

麦克风阵列二维方向估计聚焦算法研究

居太亮, 彭启琮, 邵怀宗, 林静然

(电子科技大学通信与信息工程学院 成都 610054)

【摘要】 聚焦算法是解决宽带信号方向(DOA)估计问题的重要算法。常规的宽带聚焦算法需要进行预估计和多次迭代, 算法的计算复杂度较高。该文根据宽带聚焦算法的思想, 结合语音信号的特点, 以各频率点的频域相关矩阵和聚焦点的相关矩阵的均方误差最小为最优化准则, 提出了一种不需要预估计和迭代的二维宽带聚焦算法。该算法实时性好、计算复杂度低。采用八元均匀圆环麦克风阵列, 仿真验证了该算法的良好二维DOA估计性能。

关键词 频率聚焦算法; 麦克风阵列; 子空间方法; 宽带信号方向估计

中图分类号 TN914

文献标识码 A

Research on 2D DOA Estimations Focusing Methods Based on Microphone Arrays

JU Tai-liang, PENG Qi-cong, SHAO Huai-zong, and LIN Jing-ran

(School of Communication and Information Engineering, University of Electronic Science and Technology of China Chengdu 610054)

Abstract The focusing algorithms are the key to solve the wideband signal direction of arrival (DOA) problem. Classical wideband focusing DOA estimation algorithms have such demerits as demand of initial information of DOA and several iterations as well as bad performance for real-time applications. A new type algorithm based on wideband focusing theory is presented in this paper. It does not require pre-estimations and iterations. The optimal criterion of this algorithm is the minimum mean square error (MMSE). This algorithm has the advantage of the low computational complexity for the real-time applications. Simulation based on MATLAB shows the high performance of the proposed algorithm in planar uniform circle microphone array with eight elements.

Key words frequency focusing algorithm; microphone array; subspace method; wideband DOA estimation

基于麦克风阵列的语音拾取技术, 由于结合了时、频、空的信号处理, 提高了输出信号的信噪比, 已广泛用于通信、控制等领域^[1]。信号源到达方向(DOA)/位置估计是实现空间滤波的关键问题之一。由于麦克风阵列接收到的是短时平稳、超宽带自然信号, 其位置参数估计十分困难。目前的算法主要是把传统的阵列处理算法加以推广, 以适应麦克风阵列处理的新特性。代表算法主要有: 近场二维MUSIC算法^[2]、远场二维(2D)MUSIC算法^[3]、近场三维(3D)MUSIC算法和近场子阵算法^[4], 但是这几种算法均是窄带处理算法在具体应用环境的直接推广, 计算复杂度较高。

聚焦算法是把频带内不重叠的信号空间聚焦到

参考频率点, 得到一个单一频率点(参考频率)的数据协方差矩阵, 再用窄带信号处理的方法进行DOA估计^[5-8]。目前已报道的算法一般要求对信号的方向信息进行预估计和迭代。并且这些算法主要针对平稳信号作一维参数估计, 其分数带宽比小于1.5。麦克风阵列处理的语音信号是短时平稳信号, 分数带宽比接近2。本文根据宽带聚焦算法的思想, 结合语音信号的特点提出了一种基于麦克风阵列的二维方向估计聚焦算法, 该算法不需要预估计和迭代, 计算复杂度低。

1 基于任意结构的阵列信号接收模型

设空间有 D 个声源 S_1, S_2, \dots, S_D , M 个全向无

收稿日期: 2006-10-20; 修回日期: 2007-06-18

基金项目: 四川省科技攻关项目(04GG021-020-02)

作者简介: 居太亮(1974-), 男, 博士生, 主要从事信号处理、语音信号处理、阵列信号处理等方面的研究。

差异麦克风(声压传感器) P_1, P_2, \dots, P_M 组成一个 M 元的麦克风阵列, 则阵列接收到的信号在频域可以表示为^[4]:

$$\mathbf{F}(f) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-j2\pi f\tau_{21}} & e^{-j2\pi f\tau_{22}} & \dots & e^{-j2\pi f\tau_{2D}} \\ \vdots & \vdots & \vdots & \vdots \\ e^{-j2\pi f\tau_{M1}} & e^{-j2\pi f\tau_{M2}} & \dots & e^{-j2\pi f\tau_{MD}} \end{bmatrix} \begin{bmatrix} F_{11}(f) \\ F_{12}(f) \\ \vdots \\ F_{1D}(f) \end{bmatrix} + \begin{bmatrix} n_1(f) \\ \vdots \\ n_M(f) \end{bmatrix} = \mathbf{A}(f, \theta, \phi) \mathbf{S}(f) + \mathbf{N}(f) \quad (1)$$

