

基于广义互功率谱相位法的声源定位技术

张传义, 卢晓

(东北大学 信息科学与工程学院, 辽宁 沈阳 110819)

摘要: 研究如何在复杂噪声和严重混响的室内环境下,提高声源定位系统的性能. 首先分析室内麦克风阵列的信号模型,确立一种合适的阵列拓扑结构;其次在复倒谱域对滤波器的各项参数进行仿真比较,得出滤波器参数的最佳值;然后提出一种广义互功率谱相位法,改进了峰值不明显等问题. 在 Matlab 中对多种加权函数进行了仿真比较,进一步改善声源定位结果的准确性. 最后对提出的算法进行验证,结果表明,当改进后算法的角度误差和距离误差分别为 $\pm 10^\circ$ 和 ± 15 cm 时,定位成功率能达到 80% 左右.

关键词: 声源定位;麦克风阵列;倒谱法;广义互功率谱相位法

中图分类号: TP 216.1 **文献标志码:** A **文章编号:** 1005-3026(2018)08-1075-05

Sound Source Localization Technology Based on Generalized Cross-Power Spectrum Phase

ZHANG Chuan-yi, LU Xiao

(School of Information Science & Engineering, Northeastern University, Shenyang 110819, China. Corresponding author: LU Xiao, E-mail:1073667526@qq.com)

Abstract: The paper mainly researches how to improve sound source localization effects in indoor environments with complex noise and serious reverberation. Firstly, the signal model of the indoor microphone array is analyzed and a suitable array topology is established. Secondly, a lot of parameters of the filter in the cepstral domain are determined optimally by simulation. Then, the generalized cross-power spectrum phase is proposed to make the peak value more obvious. In addition, many weight functions are simulated and compared in Matlab to further improve the accuracy of sound source localization. Finally, the proposed algorithm is verified. The result shows that the localization success rate is up to 80% when the angle error and distance error of the improved algorithm are $\pm 10^\circ$ and ± 15 cm, respectively.

Key words: sound source localization; microphone array; cepstrum; generalized cross-power spectrum phase

近年来,基于麦克风阵列的声源定位系统已成为现代信号处理领域的研究热点之一,在视频会议、语音识别等领域有着广泛应用^[1];此外,还可以用于机器人的人机对话、助听器等设备^[2-3]. Flanagan 等将麦克风阵列运用于语音加强,又利用阵列信号在混响条件下对声音进行捕获^[4-5]; Silverman 等把传声器阵列引入到语音识别系统,又将阵列信号处理技术运用在说话人的辨别上^[6-7]. 与国外的研究工作相比,无论在理论研究还是在装置开发方面,国内的工作都还处于初级

阶段,因此有必要研究这种声源定位技术.

按照利用信号的信息不同,基于麦克风阵列的方法有基于信号到达角度(DOA)、基于信号能量(RSSI)和基于时延估计(TDE)的3种定位方法. 本文主要通过对广义互功率谱相位法的研究,实现了声源定位.

1 信号模型

在理想信号模型中,麦克风阵列 M_1, \dots, M_n

接收到的信号只有直达的声源信号 $s(t)$ 和噪声信号 $n_1(t), \dots, n_n(t)$. 第 i 个麦克风 M_i 接收到的声音信号 $x_i(t)$ 为

$$x_i(t) = a_i s(t - \tau_i) + n_i(t). \quad (1)$$

式中: a_i 为声音信号的传播效应衰减因子; τ_i 为声音信号 $s(t)$ 到达第 i 个麦克风 M_i 的传播时间; $n_i(t)$ 为第 i 个麦克风 M_i 收到的直达的噪声.

实际环境中麦克风 M_i 接收到的信号为

$$x_i(t) = h_i(t) * s(t) + n_i(t). \quad (2)$$

式中 $h_i(t)$ 为环境脉冲响应函数, $*$ 表示卷积.

