力。然而，RTEFC 未考虑数据的时效性和频次，而这两项特征在本研究中十分重要。

在 RFEFC 中，实时维护固定数量的簇 $N_c$。每个第 $i$ 个簇（$i \in \{1, 2, \ldots, N_c\}$）由一个簇质心 $\hat{\mathbf{s}}_i$ 和一个簇置信度 $\rho_i$构成。这里，$\hat{\mathbf{s}}_i$ 表示单位半径半球上第 $i$ 个三维声源位置的估计值。为进行平均，必须采用笛卡尔表示。另一方面，$\rho_i$是一个介于 0 和 1 之间的数值，用于量化对估计值 $\hat{\mathbf{s}}_i$的置信度，其中 1 表示置信度高。初始时，对于所有 $i$，$\hat{\mathbf{s}}_i = \varnothing$ 和$\rho_i = 0$，这意味着簇处于非活动状态，即未检测到声源。令 $\hat{\mathbf{s}}$ 表示每隔 $\Delta t$ 秒采样一次的某个声源位置的帧级估计值。如果该估计值在滤波过程中被拒绝，则令 $\hat{\mathbf{s}} = \varnothing$。RFEFC 会搜索置信度最高且质心与 $\hat{\mathbf{s}}$之间的距离小于某个最小距离 $d_{\min}$ 的簇。如果找到这样的簇，则使用指数滤波更新其质心 $\hat{\mathbf{s}}_{\text{close}}$，并按如下方式提高其置信度 $\rho_{\text{close}}$

$$\begin{aligned} \hat{\mathbf{s}}_{\text{close}} &= \alpha\hat{\mathbf{s}}_{\text{close}} + (1-\alpha)\hat{\mathbf{s}} \\ \rho_{\text{close}} &= \min\left\{1, \rho_{\text{close}} + N_s^{-1}\right\}, \end{aligned} \tag{12}$$

其中，$N_s$ 是一个正整数，表示使 $\rho_{\text{close}}$ 达到 1 所需的连续帧数，从而记录集群的更新频率。如果找不到质心与 $\hat{\mathbf{s}}$ 距离在 $d_{\min}$ 以内的集群，RFEFC 会选择置信度最低的集群，并将其更新为

$$\begin{aligned} \hat{\mathbf{s}}_{\text{old}} &= \hat{\mathbf{s}} \\ \rho_{\text{old}} &= N_s^{-1}\,. \end{aligned} \tag{13}$$

此更新可以解释为从过时数据创建一个新簇。最后，为了体现新近性，每隔 $i$ 个久久久久久

对于未更新的簇，RFEFC 会降低置信度阈值，然后遗忘置信度降至 0 的簇。操作顺序如下：

$$\rho_i = \max\left\{0, \rho_i - \frac{\Delta t}{T_{\text{win}}}\right\}$$
$$\hat{\mathbf{s}}_i = \begin{cases} \hat{\mathbf{s}}_i & \text{if } \rho_i > 0 \\ \varnothing & \text{if } \rho_i = 0\,, \end{cases} \tag{14}$$

其中，$T_{\text{win}}$ 是一个正实数，用于控制 RFEFC 忘记不活跃簇所需的时间（以秒为单位）。

在 RFEFC 中引入 $\rho_i$，使其能够通过一种简单的机制判断估计位置上是否存在声源。当 $\rho_i$ 达到 1 并持续高于某个阈值 $T_a$时，即检测到声源 $i$。RFEFC 的内存和计算复杂度均与 $N_c$呈线性关系，因此其效率与 RTEFC 相当。此外，由于采用指数滤波，原则上 RFEFC 能够跟踪持续发声的移动声源。最后，与更新最接近位置估计值的簇的 RTEFC 不同，RFEFC 会更新指定距离范围内置信度最高的簇。此选择规则旨在尽量减少因估计值抖动而导致多个相距较近的簇长时间处于活跃状态的可能性。在这种情况下，RFEFC 只会更新置信度最高的簇，从而使附近其他簇逐渐被遗忘。

