doi: 10.11823/j.issn.1674-5795.2018.01.09

强混响条件下基于广义互相关的声源定位算法改进

蒋相斌, 白宗龙, 孙金玮, 黄博妍

(哈尔滨工业大学,黑龙江 哈尔滨 150001)

摘 要:针对特定声场环境,提出了一种强混响条件下基于广义互相关的声源定位算法的改进措施,即利用 混响成分中包含的声源位置信息,提高声源定位的正确率。混响与噪声是影响声源定位的两个主要因素。目前声 源定位的主要方法都是以降低信号中混响与噪声成分为目标的,如波束成形、空间谱估计以及到达时间差法等。 其中,基于广义互相关的声源定位算法属于到达时间法的常用方法,具有实时性强与高鲁棒性的优势。本文采用 镜像源模型对室内环境进行建模,分析声源到双麦克风的传播路径,利用混响中存在的声源位置信息剔除错误定 位结果,从而提高定位正确率。最后,利用蒙特卡洛法对本文提出的方法进行验证,结果表明:该算法提高了强 混响低信噪比条件下的声源定位正确率。

关键词: 声源定位; 到达时间法; 镜像源模型; 广义互相关 中图分类号: TB95 文献标识码: A 文章编号: 1674-5795(2018)01-0043-05

An Improvement for Accuracy of Acoustic Source Location Based on General Cross-correlation in High Reverberation Environment

JIANG Xiangbin, BAI Zonglong, SUN Jinwei, HUANG Boyan

(Harbin Institute of Technology, Harbin 150001, China)

Abstract: Sound source location based on General Cross-correlation can be divided into two parts-time delay estimation and location estimation. Considering the cost and calculation speed, the location system based on GCC receives more and more concern from researchers. This paper provides a new approach to improve the accuracy of location estimation using signals from two microphones. For that reverberation carries information of location of acoustic source, we can separate the detail which contains information of sound source under assumption of knowing coordinate of microphones. This method can be applied to many aspects especially for humanoid robot.

Key words: sound source location; TOA; image model; GCC

0 引言

声源定位技术可以应用到很多方面,例如视频会 议、人工智能、语音增强以及降噪等^[1-3]。常见的声源 定位方法有到达时间差法(TDOA),波束成形法(BF) 以及空间谱估计算法等。基于广义互相关算法是到达 时间差法的一种,具有计算量小等优势,实际应用广 泛。其中,广义互相关法是指通过对两信号的互功率 谱密度函数相乘并进行加权后进行傅里叶反变换得到 两信号时延的过程。

基于广义互相关的声源定位算法分为两步:首先, 求出两两传声器的时间延迟;然后,结合传声器的空 间位置信息确定声源的位置。本文在强混响条件下结 合镜像源模型对算法进行改进,并进行仿真实验验证。 实验结果证明:在低信噪比(SNR)的条件下,改进后 的定位准确率有一定的提高。 文章首先介绍了现有的声源定位模型;并提出了 一种利用包含在混响中空间位置信息进行声源定位的 改进方法;通过仿真实验结果比较,最终结果表明: 文中所提出的声源定位算法提高了强混响低信噪比条 件下的声源定位正确率。

1 声源定位模型

基于广义互相关的声源定位包含两个步骤:时延 估计以及位置估计。本节将分别介绍这两部分。

在自由场环境下,空间位置不同的两传声器接收 到声源 $x_1(t)$ 的数学模型如^[4]

$$x_1(t) = a_1 s(t - \tau) + n_1(t) \tag{1}$$

$$a_1(t) = a_2 s(t - \tau - \tau_0) + n_2(t)$$
(2)

式中: $x_1(t)$, $x_2(t)$ 分别为1号传声器2号传声器接收 到的信号; $n_1(t)$, $n_2(t)$ 分别为声源传播到两传声器过 程中的噪声; a_1 , a_2 分别为声源传播到两传声器过程 中的衰减,并且有 $0 < a_1$, $a_2 < 1$; τ 为声源传播到 1 号 传声器的时间; τ_0 为声源传播到 1 号传声器与 2 号传 声器的时间差。s(t), $n_1(t)$ 和 $n_2(t)$ 为联合平稳过程, 假设s(t) 与 $n_1(t)$, $n_2(t)$ 不相关。