2 复倒谱混响消减原理及仿真

2.1 复倒谱混响消减原理

用离散方式表征声音信号时,室内的麦克风阵列采集到的声音信号 $x(t)$ 的表达式为

$$x(m) = \sum_{k=0}^{P-1} h(k)s(m-k). \quad (3)$$

式中 P 是采样长度. 把式(3)变换到频域:

$$X(w) = H(w)S(w). \quad (4)$$

式(4)两边取对数,并做傅里叶逆变换,得出复倒谱关系:

$$\hat{x}(n) = F^{-1}[\ln X(w)] = F^{-1}[\ln H(w)] + F^{-1}[\ln S(w)] = \hat{h}(n) + \hat{s}(n). \quad (5)$$

从复倒谱的特征和实验结果可以得出:当语音信号变换至复倒谱域后,就可以运用“低时窗”过滤掉混响 $\hat{h}(n)$ 部分的影响,然后经过逆向操作,就可以得到相对纯净的语音信号.

2.2 复倒谱域滤波器相关参数的设计

实验的语音信号为手机播放的一段音乐,并在其中加入一段白噪声,其中采样频率为 21.5 kHz,帧长 N 为 512 点,帧移为 128. 复倒谱域滤波器本质上是一种带阻滤波器. 实验中麦克风阵列采集到的声音信号通过复倒谱变换后的波形如图 1 所示.

从图 1 可知,应当设置适当的滤波器滤除中间部分. 滤波器的基本构成如图 2 所示,其中 L 为滤波器最高截止点, h 为过渡带宽.

图 3 所示是滤波器最高截止点对消减混响效果的影响:最高截止点和混响消减效果成反比,当 $L = 4$, 即 $L/N = 1/128$ 时,消减混响的指标 $\Delta SRNR$ (时域信噪比改善指标), $\Delta LPCC$ (线性预测倒谱系数距离改善指标) 和 $\Delta MFCC$ (频率倒谱系数距离改善指标) 达到最优值. 图 4 所示为 $L = 4$ 时过渡带曲线阶数对混响消减效果的影响.

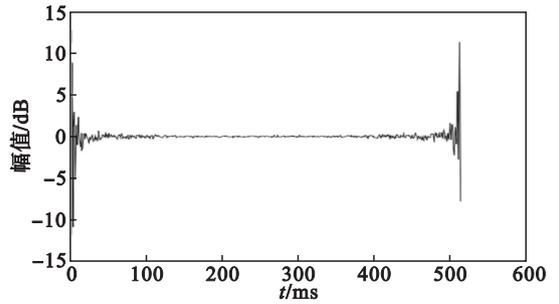


图 1 复倒谱域的语音信号

Fig. 1 Speech signal of complex cepstrum

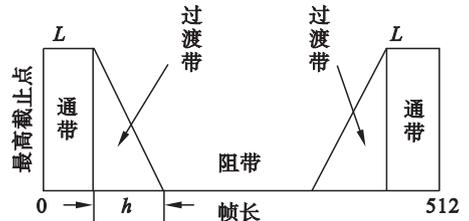


图 2 滤波器特性

Fig. 2 Feature of the filter

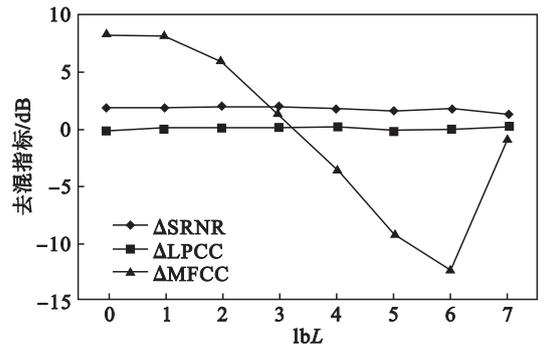


图 3 滤波器不同截止点去混响效果

Fig. 3 Effect of the cut-off point of filter on dereverberation

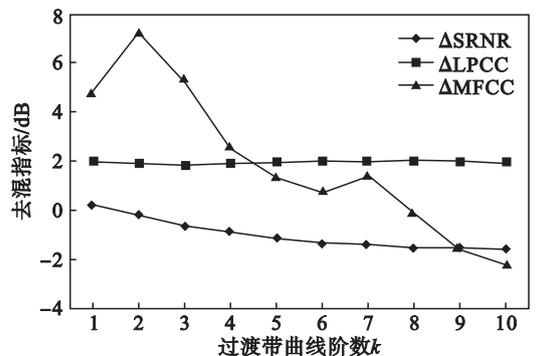


图 4 过渡带曲线阶数对去混响效果的影响

Fig. 4 Effect of the order of transition band on dereverberation

过渡带宽的取值也会影响混响消减的效果. 当最高截止点 $L = 4$, 过渡带曲线阶数 $k = 2$ 时, 过渡带宽对混响消减效果的影响如图 5 所示.