## IV. 实验

我们在实地进行了三项实验，使用图 1 所示的 Pixel 3 智能手机的非线性三麦克风阵列评估所提出方法的性能。实验在一间大小适中的办公室内进行，环境中存在典型的办公室噪声。

为填充式（5）中的 $Q$，数据采集按如下方式进行：将智能手机放置在转台上，并将扬声器放在固定表面上，距离转台 0.6 米；扬声器分别位于十个不同仰角处，仰角均匀分布在 $0^\circ$ 至 $90^\circ$ 之间。数据采集期间，扬声器播放白噪声，同时转台以固定速度旋转，直至完成一整圈，从而覆盖整个方位角范围。白噪声被用作声源信号，因为它具有良好的自相关特性。测量到达时间差，并将其映射到经纬度格点上的已知点。通过插值生成式（7）中的斐波那契格点。

参数定义如下：采样率$f_s$ 为 48 kHz，使用帧长为 $L = 1024$ 、重叠率为 50% 的帧。第 III-B 节中的格点数 $N$ 设为 $10^4$。第 III-C 节中的阈值 $T_R$、 $T_\beta$和 $T_q$ 分别设为 1e-2、0.5 和 5e-5。参数

第 III-D 节中的 $N_c$、 $\Delta t$、 $d_{\min}$、 $N_s$、 $\alpha$、 $T_{\text{win}}$和 $T_a$ 分别设为 10、 $L/(2f_s)$ 秒、0.25、5、0.75、5 秒和 0.5。

在第一个实验中，我们使用采集到的到达时间差，评估监督学习相较于更简单的 CF 在将到达时间差映射到二维到达方向方面的实际效果；后者是 [3]–[5] 中所用的映射方案类型。为此，我们将式 (6) 中 NNS 的二维到达方向估计性能与式 (3) 中对应的 CF 进行基准比较。

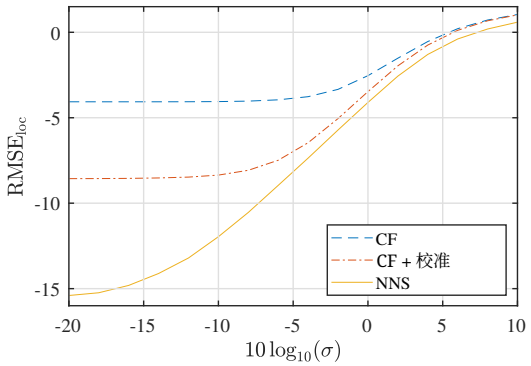



图 4. 实验 1. 不同到达时间差测量噪声下，NNS 与 CF 的二维到达方向估计性能。

采用定位均方根误差（RMSE$_{\text{loc}}$）作为性能指标。RMSE$_{\text{loc}}$ 衡量在单位半球上进行 $M$ 次试验得到的定位误差，定义如下：

$$\text{RMSE}_{\text{loc}} = \sqrt{\frac{1}{M}\sum_{i=1}^{M} ||\mathbf{s}^{(i)} - \hat{\mathbf{s}}^{(i)}||^2}\,, \tag{15}$$

其中，$\mathbf{s}^{(i)}$ 表示在第 $i$次试验中生成的单位半球上的声源位置，$\hat{\mathbf{s}}^{(i)}$ 表示相应的估计值。为生成$\mathbf{s}^{(i)}$ 以供评估，我们分别从各自范围内独立、均匀地随机抽取方位角和仰角，并将其映射到笛卡尔坐标。利用现场采集的数据集对相应的到达时间差进行插值。最后，为模拟测量噪声，在插值后的到达时间差中加入标准差为 $\sigma \in [0.01, 10]$ 厘米的加性高斯白噪声（AWGN）。由于 CF 需要准确掌握麦克风阵列几何结构，我们考虑了两种变体。第一种简称 CF，其阵列几何结构通过手动测量获得。第二种变体（CF + 校准）使用为 NNS 采集的数据集对阵列进行校准，并采用 Levenberg–Marquardt算法作为优化算法。图 4 给出了结果。可以看到，未经校准的 CF 性能较差，这归因于其对不精确的麦克风位置估计值较为敏感。尽管引入校准显著改善了结果，但该方案仍无法达到 NNS 所展现的高性能，尤其是在低噪声和中等噪声条件下。因此，NNS 可用于提高先前工作在 [3]–[5]中实现的定位分辨率。

