

# 基于圆形麦克风阵列的声源定位改进算法

郭业才<sup>1,2</sup> 宋官琨琨<sup>2</sup> 吴礼福<sup>1,2</sup> 孙心宇<sup>1,2</sup> 王丽华<sup>3</sup>

(1. 南京信息工程大学电子与信息工程学院, 南京, 210044; 2. 江苏省大气环境与装备技术协同创新中心, 南京, 210044; 3. 南京信息工程大学信息与控制学院, 南京, 210044)

**摘要:** 针对波达方向估计中传统互功率谱法声源方位估计准确性差、方位模糊的问题, 提出了一种基于圆形麦克风阵列的声源定位改进算法, 并进行了实验验证。在该改进算法中, 先设计了十二元圆形麦克风阵列, 由麦克风对接收语音信号的时延与相位得到相位旋转因子, 再将其引入到语音信号的互功率谱中, 新定义了圆形集成互功率谱, 由该功率谱进行声源方位估计。仿真与实测实验结果表明, 本文的圆形集成互功率谱法对声源方位进行估计, 估计的准确度高于传统互功率谱法。

**关键词:** 波达方向; 互功率谱; 时间延迟; 相位旋转因子; 麦克风阵列

**中图分类号:** O429 **文献标志码:** A

## Improved Algorithm of Sound Source Localization Using Circular Microphone Array

Guo Yecai<sup>1,2</sup>, Song Gong Kunkun<sup>2</sup>, Wu Lifu<sup>1,2</sup>, Sun Xinyu<sup>1,2</sup>, Wang Lihua<sup>3</sup>

(1. College of Electronic & Information Engineering, Nanjing University of Information Science & Technology, Nanjing, 210044, China; 2. Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology (CICAET), Nanjing, 210044, China; 3. School of Information & Control, Nanjing University of Information Science & Technology, Nanjing, 210044, China)

**Abstract:** In view of the problems of inaccurate localization and ambiguous orientation of traditional cross power spectrum (CS) in direction of arrival estimation, an improved algorithm of sound source localization using a circular microphone array is proposed and demonstrated by practical experiments. In the proposed algorithm, the twelve-element circular microphone array is firstly designed, then the phase rotation factor is defined by combining the time delay with the phase information of speech signals received by each pair of microphone and introduced into the traditional cross power spectrum to define novel circular integrated cross power spectrum (CICS), finally the sound source orientation can be estimated by CICS algorithm. The simulation and practical experiment results show that the proposed algorithm can greatly improve the estimated accuracy of source orientation compared with the traditional cross spectrum algorithm.

**Key words:** direction of arrival (DOA); cross power spectrum; time delay; phase rotation factor; microphone array

**基金项目:** 江苏省高校自然科学基金(13KJA510001)重大资助项目; 江苏省高校科研成果产业化推进 (JHB2012-9)资助项目; 江苏省高校“信息与通信工程”优势学科建设工程(2014)资助项目; 江苏省六大人才高峰(2008026)资助项目; 国家自然科学基金青年基金(51405241)资助项目。

**收稿日期:** 2015-01-08; **修订日期:** 2015-01-28

## 引言

基于波达方向(Direction of arrival, DOA)估计的麦克风阵列声源定位技术在很多领域都是研究的热点,主要应用于声呐、视频会议、人工智能、地震研究、语音追踪与识别、监控系统等方面<sup>[1]</sup>。波达方向估计的基本原理是运用麦克风采集到语音信号的相位信息,通过计算进行声源定位。目前常用的波达方向估计方法主要有:基于可控波束形成器输出功率最大化方法,例如:最小方差波束形成器<sup>[2]</sup>(Minimum variance beamformer, MVB)、延迟和求和波束形成器<sup>[3]</sup>(Delay and sum beamformer, DSB);基于高分辨率谱估计的声源定位法,例如:MUSIC法<sup>[4]</sup>,ESPRIT法<sup>[5]</sup>。