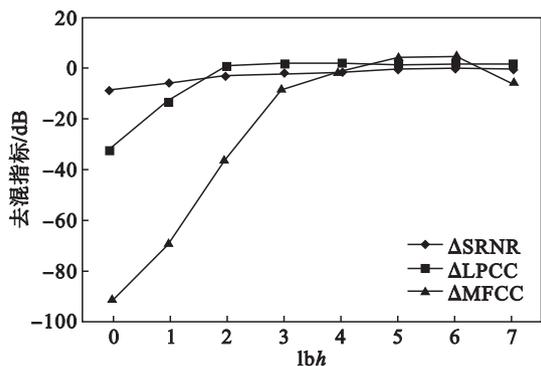


图 5 过渡带宽对去混响效果的影响

Fig. 5 Effect of transition bandwidth on dereverberation

通过对上面各参数的研究得到滤波器的最佳参数值. 当混响时间为 $2^6, 2^7, 2^8, 2^9, 2^{10}$ ms 时, 滤波器最高截止点均为 4, 表明在一定的混响时间范围内, 最高截止点与混响时间无关.

2.3 复倒谱混响消减仿真

取采样频率为 21.5 kHz, 帧长为 512 点, 帧移为 128 点, 汉宁窗的系数为 2.8, 最高截止点 $L = 4$, 过渡带阶数 $k = 2$, 过渡带宽 $h = 64$ 时, 混响消减前后的波形如图 6 所示. 对比图 6b 和图 6c 可以发现, 掺杂在语音信号中的混响, 在经过复倒谱域滤波之后, 可以得到比较纯净的语音信号.

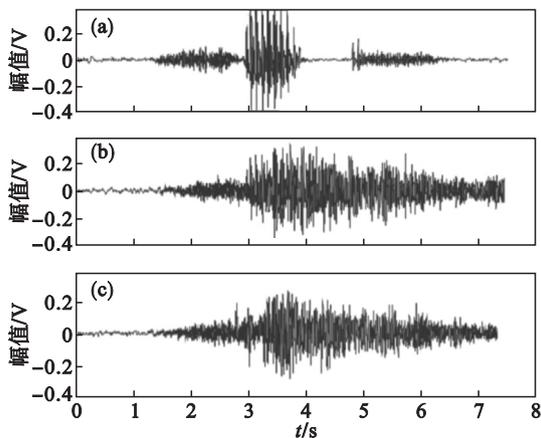


图 6 混响消减前后波形

Fig. 6 Waveform with/without reverberation

(a) — 纯净波形; (b) — 加混响波形; (c) — 去混响波形.

3 广义互功率谱相位法

互功率谱相位法首先对采集到的声源信号加窗分帧, 求其互功率谱函数, 然后在频域对求得的函数加权, 再取反变换得到互相关函数.

在弱噪声和弱混响环境中, 计算两个信号的时延差可以采用互功率谱相位法; 但是, 在强噪声和强混响环境中, 互相关函数的峰值并不明显. 针对

这一问题, 可以对互功率谱相位法进行改进.

3.1 加权函数

首先对互功率谱函数进行频域加权, 常用加权函数有 ROTH 函数、平滑相干 (SCOT) 函数和相位变换 (PHAT) 函数等^[8]. 常用权值 A 见表 1^[9]. 各函数相关系数仿真如图 7 所示.

表 1 常用加权函数
Table 1 Weighting functions commonly used

基本互相关函数	$A = 1$
ROTH 函数	$A = 1/G_{12}(f)$
SCOT 函数	$A = 1/\sqrt{G_{11}(f)G_{22}(f)}$
PHAT 函数	$A = 1/ \sqrt{G_{12}(f)} $