在第二项实验中，我们测试所提出的方法检测并定位两个重叠声源的能力。一个声源是播放音乐的扬声器，放置在距离两米处，方位角为 $55^\circ$，仰角为 $0^\circ$。另一个声源是正在说话的人类说话者，站在距离半米处，方位角为 $145^\circ$，仰角为 $40^\circ$。结果如图5所示。帧级定位估计大多准确，但在 100°处出现了一个“幽灵”声源，我们将其归因于一帧中存在多个主导声源导致互相关函数平滑所致

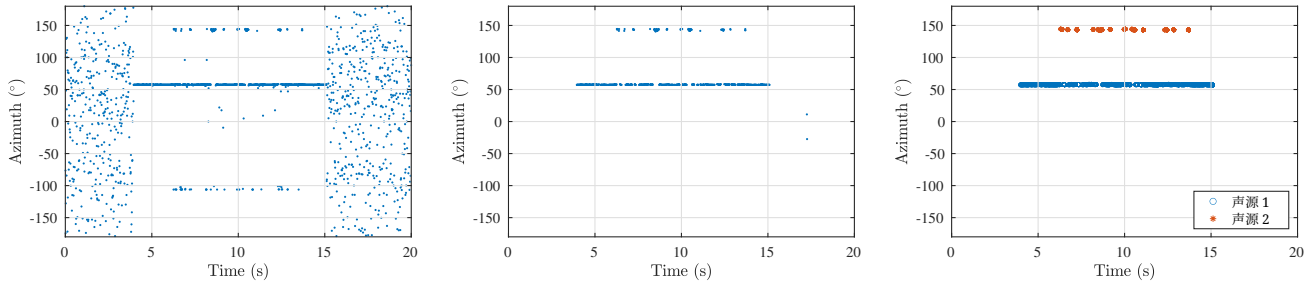



图 5. 实验 2。两个重叠声源的检测与定位处理阶段。为简洁起见，省略仰角。音乐和语音分别在 4 秒和 6 秒时开始，在 14 秒和 15 秒时结束。(a) 原始 DOA 估计值。(b) 滤波后的估计值。(c) 滤波并聚类后的估计值。

。不过，滤波会去除大多数离群值。最后，RFEFC检测到这两个声源，并由于具有 $N_s > 1$，通过剔除剩余离群值增加了一层滤波。我们注意到，与[3]–[5], 不同，这里没有使用语音活动检测器，因此可以检测和定位语音以外的声源。

与 RTEFC 一样，在 RFEFC 中使用指数滤波表明，原则上，所提出的方法能够跟踪产生连续声音的移动声源。为验证这一说法，在第三个也是最后一个实验中，我们测试所提出的方法跟踪两个重叠声源的能力。我们考虑与第二个实验相同的双声源定位场景，不同之处在于智能手机放置在旋转转台上，并在 20 秒内完成一整圈旋转。因此，我们预计声源位置会在半球上形成两个同心圆，每个圆分别由相应声源的仰角决定。图 6 中的结果表明，所提出的方法确实具有出色的跟踪能力。