式中 f 为信号频率, 可以通过 FFT 或者 DFT 变换得到; τ_{ij} 为第 i 个传感器接收到第 j 个信号源发出的信号的相对时延, $i=1, 2, \dots, M$, $j=1, 2, \dots, D$, $\tau_{1j}=0$; $F_{1j}(f)$ 为第 1 个传感器接收到来自第 j 个信号源, 并且经过 DFT 或 FFT 变换后的离散频域信号; $n_i(f)$ 为第 i 个传感器上收到的噪声信号。假设第 j 个信号源的方向矢量为 \mathbf{u}_j , 第 i 个传感器的位置矢量为 \mathbf{P}_i , 且 $\mathbf{u}_j = [-\sin\theta_j \cos\phi_j - \sin\theta_j \sin\phi_j - \cos\theta_j]^T$, $\mathbf{P}_i = (x_i, y_i, z_i)$, 参考传感器设在坐标原点, 则时延 $\tau_{ij} = \mathbf{u}_j^T \mathbf{P}_i / c$, c 为声波波速, 在室温下取 343 m/s; θ_j 为第 j 个信源与 Z 轴的夹角; ϕ_j 为第 j 个信源在 XOY 平面的投影与 X 轴的夹角。

2 基于相关矩阵的二维宽带聚焦算法

对于语音信号, 信号能量一般集中在 100~4 000 Hz 之间, 频带较宽, 直接采用聚焦算法很难得到理想的效果。假设式(1)中的数据是通过 N_{FFT} 点 FFT 获得的, 有用信号的对应离散频率点为: $[N_1, N_{\text{max}}]$, 把此宽带信号按频率点均匀分成 P 个子带信号: $[N_1, N_{\text{num}}], \dots, [N_{\text{num} \times (p-1)+1}, N_{\text{num} \times p}], \dots, [N_{\text{num} \times (p-1)+1}, N_{\text{max}}]$, num 为每个子带的离散频率点数, $\text{num} = (N_{\text{max}} - N_1) / P$ 。从上面介绍可以看出每个子带还是宽带信号, 只是带宽相对小一些。设第 p 个子带的频率范围为: $[N_{\text{num} \times (p-1)+1}, N_{\text{num} \times p}]$, 对于任意一频率点 k , 均有下式成立:

$$\mathbf{F}(k) = \mathbf{A}(k, \theta, \phi) \mathbf{S}(k) + \mathbf{N}(k) \quad k=1, 2, \dots, N/2 \quad (2)$$

定义频域相关矩阵 $\mathbf{R}(k)$ 为:

$$\mathbf{R}(k) = E\{\mathbf{F}(k) \times \mathbf{F}^H(k)\} \quad (3)$$

则接收信号的相关矩阵为:

$$\mathbf{R}(k) = \mathbf{A}(k, \theta, \phi) \mathbf{R}_{\text{ss}}(k) \mathbf{A}^H(k, \theta, \phi) + \sigma^2 \mathbf{I} \quad k=1, 2, \dots, N/2 \quad (4)$$

定义聚焦变换:

$$\mathbf{Q}(k) = \mathbf{T}(k) \times \mathbf{R}(k) \times \mathbf{T}^H(k) \quad (5)$$

并且使变换矩阵 $\mathbf{T}(k)$ 满足:

$$\begin{aligned} \min_{\mathbf{T}(k)} \|\mathbf{R}(k_{p_0}) - \mathbf{T}(k) \times \mathbf{R}(k) \times \mathbf{T}^H(k)\| \\ \text{s.t. } \mathbf{T}(k) \times \mathbf{T}^H(k) = \mathbf{I} \\ k = N_{\text{num} \times (p-1)+1}, N_{\text{num} \times (p-1)+2}, \dots, N_{\text{num} \times p} \end{aligned} \quad (6)$$

并求得 $\mathbf{T}(k)$ 的一个解为^[5]:

$$\mathbf{T}(k) = \mathbf{U}(k_{p_0}) \mathbf{U}^H(k) \quad (7)$$

式中 $\mathbf{U}(k_{p_0})$ 和 $\mathbf{U}(k)$ 分别为相关矩阵 $\mathbf{R}(k_{p_0})$ 和 $\mathbf{R}(k)$ 的特征向量矩阵; k_{p_0} 为选定的第 p 个子带的聚焦频率。聚焦后的相关矩阵为:

$$\mathbf{R}_{\text{focus}}^p = \sum_{k=N_{\text{num} \times (p-1)+1}}^{N_{\text{num} \times p}} \mathbf{Q}(k) = \sum_{k=N_{\text{num} \times (p-1)+1}}^{N_{\text{num} \times p}} \mathbf{T}(k) \times \mathbf{R}(k) \times \mathbf{T}^H(k) \quad (8)$$