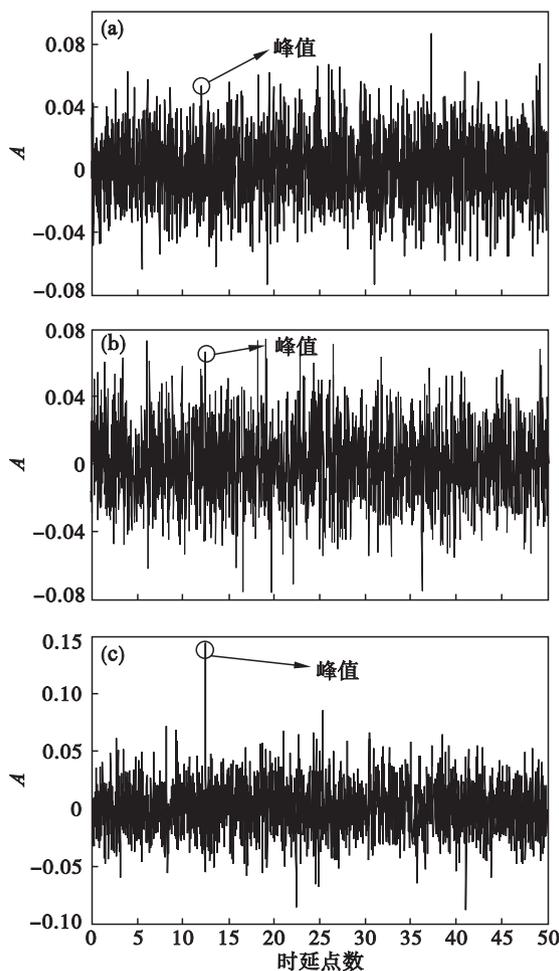


图 7 不同函数的相关系数仿真

Fig. 7 Correlation coefficient simulation with different functions

(a) — ROTH 函数; (b) — SCOT 函数; (c) — PHAT 函数.

仿真研究表明, 在同一实验条件下, ROTH 函数和 SCOT 函数处理后的互相关函数的最大峰值不明显, 得到的不一定是准确的时延估计值, 导致声源位置估计失败; 然而经过 PHAT 函数处理后的互相关函数的最大峰值十分突出, 所以, 本文选择 PHAT 函数.

3.2 限制互相关函数的搜索区间

如图 8 所示,假定相邻的两个麦克风之间的距离是 D ,而声源信号到达麦克风 M_1 和到达麦克风 M_2 的距离差是 d ,显然 $d < D$,两边同时除以声速 c 就可以得到声源信号到达麦克风 M_1 和到达麦克风 M_2 的时间差,搜索的区间应该限制在 $[-Df_s/c, +Df_s/c]$ 范围内,其中 f_s 为采样频率。

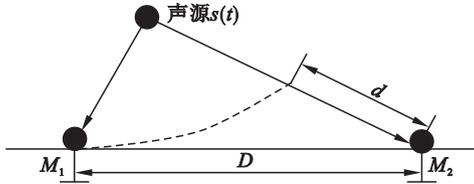


图 8 声源定位示意图

Fig. 8 Schematic of sound source localization

3.3 剔除零点的伪峰值

麦克风在接收语音信号时会引入同步噪声,虽然设计的带通滤波器可以滤除一部分噪声,然而通带内的噪声无法滤除,导致在后续的处理算法中,互相关函数原点处有可能产生不正确的峰值,从而造成估计的时延值不正确^[10]。

综上所述,本文采用广义互功率谱相位法求时延,并对其作进一步改进。

4 麦克风阵列设计

4.1 阵列结构选择

一维均匀直线阵列构造简易、易于理解,算法的难度和复杂度较低;但只能给出声源相对于麦克风阵列的距离和角度。综合考虑系统实现的难易程度、实时性要求、数字信号处理特性等,选择一维均匀线性阵列^[11]。

4.2 阵元的间距和数目

要选定一个合适的麦克风阵元间距,需要考虑音频信号频率和采样频率等因素。查阅文献可知阵元间距必须符合的要求是:最小的阵元间距必须不大于音频信号的一半。本文中音频信号的频率范围在一般情况下为 300 Hz ~ 3 kHz,它的波长范围在 11.3 ~ 113 cm,阵元间距为 5.65 ~ 56.5 cm。当阵元间距为 5.65 cm 时,可分辨的最大时延值为 0.17 ms,而将阵元间距定为 56.5 cm 时,可分辨的最大时延值为 1.7 ms;因此本文把麦克风阵元的间距设置为 57 cm。