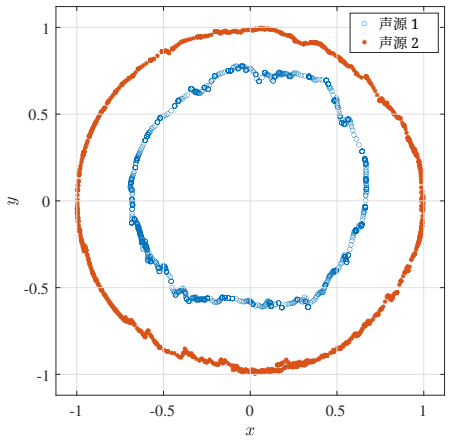



图 6. 实验 3。单位半径半球极点为中心的正投影上的聚类位置估计值。

## V. 结论

提出了一种使用非线性三麦克风阵列进行精确二维到达方向估计的方法。该问题被建模为在阵列上方单位半径半球内对一个或多个声源进行定位。为了获得最佳实际效果，将推导出的 CF 替换为应用于球面斐波那契晶格的 NNS，该晶格包含在实际环境中学习得到的 TDOA到二维到达方向映射。此外，还引入了滤波和聚类后处理器，以剔除不可靠的测量结果，并实现对多个声源更稳健的检测和定位。在实际环境中评估时，所提出的方法展现出出色的二维到达方向估计精度，以及对两个重叠声源的跟踪能力。

## 参考文献

[1] Z.-Q. Wang 和 D. Wang，“全神经网络多通道语音增强”。载于 *Interspeech*，2018 年，第 3234–3238 页。[2] G. S. Bhat、C. K. A. Reddy、N. Shankar 和 I. Panahi，“一种计算效率高的盲源分离方法，用于助听器应用并在智能手机上实时实现”，载于 *Proceedings of Meetings on Acoustics 178ASA*，第 39 卷，第 1 期。美国声学学会，2019 年，第 055008 页。[3] A. Ganguly、A. K¨uç¨uk 和 I. Panahi，“用于提升助听设备空间感知能力的实时智能手机应用程序”，载于 *201840th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*。IEEE，2018 年，第 433–436 页。[4] S. Tokg¨oz、A. Kovalyov 和 I. Panahi，“使用三个麦克风实时估计语音源的到达方向”，载于 *2020 IEEE Workshop on Signal Processing Systems (SiPS)*，2020 年，第 1–5 页。[5] S. Tokgoz 和 I. M. S. Panahi，“使用随机奇异值分解的稳健三麦克风语音源定位”，*IEEE Access*，第 9 卷，第 157 800–157 811 页，2021 年。[6] C. Knapp 和 G. Carter，“用于时延估计的广义相关法”，*IEEE transactions on acoustics, speech, and signal processing*，第 24 卷，第 4 期，第 320–327 页，1976 年。[7] D. V. Rabinkin、R. J. Renomeron、A. Dahl、J. C. French、J. L. Flanagan 和 M. Bianchi，“使用麦克风阵列进行声源位置定位的 DSP 实现”，载于 *Advanced signal processing algorithms, architectures, and i mplementations VI*，第 2846 卷。国际光学与光子学学会，1996 年，第 88–99 页。[8] I. J. Tashev，声音采集与处理：实用方法。John Wiley & Sons，2009 年，第 300–301 页。[9] J. L. Bentley，“用于关联搜索的多维二叉搜索树”，*Communications of the ACM*，第 18 卷，517，1975 年 9 月，509no. 第 – 页。[10] J. H. Friedman、J. L. Bentley 和 R. A. Finkel，“一种以对数期望时间寻找最佳匹配的算法”，*ACM Transactions on Mathematical Software (TOMS)*，第 3 卷，第 3 期，第 209–226 页，1977 年。[11] ´A. Gonz´alez，“使用斐波那契网格和经纬度网格测量球面面积”，*Mathematical Geosciences*，第 42 卷，第 1 期，第 49–64 页，2010 年。[12] R. Marques、C. Bouville、M. Ribardi`ere、L. P. Santos 和 K. Bouatouch，“用于光照积分的球面斐波那契点集”，载于 *Computer Graphics Forum*，第 32 卷，第 8 期。Wiley Online Library，2013 年，第 134–143 页。[13] G. Stanley，“大数据近似控制（BDAC）——一种基于模式匹配与近似的无模型估计和控制新范式”，*Journalof Process Control*，第 67 卷，第 141–159 页，2018 年。