理论上,随着麦克风个数的增加,波达方向估计的定位精度随之提高,但计算复杂度也会提高。为了降低复杂度,一方面运用非线性的方法对计算进行优化,这样会产生定位不精确问题<sup>[6]</sup>,另一方面运用线性麦克风阵列对阵列模型进行简化,这样会产生定位模糊问题<sup>[7]</sup>(无法准确区分声源位于阵列的前方或者后方)。针对以上问题,本文构建了十二元圆形麦克风阵列,由麦克风对接收语音信号的时延与相位信息得到相位旋转因子,再将其引入到语音信号的互功率谱中,得到圆形集成互功率,由该功率谱进行声源方位估计,并由仿真与实验来验证声源方位估计的准确性。

## 1 圆形麦克风阵列

如图1所示,建立圆形麦克风阵列模型。由 $m$ 个相同的麦克风等间距的排列组成,阵列的外接圆半径为 $R$ ,2个毗连的麦克风之间的夹角为 $\alpha$ (如 $\angle M_1OM_2$ ),间距为 $l$ (如 $M_1M_2=l$ )。在远场条件下,声源 $S(t)$ 位于 $\theta$ 方向。 $m$ 个阵元围绕旋转成一个圆阵,本文约定:当且仅当 $i \equiv j \pmod{m}$ 时,阵元 $M_i$ 与 $M_j$ 为同一个麦克风。

如图2所示,根据几何学知识,作垂直于弦 $\overline{M_1M_2}$ 的垂线 $QM_1$ ,声源位于远场,声波的入射波为相互平行的平面波,则 $\angle QM_1N_1 = \theta$ ,作 $M_2N_1 \perp M_1N_1, M_3N_2 \perp M_2N_2$ ,可得

$$M_iN_i = l \sin(\theta + (i-1)\alpha) \quad (1)$$

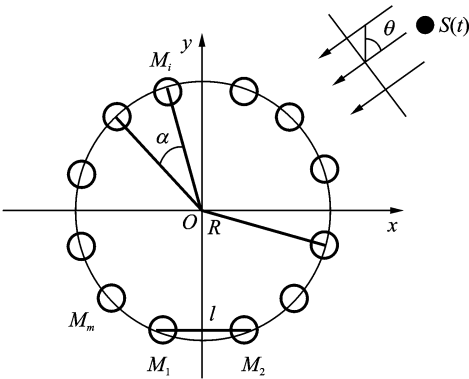


图1 圆形麦克风阵列

Fig. 1 Circular microphone array

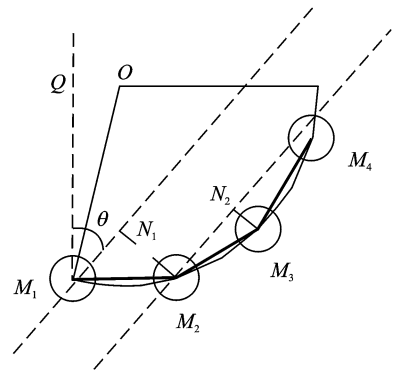


图2 阵列几何作图

Fig. 2 Geometry drawing of array

## 2 波达方向估计

根据图1,2个毗连的麦克风间的时延值为

$$\tau_{M_{i-1}M_i}(\theta) = \frac{l}{c} \sin(\theta + (i-1)\alpha) \quad (2)$$

式中: $c$ 为声速。 $x_{i-1}(t), x_i(t)$ 分别为第 $i-1$ 个、第 $i$ 个麦克风采集到的语音信号; $\tau_{M_{i-1}M_i}$ 为2个麦克风间的时延值,则

$$x_i(t) = s(t) + n_i(t) \quad (3)$$

$$x_{i-1}(t) = s(t - \tau_{M_{i-1}M_i}) + n_{i-1}(t) \quad (4)$$

式中:  $s(t)$  为声源入射信号;  $n_i(t)$  为环境噪声。

对  $x_{i-1}(t)$ ,  $x_i(t)$  进行傅里叶变换, 得

$$X_i(\omega) = S(\omega) + N_i(\omega) \quad (5)$$

$$X_{i-1}(\omega) = S(\omega)e^{-j\omega\tau_{M_{i-1}M_i}} + N_{i-1}(\omega) \quad (6)$$

式中:  $\omega$  为频率。

麦克风间的噪声信号不相关, 在频域  $x_{i-1}(t)$ ,  $x_i(t)$  的互相关函数<sup>[8]</sup>为

$$\Phi_{M_{i-1}M_i}^{\omega} = E[X_{i-1}(\omega)X_i^*(\omega)] = \Phi_{ss}(\omega)e^{-j\omega\tau_{M_{i-1}M_i}(\theta)} \quad (7)$$