确定阵元间距后需要确定阵元的数目。麦克风的数目较少时,麦克风阵列的结构简单,对应的几何算法也不复杂;但会导致系统获得的声源位

置信息比较少,间接影响声源位置估计的准确度。综合考虑,本文选用 4 个麦克风组成阵列。

5 实验与分析

实验测试环境在一间封闭的实验室内,房间内有桌椅、电脑等障碍物,因此麦克风接收到的声源信号会受到噪声和混响的影响。把麦克风阵列水平放置,使声源和阵列处于同一水平线。选择麦克风 M_1 作为参考点,并以其为圆心,在半径 1 ~ 3 m 范围内,角度在阵列正前方的 $0^\circ \sim 180^\circ$ 内,随机选取 10 个点。定位系统配置的采样频率是 21.5 kHz,采样点数是 1024,声源为 1 kHz 的单频信号,幅值可调。将改进前后的定位算法在不同的声源位置处测试 30 次,通过 PC 机界面直接查看声源距离阵列参考点的距离变量 R 和角度变量 θ 。图 9 所示为该实验测试点分布图。

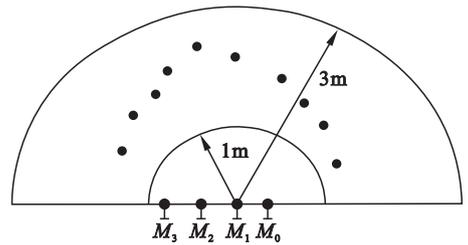


图 9 测试点布局

Fig. 9 Layout of test points

算法改进前后估计值的对比见表 2。可以看出,用改进后算法估计出的声源位置更精确。

表 2 两种算法的估算均值和实际值

Table 2 Actual value and estimated mean value of the two algorithms

声源位置 (θ, R)	改进前估算出的均值 (θ, R)	改进后估算出的均值 (θ, R)
$(30^\circ, 1.30 \text{ m})$	$(9.5^\circ, 1.01 \text{ m})$	$(15.2^\circ, 1.12 \text{ m})$
$(45^\circ, 1.45 \text{ m})$	$(30.5^\circ, 1.30 \text{ m})$	$(36.1^\circ, 1.37 \text{ m})$
$(60^\circ, 1.7 \text{ m})$	$(46.2^\circ, 1.56 \text{ m})$	$(51.5^\circ, 1.84 \text{ m})$
$(90^\circ, 2.3 \text{ m})$	$(101.6^\circ, 2.12 \text{ m})$	$(96.4^\circ, 2.20 \text{ m})$
$(120^\circ, 2.5 \text{ m})$	$(107.3^\circ, 2.38 \text{ m})$	$(113.5^\circ, 2.41 \text{ m})$
$(135^\circ, 2.1 \text{ m})$	$(148.8^\circ, 2.27 \text{ m})$	$(143.7^\circ, 2.21 \text{ m})$
$(165^\circ, 1.5 \text{ m})$	$(142.2^\circ, 1.82 \text{ m})$	$(154.6^\circ, 1.71 \text{ m})$

6 结 语

本文研究了室内麦克风阵列的理想信号模型和实际信号模型,提出一种互功率谱相位法,并对其进行了改进。对本文设计的声源定位系统进行

了实验论证. 假定估计的角度误差为 $\pm 10^\circ$, 距离误差为 ± 15 cm, 则改进后的算法成功率能达到 80%, 明显高于改进前的成功率.

参考文献:

- [1] Huang Y T, Benesty J, Elko G W, et al. Real-time passive source localization; a practical linear-correction least-square approach [J]. *IEEE Transactions on Speech and Audio Processing*, 2001, 9(8): 943 - 956.
- [2] Widrow B, Luo F L. Microphone arrays for hearing aids: an overview [J]. *Speech Communication*, 2003, 39 (1/2): 139 - 146.
- [3] Wang Q H, Ivanov T, Aarabi P. Acoustic robot navigation using distributed microphone arrays [J]. *Information Fusion*, 2004, 5(2): 131 - 140.
- [4] Flanagan J L. Beam width and usable bandwidth of delay-steered microphone arrays [J]. *Bell Labs Technical Journal*, 1985(4): 983 - 985.
- [5] Flanagan J L, Berkley D A, Elko G W, et al. Autodirective microphone systems [J]. *Acta Acustica united with Acustica*, 1991, 73(2): 58 - 71.