利用经典的 DOA 估计算法如 MUSIC 算法可以求得每个子带的空间谱 $\mathbf{P}_{\text{music}}^p$, 然后平均这 p 个空间谱, 得到最终的空间谱矩阵 $\mathbf{P}_{\text{music}}$:

$$\mathbf{P}_{\text{music}} = \sum_{p=1}^P \mathbf{P}_{\text{music}}^p \quad (9)$$

二维思索空间谱矩阵 $\mathbf{P}_{\text{music}}$, 找出谱峰可以求得信号源的方位信息。

为了求得第 p 个子带的最佳聚焦频率 k_{p_0} , 设聚焦误差为 $\boldsymbol{\varepsilon}$, 把 $\mathbf{T}(k)$ 的解代入式(5)得:

$$\begin{aligned} \boldsymbol{\varepsilon} = \mathbf{R}(k_{p_0}) - \mathbf{U}(k_{p_0}) \mathbf{U}^H(k) \times \mathbf{R}(k) \times \mathbf{U}(k) \mathbf{U}^H(k_{p_0}) = \\ \mathbf{R}(k_{p_0}) - \mathbf{U}(k_{p_0}) \mathbf{A}(k) \times \mathbf{U}^H(k_{p_0}) \end{aligned} \quad (10)$$

对式(10)稍加变换得:

$$\boldsymbol{\varepsilon} = \mathbf{A}(k_{p_0}) - \mathbf{A}(k) \quad (11)$$

即最佳聚焦频率 k_{p_0} 的选择依据为:

$$\begin{aligned} \min_{k_{p_0}} \|\boldsymbol{\varepsilon}\| = \min_{k_{p_0}} \|\mathbf{A}(k_0) - \mathbf{A}(k)\| \\ k = N_{\text{num} \times (p-1)+1}, N_{\text{num} \times (p-1)+2}, \dots, N_{\text{num} \times p} \end{aligned} \quad (12)$$

式(12)描述了一个一维优化问题, 通过简单的一维搜索可以求出第 p 个子带的最优聚焦频率 k_{p_0} 。

从式(11)可以看出, 如果增加信号带宽, 聚焦误差也将相应增加。当子带数取 1 时, 即不划分子带时, 该算法退化为普通宽带聚焦算法, 聚焦误差增大, 估计性能下降, 但是算法的计算复杂度最小。

当子带个数增加, 每个子带的频率范围变窄, 聚焦误差减小, 估计性能提高, 但计算复杂度增加。当每个子带成为窄带信号时, 聚焦已经失去意义, 该算法退化为非聚焦类算法。

3 仿真和讨论

本文用计算机仿真的方法来验证前面的算法, 仿真使用MATLAB语言。在仿真中, 采用16 kHz、16 bit采样、512点FFT、256点汉明窗, 八个全向无差异麦克风均匀分布于XOY平面的半径为20 cm的圆周上。两个信号源均为中文标准普通话, 信号源1为男声, 方向为: $\theta = 20^\circ$, $\phi = 65^\circ$; 信号源2为女声, 方向: $\theta = 70^\circ$, $\phi = 25^\circ$ 。噪声为高斯白噪声, 信噪比分别为: 10 dB、0 dB、-5 dB。搜索步长为 1° , 水平角和俯仰角的搜索范围为 $0^\circ \sim 90^\circ$ 。当取十个子带时, 仿真结果如图1~图3所示。