式中:  $E[\cdot]$  表示数学中的期望;  $\Phi_{ss}(\omega)$  为  $s(t)$  的功率谱密度。在 DOA 估计求取入射角, 用频域代替了时域, 这里引入互功率谱的相位为

$$P_{M_{i-1}M_i}^{\omega}(\omega) = \frac{\Phi_{M_{i-1}M_i}^{\omega}}{|\Phi_{M_{i-1}M_i}^{\omega}|} \quad (8)$$

考虑到实际应用, 选取互功率谱峰值所对应的频率值为

$$\omega_i^{\max} = \arg \max |\Phi_{M_{i-1}M_i}^{\omega}| \quad (9)$$

由  $\varphi(0 \leq \varphi \leq 2\pi)$  方向传递的语音信号, 对于麦克风  $M_1M_2$  与  $M_{i-1}M_i$  的时延差值为

$$\tau_{M_{i-1}M_i}(\varphi) = \tau_{M_1M_2}(\varphi) - \tau_{M_{i-1}M_i}(\varphi) \quad (10)$$

那么, 相位旋转因子(Phase rotation factor, PRF)定义为

$$\text{PRF}_{M_{i-1}M_i}^{\omega}(\varphi) = e^{-j\omega\tau_{M_{i-1}M_i}(\varphi)} \quad (11)$$

结合式(8, 11), 将圆形集成互功率谱(Circular integrated cross spectrum, CICS)定义为

$$\text{CICS}(\varphi) = \sum_{i=1}^M \text{PRF}_{M_{i-1}M_i}^{\omega}(\varphi) P_{M_{i-1}M_i}^{\omega}(\omega) \quad (12)$$

在实际中, 依据式(9), 则式(12)可进一步改写为

$$\text{CICS}(\varphi) = \sum_{i=1}^M \text{PRF}_{M_{i-1}M_i}^{\omega_{i-1}^{\max}}(\varphi) P_{M_{i-1}M_i}^{\omega}(\omega_{i-1}^{\max}) \quad (13)$$

根据 DOA 所求得的声源方位角为

$$\hat{\theta} = \arg \max_{0 \leq \varphi \leq 2\pi} \text{CICS}(\varphi) \quad (14)$$

## 3 误差分析

### 3.1 混响影响分析

混响是由于空间中存在多个传播反射路径造成的<sup>[9]</sup>。本文运用基于 CICS 的 DOA 算法, 其中每个部分都是在不同的空间位置获得数据, 故式(12)中的混响成分都是不相关的。本文基于 CICS 的 DOA 算法对混响有较好的抑制效果。

### 3.2 噪声影响分析

由于麦克风在采集语音信号的同时也会采集到环境噪声。由式(2), 有

$$\theta = \arcsin\left(\frac{\tau_{M_{i-1}M_i}C}{l}\right) - (i-1)\alpha \quad (15)$$

由式(15)可知, 时延值  $\tau_{M_{i-1}M_i}$  会受到噪声影响, 进而对声源角度  $\theta$  产生影响。为了评估噪声对于角度估计产生的影响, 用平均绝对误差来评估本文算法的角度估计性能。

### 3.3 语音信号预处理

由于语音信号是一个时变的、非稳态的信号, 直接处理会使定位结果出现较大的误差。故需对采集到的语音信号进行预处理: 分帧、加窗, 这样每帧的语音信号在短时范围内特征变化较小, 可以作为稳态