如图1所示,本文中假设环境为室内,有较强的 混响,传声器数量为4个,声源为平稳宽带声源,则 式(1)和(2)所使用的数学模型过于简单而不能准确描 述声源传播路径。所以引入镜像源模型描述强混响条 件下传声器接收到的信号模型。该模型假设存在关于 各个反射面相对称的镜像源,多个镜像源产生一系列 的冲击响应,以此模拟混响的形成过程。此模型规定 传声器接收到的信号由直达声、混响以及噪声组成, 数学模型如式(3)、式(4)所示。

$$x_{1}(t) = h_{1} \otimes s(t) + h_{1}(t)$$
(3)

$$x_2(t) = h_2 \bigotimes s(t) + n_2(t) \tag{4}$$

式中: h1, h2 为混响过程的冲击响应; ⊗卷积运算。

基于广义互相关的声源定位算法的关键问题在于 计算两两传声器之间的时间差。本文介绍广义互相关 的时延获取以及其改进方法。文中使用的物理模型假 设为平稳的,因此适用于变化缓慢的环境,即信号以 及噪声的参数在整个测量阶段保持平稳。



图1 室内强混响条件下声源定位

基于广义互相关的声源定位算法通过互相关函数计 算两传声器之间的时间差,互相关函数的定义如式(5)。

 $R_{x_{1}x_{2}}(\tau) = (\tau)E[x_{1}(t)x_{2}(t-\tau)]$ (5) 式中: E[·]为求数学期望。

将式(5)最大化,峰值位置对应的 $x_1(t)$ 值就是两 传声器的时间差。在高信噪比(SNR)的情况下,由式 (5)会得到明显的峰值,但是存在混响以及噪声以至于 SNR 较低时,会出现多个峰值甚至无法确定明显的峰 值,导致时延 $x_1(t)$ 的估计出现严重的偏差,最终导致 声源定位算法的失效。为了解决存在噪声以及混响情况下互相关函数求时延失效的问题,引入了广义互相关方法。本文在节中也提出了一种提高时延估计准确率的方法。

通过傅里叶变换可以将 x₁(t)和 x₂(t)的互相关函数与互功率谱密度函数联系起来,如式(6)所示

$$R_{x_1x_2}(\tau) = \int_{-\infty}^{+\infty} G_{x_1x_2}(\omega) e^{jwt} d\omega$$
 (6)

式中: $G_{x_1x_2}(\omega)$ 为互动率谱密度函数,定义如式(7)。

$$G_{x_1x_2}(\boldsymbol{\omega}) = X_1(\boldsymbol{\omega})X_2^*(\boldsymbol{\omega})$$
(7)

式中: $(\cdot)^*$ 为复共轭; $X_1(\omega)$, $X_2(\omega)$ 分别为信号 x_1 (t)和 $x_2(t)$ 的傅里叶变换。

为了降低噪声以及混响的影响,利用滤波器 $f_i(t)$ 对 $x_i(t)$ 进行滤波,滤波后的信号表示为 $y_1(t)$,其中 i=1, 2。所以, $y_1(t)$ 和 $y_2(t)$ 的互动率谱密度函数 $G_{y_{y_2}}$ (ω)表示如式(8)。

$$G_{y_1y_2}(\boldsymbol{\omega}) = F_1(\boldsymbol{\omega}) X_1(\boldsymbol{\omega}) F_2^*(\boldsymbol{\omega}) X_2^*(\boldsymbol{\omega})$$
$$= \Psi(\boldsymbol{\omega}) X_1(\boldsymbol{\omega}) X_2^*(\boldsymbol{\omega})$$
(8)

其中, $\Psi(\omega) = F_1(\omega)F_1^*(\omega)$ 表示频率权重。



图 2 广义互相关流程

表1列举了应用最为广泛的四种加权方法,分别为互相关、ROTH,SCOT 以及 PHAT 加权。图 2 显示 了广义互相关方法的计算流程,其中, $\Psi_{12}(\omega)$ 表示在 表1中选取的加权函数。

表1 常用的广义互相关加权函数

加权函数	权重
传统互相关	1
ROTH	$\frac{1}{G_{x_1x_2}(\omega)}$
SCOT	$rac{1}{\sqrt{G_{x_1x_2}(oldsymbol{\omega})}}$
PHAT	$\frac{1}{\mid G_{x_1x_2}(\boldsymbol{\omega})\mid}$

式(9)表示信号源与传声器之间的关系,其中信号 源位于 $S = (x_0, y_0)$,三个传声器位于 $A = (x_1, y_1)$, $B = (x_2, y_2)$ 以及 $C = (x_3, y_3)$ 。