- [6] Silverman H F. Some analysis of microphone arrays for speech data acquisition [J]. *IEEE Transactions on Acoustics Speech & Signal Processing*, 2003, 35(12): 1699 - 1712.
- [7] Silverman H F, Kirtman S E. A two-stage algorithm for determining talker location from linear microphone array data [J]. *Computer Speech & Language*, 1992, 6(2): 129 - 152.
- [8] Arslan G, Sakarya F A. A unified neural-network-based speaker localization technique [J]. *IEEE Transactions on Neural Networks*, 2000, 11(4): 997 - 1002.
- [9] Chen J D, Benesty J, Huang Y T. Performance of GCC- and AMDF-based time-delay estimation in practical reverberant environments [J]. *EURASIP Journal on Applied Signal Processing*, 2005 (1): 25 - 36.
- [10] Bian X H, Rehag J M, Abowd G D. Sound source localization in domestic environment [R]. Atlanta: Gvu Center, Georgia Institute of Technology, 2004: 19 - 36.
- [11] 张传义, 米常伟. 基于 TDE 技术的声源定位算法研究 [J]. 东北大学学报 (自然科学版), 2014, 35(3): 333 - 336. (Zhang Chuan-yi, Mi Chang-wei. Research on sound source localization system based on time delay estimation [J]. *Journal of Northeastern University (Natural Science)*, 2014, 35(3): 333 - 336.)

(上接第 1068 页)

4 结 语

1) 基于伸缩杆结构的光纤光栅多级传感系统可对温度漂移影响进行自补偿.

2) 在保证测量分辨力及灵敏度的同时扩大了传感器的测量范围.

3) 传感器结构紧凑, 易于多路复用.

4) 传感系统测量范围大, 系统稳定性高, 调整结构参数后, 有望应用于汽车载荷动态称重领域.

参考文献:

- [1] Mihailov S J. Fiber Bragg grating sensors for harsh environments [J]. *Sensors*, 2012, 12(2): 1898 - 1918.
- [2] Kesavan K, Sundaram B A, Ahmed A K F, et al. Performance assessment of indigenously developed FBG strain sensors under short-term and long-term loadings [J]. *Sādhanā*, 2015, 40(2): 577 - 590.
- [3] Panopoulou A, Loutas T, Roulias D, et al. Dynamic fiber Bragg gratings based health monitoring system of composite aerospace structures [J]. *Acta Astronautica*, 2011, 69(7/8): 445 - 457.

- [4] Lydon M, Taylor S E, Robinson D, et al. Development of a bridge weigh-in-motion sensor; performance comparison using fiber optic and electric resistance strain sensor systems [J]. *IEEE Sensors Journal*, 2014, 14(12): 4284 - 4296.
- [5] Qiao X, Shao Z, Bao W, et al. Fiber Bragg grating sensors for the oil industry [J]. *Sensors*, 2017, 17(3): 429 - 438.
- [6] Hussaini S K K, Indraratna B, Vinod J S. Application of optical-fiber Bragg grating sensors in monitoring the rail track deformations [J]. *Geotechnical Testing Journal*, 2015, 38(4): 387 - 396.
- [7] Kuang K S C, Quek S T, Tan C Y, et al. Plastic optical fiber sensors for measurement of large strain in geotextile materials [J]. *Advanced Materials Research*, 2008, 47/48/49/50: 1233 - 1236.
- [8] Zhao Y, Chen K, Yang J. Novel target type flowmeter based on a differential fiber Bragg grating sensor [J]. *Measurement*, 2005, 38(3): 230 - 235.
- [9] Zhao Y, Meng Q, Chen K. Novel current measurement method based on fiber Bragg grating sensor technology [J]. *Sensors and Actuators A: Physical*, 2006, 126(1): 112 - 116.
- [10] 吴付岗, 张庆山, 姜德生, 等. 光纤光栅 Bragg 波长的高斯曲线拟合求法 [J]. 武汉理工大学学报, 2007, 29(12): 116 - 118. (Wu Fu-gang, Zhang Qing-shan, Jiang De-sheng, et al. Method of Gaussian curve fitting for measuring fiber Bragg wavelength [J]. *Journal of Wuhan University of Technology*, 2007, 29(12): 116 - 118.)