来处理,帧长为 30 ms,帧移为 15 ms(保证相邻帧之间有重叠),窗函数为汉明窗。

### 4 仿真实验

仿真环境为 5.5 m×3.3 m×2.3 m 的房间冲激响应模型<sup>[10]</sup>。所用阵列如图 1 所示,阵列半径  $R=0.05$  m,阵列位于房间的西南角位置,采样率  $f_s=16\ 000$  Hz,声速  $c=343$  m/s;声源为成年人声音(男声、女声),位于  $\theta$  方位角,距离阵元中心 1.5 m,噪声选取高斯白噪声。仿真实验中,以基于传统互功率谱法(Cross power spectrum,CS)的声源方位估计算法为比较对象,对基于圆形集成互功率谱法(Circular integrated cross power spectrum,CICS)的声源定位估计算法性能进行验证。声源位置不变时,声源位于  $\theta=70^\circ$ ,对于不同信噪比(SNR={5,10,15,20} dB),运用基于 CICS 的 DOA 算法,估计声源方位角,如图 3 所示。声源位置变化时,声源持续时间为 12 s,距离阵元中心 1.5 m,从  $\theta=70^\circ$  时开始运动,沿顺时针方向,以  $20^\circ$  为一步围绕圆形阵列移动,移至  $\theta=170^\circ$  时为止。运用基于 CICS 的 DOA 算法,估计声源方位角。麦克风个数  $m=12$ ,信噪比 SNR=20 dB。方位角估计值和误差值如图 4,5 所示。

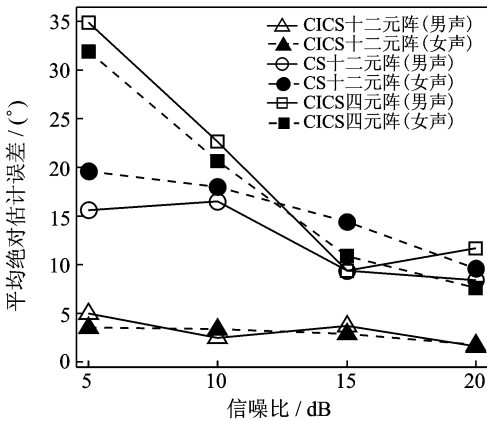


图 3 方位角估计的平均绝对误差仿真值  
Fig. 3 Mean absolute error simulation of azimuth estimation

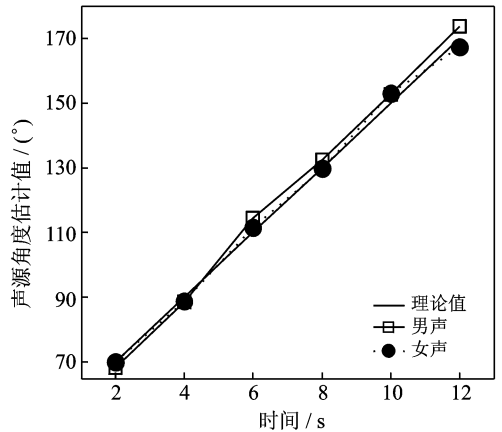


图 4 方位角估计仿真值  
Fig. 4 Simulation value of azimuth estimation

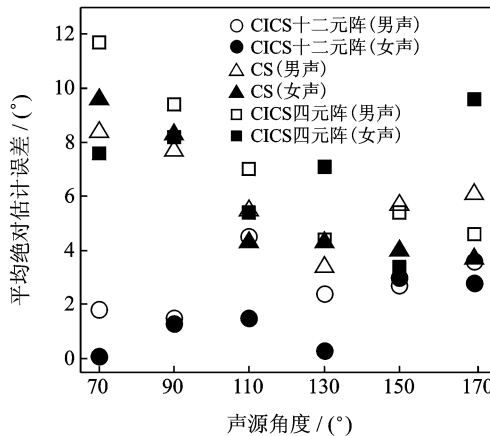


图 5 不同声源角度时方位角估计平均绝对误差仿真值  
Fig. 5 Mean absolute error simulation value of azimuth estimation

图 3 表明,经过多次实验取平均值,随着信噪比的提高,声源方位角估计的精度提高,且在 SNR=20 dB 时,方位角估计误差小于  $2^\circ$ ;在信噪比不变的条件下,与基于 CS 算法的十二元圆形阵列和基于 CICS 算

法的四元十字阵列相比,基于 CICS 算法的十二元圆形阵列声源方位角估计的平均绝对误差最低,在 SNR = {5, 10, 15, 20} dB 时,方位角估计误差都小于  $5^\circ$ ,声源方位角估计的精度提高。图 4 表明,在信噪比不变的条件下,随着声源位置的变动,方位角估计值与理论值相比较为接近。图 5 表明,在相同信噪比条件下,与基于 CS 的十二元圆形阵列和基于 CICS 算法的四元十字阵列对比,基于 CICS 算法的十二元圆形阵列声源方位角估计的平均绝对误差最低,声源方位角估计的精度得到了提高。CICS 算法的最大平均绝对误差为:十二元阵:男声  $4.5^\circ$ ,女声  $3^\circ$ ;四元阵:男声  $11.7^\circ$ ,女声  $8.2^\circ$ 。CS 算法的十二元阵为:男声  $8.4^\circ$ ,女声  $9.6^\circ$ 。