其中,

$$\begin{cases} SA = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} \\ SB = \sqrt{(x_2 - x_0)^2 + (y_2 - y_0)^2} \\ SC = \sqrt{(x_3 - x_0)^2 + (y_3 - y_0)^2} \end{cases}$$
(10)

式中: τ_{21} 和 τ_{31} 分别为传声器 B 到传声器 A 的时延以及 传声器 C 到传声器 A 的时延: c 为声速。

式(9)通过牛顿法进行求解为

$$\hat{S}_{n+1} = \hat{S}_n - F^{-1}(\hat{S}_n) f(\hat{S}_n)$$
(11)
$$\left[\sqrt{(x - x_n)^2 + (x - x_n)^2} - \right]$$

$$f(\hat{S}_n) = \begin{vmatrix} \sqrt{(x_2 - x_{0n})^2 + (y_1 - y_{0n})^2} \\ \sqrt{(x_1 - x_{0n})^2 + (y_1 - y_{0n})^2} - c\tau_{21} \\ \sqrt{(x_3 - x_{0n})^2 + (y_3 - y_{0n})^2} - \end{vmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$
(12)

$$\left[\sqrt{(x_1 - x_{0n})^2 + (y_1 - y_{0n})^2 - c\tau_{31}}\right]$$
$$F(\hat{S}_n) = \begin{bmatrix} \frac{\partial f_1}{\partial x_0} & \frac{\partial f_1}{\partial y_0} \\ \frac{\partial f_2}{\partial x_0} & \frac{\partial f_2}{\partial y_0} \end{bmatrix}$$
(13)

式中: $\hat{S}_n = (x_{0n}, y_{0n})$ 为声源位置估计: n 为迭代次数。

由公式(11)可以看出声源位置估计是由时延决定 的,所以时延估计的准确度直接影响最终声源位置估 计的正确率。在第3节中给出了强混响条件下的利用 镜像源信息提高时延估计准确度的方法。

2 提高时延估计准确度的方法

在立方体室内空间中,接收到的声音信号存在比 较强的混响。利用镜像源模型得到立方体室内冲击响 应如式所示^[6]

$$h(t) = \sum_{i=-N_x}^{N_x} \sum_{j=-N_y}^{N_y} \sum_{k=-N_z}^{N_z} p(t)_{ijk} \text{ for } 0 \le t \le t_{length}$$
(14)

式中: t_{length} 为冲击响应的长度; $p(t)_{ijk}$ 为来自 ijk 镜 像源的声压信号; N_x , N_y 和 N_z 分别为镜像源的 数目。

如图 3 所示,如果只考虑距离声源比较近的反射

面,式(3)和(4)可以简化为



图 3 单个镜像源传播模型

$$x_1(t) = a_{11}s(t) + a_{12}s(t-\delta_1) + n_1(t)$$
(15)

 $x_{2}(t) = a_{21}s(t-\tau_{21}) + a_{22}s(t-\delta_{2}-\tau_{21}) + n_{2}(t)$ (16)

式中: δ_1 和 δ_2 分别为镜像源 *S*'传播到 1 号传声器的时 延以及传播到 2 号传声器的时延。由于反射面吸收以 及传播衰减的因素,衰减系数 a_{11} , a_{12} , a_{21} , a_{22} 满足 关系 $a_{11}>a_{12}$, $a_{21}>a_{22}$ 。

由图 3 中几何关系,实际声源与镜像源相对于麦 克风满足式(17)和式(18)的关系

$$S'A - SA = c\delta_1 \tag{17}$$

$$S'B - SB = c\delta_2 \tag{18}$$

将互相关函数写为

$$R_{x_{1}x_{2}}(\tau) = E[x_{1}(t-\tau)x_{2}(t)]$$
(19)
将式(17)和(18)带入公式(19),得到
$$R_{x_{1}x_{2}}(\tau) = E[a_{11}a_{21}s(t-\tau)s(t-\tau 21) + a_{11}a_{21}s(t-\tau)s(t-\delta_{2}-\tau 21) + (20)$$

$$a_{12}a_{21}s(t-\tau-\delta_{1})s(t-\tau 21) + a_{12}a_{22}s(t-\tau-\delta_{1})s(t-\delta_{2}-\tau 21)]$$