## 5 实测实验

实测环境为全消声实验室,房间尺寸为  $5.5\text{ m} \times 3.3\text{ m} \times 2.3\text{ m}$ ,阵列摆放位置与仿真环境相同,声源高度、阵列高度都为  $1.2\text{ m}$ ,声源距离阵列中心距离为  $1.5\text{ m}$ ,阵列为 12 个麦克风的圆形阵列。实验器材:数据采集设备为 16 通道的 PXIE-4496 数据采集卡、配套 PC 机(Intel 2 GHz Core i7 CPU, 2 GB RAM);声源为某公司的 AM012 人工嘴(输出声压级  $94\text{ dB}$ ),麦克风为简易声音传感器模块(全向性,工作电压  $5\text{ V}$ )。

由于人工嘴在通电工作时产生的嘶嘶声、房间换气扇转动时产生的呼呼声,实测环境下的信噪比平均为  $20\text{ dB}$ 。将本文基于 CICS 的声源定位算法的实测结果与仿真结果进行对比。声源位置不变时,方位角估计的平均误差如图 6 所示。声源位置变化时,声源方位角估计结果如图 7, 8 所示。

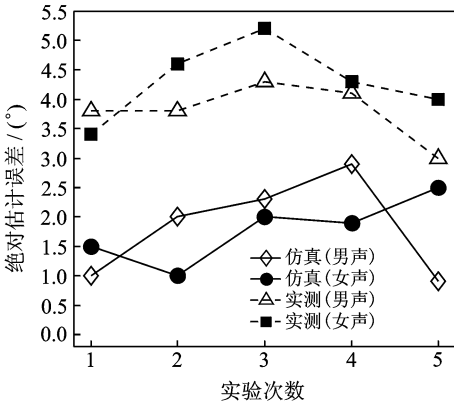


图 6 方位角估计的平均绝对误差实测值

Fig. 6 Mean absolute error measured value of azimuth estimation

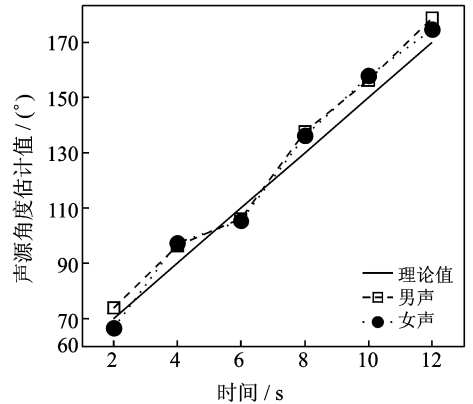


图 7 方位角估计实测值

Fig. 7 Measured value of azimuth estimation

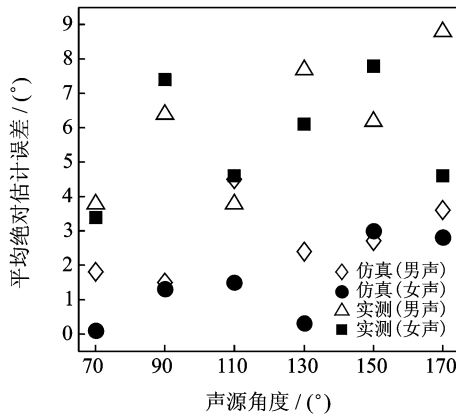


图 8 方位角估计平均绝对误差实测值

Fig. 8 Mean absolute error measured value of azimuth estimation

图6表明,由于存在实际环境噪声以及采集设备A/D转换等影响因素,因此实测结果与仿真结果会有一定偏差,仿真与实测结果的平均绝对估计误差分别为:仿真:男声 $1.8^\circ$ ,女声 $1.7^\circ$ ;实测:男声 $3.8^\circ$ ,女声 $4.3^\circ$ 。图7,8表明,由于存在实际因素影响造成实测与仿真结果有一定偏差,且实测结果与仿真结果的最大平均绝对误差分别为:仿真:男声 $4.5^\circ$ ,女声 $3^\circ$ ;实测:男声 $8.8^\circ$ ,女声 $7.8^\circ$ 。

## 6 结束语

本文通过定义相位旋转因子并将其引入到传统互功率谱中,提出了一种基于圆形麦克风阵列的声源定位改进算法,该算法通过新定义的CICS法来进行声源方位估计。仿真与实际实验结果表明,本文的改进算法有较好的方位估计性能。未来将考虑多声源定位研究,进一步改进本文算法,如通过在语音信号处理中运用盲源分离方法来分离麦克风采集到的多路声源信号,然后用分离后的语音信号分别进行声源定位,达到多声源定位效果,或者是运用短时傅里叶变换,在“时间-频域”域上运用稀疏成分分析法来确定主声源域,然后对各主声源域进行声源定位,达到多声源定位效果;同时也将进一步考虑噪声等影响因素对声源方位估计的影响,完善相应的算法研究,提高定位精度。

### 参考文献:

- [1] 居太亮. 基于麦克风阵列的声源定位算法研究[D]. 成都:电子科技大学,2006.  
Ju Tailiang. Research on speech source localization methods based on microphone arrays [D]. Chengdu: University of Electronic Science and Technology, 2006.
- [2] Barbara R, Friedrich F, Dietrich K. An analysis of nonstationary variance estimates in the maximum negentropy beamformer [C]//Joint Workshop on Hands-free Speech Communication and Microphone Arrays. Edinburgh, England: IEEE, 2011: 201-206.
- [3] 陈华伟,王天南,张峰,等. 基于抽头稀疏化的最小二乘频率不变形成器设计[J]. 数据采集与处理,2014,29(2):191-197.  
Chen Huawei, Wang Tiannan, Zhang Feng, et al. Least square approach to design of frequency invariant beamformers with sparse tap coefficients [J]. Journal of Data Acquisition and Processing, 2014,29(2):191-197.
- [4] Junli Liang, Ding Liu. Passive localization of mixed near-field and far-field sources using two-stage MUSIC algorithm[J]. IEEE Transactions on Signal Processing, 2010,58(1):108-120.
- [5] Haohai S, Heinz T, Edwin M. Robust localization of multiple sources in reverberant environments using EB-ESPRIT with spherical microphone arrays[C]//IEEE International Conference on Acoustics Speech and Signal Processing. Prague, Czech Republic: IEEE, 2011:117-120.
- [6] Dmochowski J, Zibulevsky M. Direction of arrival estimation using the parameterized spatial correlation matrix[J]. IEEE Transactions on Audio, Speech and Language Processing, 2007,15(4):1327-1339.
- [7] 王瑞,陈砚圃,张恒. 声源定位中波达角一致性检测方法[J]. 数据采集与处理,2010,25(6):756-760.  
Wang Rui, Chen Yanpu, Zhang Heng. Method for detecting DOA homogeneity in sound source localization[J]. Journal of Data Acquisition and Processing, 2010,25(6):756-760.
- [8] Amin K, Akihiko S. A DOA estimation method for an arbitrary triangular microphone arrangement [C]//14th European Signal Processing Conference. Florence, Italy: EURASIP, 2006:1-5.
- [9] 张德会,陈光冶. 复倒谱域语音信号去混响研究[J]. 声学技术,2009,28(1):39-44.  
Zhang Dehui, Chen Guangye. Speech signal dereverberation with cepstral processing[J]. Technical Acoustics, 2009,28(1): 39-44.
- [10] Antonacci F, Filos J, Thomas M R P. Inference of room geometry from acoustic impulse responses[J]. IEEE Transactions on Audio, Speech and Language Processing, 2012,20(10):2683-2695.

作者简介:郭业才(1962-),男,教授,博士生导师,研究方向:信号处理、盲均衡技术,E-mail:guo-yecai@163.com;宋宫琨(1989-),男,硕士研究生,研究方向:语音信号信号处理;吴礼福(1981-),男,博士,研究方向:语音信号信号处理;孙心宇(1990-),女,博士,研究方向:阵列信号信号处理;王丽华(1974-),女,博士,研究方向:机械电子与图像处理。