提出衰减系数项,则
$$R_{x_{1}x_{2}}(\tau) = a_{11}a_{21}E[s(t-\tau)s(t-\tau_{21})] + a_{12}a_{22}E[s(t-\tau)s(t-\delta_{2}-\tau_{21})] + a_{12}a_{21}E[s(t-\tau-\delta_{1})s(t-\delta_{2}-\tau_{21})] + a_{12}a_{22}E[s(t-\tau-\delta_{1})s(t-\delta_{2}-\tau_{21})] + a_{12}a_{2}E[s(t-\tau-\delta_{1})s(t-\delta_{2}-\tau_{21})] + a_{12}a_{2}E[s(t-\tau-\delta_{1})s(t-\delta_{2}-\tau_{21})] + a_{12}a_{2}E[s(t-\tau-\delta_{1})s(t-\delta_{2}-\tau_{21})] + a_{12}a_{2}E[s(t-\tau-\delta_{1})s(t-\delta_{2}-\tau_{21})] + a_{1$$

 $R_{x_{1}x_{2}}(\tau_{21}) = a_{11}a_{21}R_{ss}(\tau_{21}) + a_{11}a_{22}R_{ss}(\tau_{21} + \delta_{2}) + a_{12}a_{21}R_{ss}(\tau_{21} - \delta_{1}) + a_{12}a_{22}R_{ss}(\tau_{21} + \delta_{2} - \delta_{1})$ $\vec{x} \oplus : R_{ss}(\cdot) \end{pmatrix} \vec{y} = \vec{y} \cdot \vec{h} \cdot \vec{$

在仅考虑距离声源最近的反射面以及单次反射的 情况下,由式(22)可知两传声器接收到信号的互相关 函数具有四个峰值,且由 *a*₁₁*a*₂₁>*a*₁₁*a*₂₂, *a*₁₁*a*₂₁>*a*₁₂*a*₂₁, 可以确定四个峰值之间的相对位置关系。最大峰值位 置与其余四个峰值位置关系为

$$(\tau_{21} + \delta_2) + (\tau_{21} - \delta_1) - (\tau_{21} + \delta_2 - \delta_1) = \tau_{21}$$
(23)

式(23)显示了峰值之间位置关系,利用该关系可 以判定由互相关函数确定的时延是否正确,以避免错 误的时延导致错误的声源定位结果。最终,提高声源 定位的准确率,其中,准确率的定义为

$$AC = \frac{N_{\text{correct}}}{N_{\text{total}}}$$
(24)

式中: N_{correct} 为估计正确的实验次数; N_{total} 为实验的总次数。

3 仿真实验

在本节中,给出传统 GCC(Generalized Cross Crorrelation 广义互相关,这种算法可用于声源定位), GCC_ROTH,GCC_SCOT 以及 GCC_PHAT 的随 SNR 变 化的时延估计准确率比较结果。然后利用式(23)分别 对四种方法进行改进,并给出改进后的方法与为改进 方法的比较结果。

图 4 显示了四种方法随 SNR(信噪比)变化的时延 估计准确率的比较结果。其中进行了 1000 次实验,快 拍数为 1024,信号频率为 500 Hz。通过图 4 可知,随 SNR 降低,四种算法的时延估计准确度都会降低,这 种情况在 SNR=-4 dB 之后十分明显。



图 4 基于广义互相关时延估计准确度随 SNR 的变化

为了改善在强混响低 SNR 下的时延估计准确度, 利用式(23)对以上四种算法进行改进。图 5~8 分别显 示了改进后的方法与原始方法时延估计准确度的比较 结果。每次比较进行了 20 次蒙特卡洛实验,每次试验进 行 1000 次时延估计,每次时延估计的快拍数为 1024, 声源的频率为 500 Hz,且仿真在 SNR = -10 dB 情况下 进行。



图 6 GCC_PHAT 与改进后 GCC_PHAT 时延估计准确度的比较



图 7 GCC_ROTH 与改进后 GCC_ROTH 时延估计准确度的比较



图 8 GCC_SCOT 与改进后 GCC_SCOT 时延估计准确度的比较

由图 5~8 可知,再加入式(23)的限制条件之后, 各个方法的时延估计准确度在强混响低 SNR 的情况下 都有一定程度的改善。因此提高了声源定位的准确度, 避免因时延估计错误导致声源定位的错误结果。

4 结论

通过介绍基于广义互相关的声源定位算法的基础 上,针对该算法在强混响低信噪比的情况下时延估计 准确度较低的问题,利用镜像源模型提出了一种提高 时延估计准确度的方法。同过仿真实验结果的分析可 知,本方法利用混响中的位置信息增强抵抗噪声干扰 的能力,进而提高实时延估计的准确度。混响区别于 一般噪声,原因是其含有声源的空间位置信息,如何 利用混响中的声源位置信息进一步提高声源定位的准 确度以及进行三维空间内的镜像源模型分析是我们下 一步的主要研究内容。

参考文献

[1] Escolano J, Xiang N, Perezlorenzo J M, et al. A Bayesian Di-

收稿日期: 2017-08-09; 修回日期: 2017-09-01

基金项目: 宽窄带混合主动噪声控制系统高效稳健算法及应用研究(JJ20140176)



蒋相斌(1975-),男,博士研究生,毕业 于国防大学联合勤务学院,主要研究方向 为噪声特性分析、主动噪声控制。参与项 目有"某型直升机舱内噪声特性分析", "复杂噪声环境噪声测试及主动噪声控制 算法研究"等。



孙金玮,哈尔滨工业大学,仪器科学与技术专业教授,研究方向为信号处理,主动噪声控制。国家自然基金"基于系统自激的 窄带主动噪声控制技术及其应用研究"及 "宽窄带混合主动噪声控制算法研究"负责 人,主持纵向及横向相关课题12项。

美科学家证实量子无线电可助地下水下通信和测绘

美国国家标准技术研究院(NIST)的科学家证实,量子无线 电可以在 GPS、普通手机、无线电信息难以抵达区域,甚至完 全不能工作的地方(例如峡谷、水下和地下)实现通信和测绘。 当 GPS 信号难以穿透水、土壤、建筑物墙壁、摩天大楼时,难 以用于潜艇、扫雷、军事或救灾时,以及无线电信号因瓦砾或 电磁设备干扰的混乱环境下受阻时,该技术可为水手、士兵和 测绘员提供技术支撑。

NIST 科学家正在试验一种低频磁无线电——一种超低频 (VLF)数字调制磁信号。与传统的电磁通信信号不同,该无线 电可以在更高的频率下穿透建筑材料、水和土壤。超低频电磁 场已经在水下通信中得到运用,但音频或视频的数据传输能力 有限,且只能传输单向文本,潜艇还必须拖曳繁琐的无线电缆, 放慢速度并上升到潜望镜深度(18 米或地面以下约 60 英尺)进 行通信。低频通信的最大问题是接收机灵敏度低,现有发射机 的带宽有限,导致数据传输速率低。使用量子传感器将获得最 佳的磁场灵敏度和更长的通信范围,也可以提供像手机一样获 得高带宽通信的可能性。NIST 科学家依靠铷原子的量子特性制 成磁场传感器来检测数字调制磁信号,并通过改变磁场来调制 或控制频率,特别是原子产生的水平和垂直的信号波形。科学 家下一步计划开发改进发射机。

NIST 开发了一种直流磁强计,其中使用偏振光作为检测器 来监测由磁场引起的铷原子"自传"。除了具有高灵敏度以外, 该磁强计还具有室温运行、体积小、功耗低、成本低、干扰小 等特点,不需要移动或校准。下一步,NIST 拟建立和测试一个 定制的量子磁力计,类似原子钟,该磁力计将通过在原子内部 能级及其它特性之间的切换来监测信号,通过提高传感器灵敏 度来扩大低频磁场信号的范围,更好地抑制噪声,增加并有效 利用传感器的带宽。 (摘自计量测控)

rection – of – arrival Model for an Undetermined Number of Sources Using a Two-microphone Array [J]. Journal of the A-coustical Society of America, 2014, 135(135): 742–753.

- [2] Liang R, Wang Q, Tang G, et al. Multiband sound source localization algorithm for directional enhancement in hearing aids
 [J]. Ieej Transactions on Electrical & Electronic Engineering, 2016, 11(3): 331-338.
- [3] Yang S. Research and Realization of Sound Source Localization Algorithms Based on the Time Difference of Arrival [J]. Journal of Information & Computational Science, 2015, 12(16): 5995 -6003.
- [4] Knapp C, Carter G. The Generalized Correlation Method for Estimation of Time Delay [J]. IEEE Transactions on Acoustics Speech & Signal Processing, 1976, 24(4): 320-327.
- [5] Allen J B, Berkley D A. Image method for efficiently simulating small-room acoustics Allen J B, Berkley D A. 1979, 65(4): 943-950.
- [6] Mcgovern S G. Fast Image Method for Impulse Response Calculations of Box-shaped Rooms [J]. Applied Acoustics, 2009, 70 (1): 182-189.