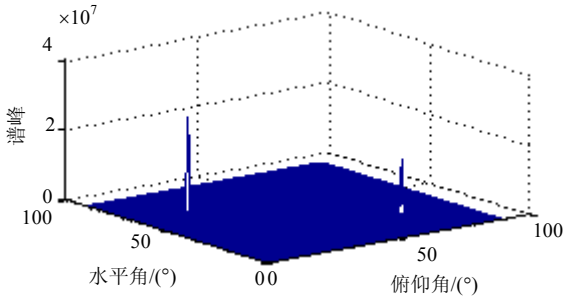


图1 信噪比为10 dB的仿真结果

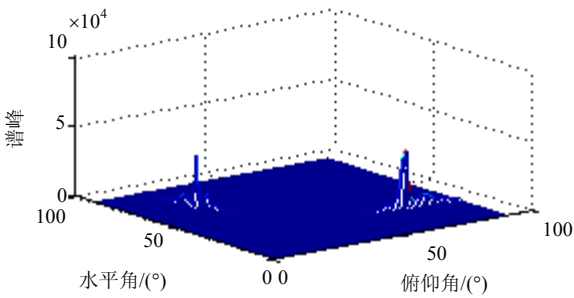


图2 信噪比为0 dB的仿真结果

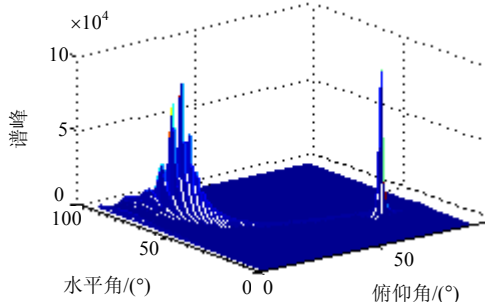


图3 信噪比为-5 dB的仿真结果

为了检验子带数的选取对算法的影响, 在信噪比为10 dB时, 取子带数分别为1、5、20进行了仿真, 其结果如图4~图6所示。当子带数较少时, 空间谱的峰值变小, 主要是因为子带数较小时, 聚焦误差增加, 导致噪声子空间的正交性和峰值下降, 但准确识别出信号的方位信息不变。

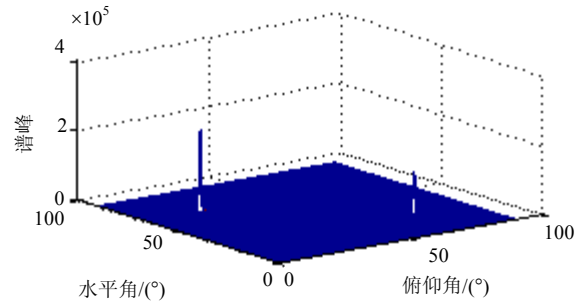


图4 子带数为1时仿真结果

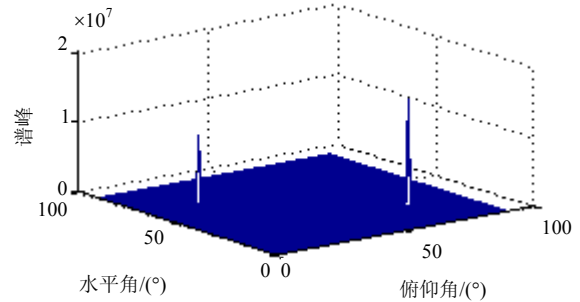


图5 子带数为5时仿真结果

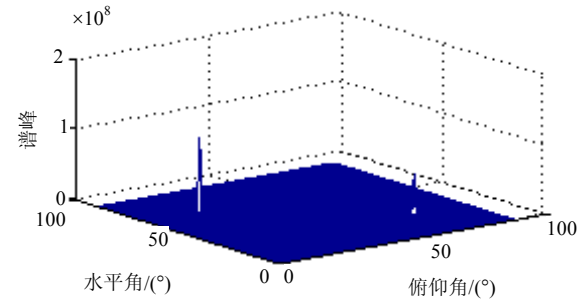


图6 子带数为20时仿真结果

在本文仿真条件下, 当信噪比为10 dB时, 文献[4]中提出的2D-MUSIC的仿真结果如图7所示。文献[5]中提出的TCT算法需要预估计信号的方向信息, 当预估计的方向为 $\theta = 20^\circ \pm 1^\circ$ 、 $\phi = 65^\circ \pm 1^\circ$ 和 $\theta = 70^\circ \pm 1^\circ$ 、 $\phi = 25^\circ \pm 1^\circ$ 时, 仿真结果如图8所示; 当预估计的方向为 $\theta = 20^\circ \pm 0.5^\circ$ 、 $\phi = 65^\circ \pm 0.5^\circ$ 和 $\theta = 70^\circ \pm 0.5^\circ$ 、 $\phi = 25^\circ \pm 0.5^\circ$ 时, 仿真结果如图9所示。从仿真结果可以看出, 本文提出的算法在计算复杂度、定位性能和实时性等方面明显优于文献[4]和文献[5]提出的算法。

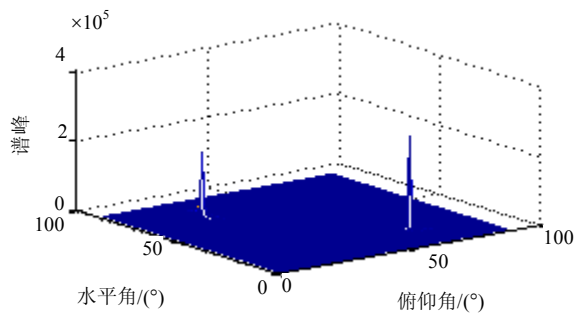


图7 2D-MUSIC算法仿真结果

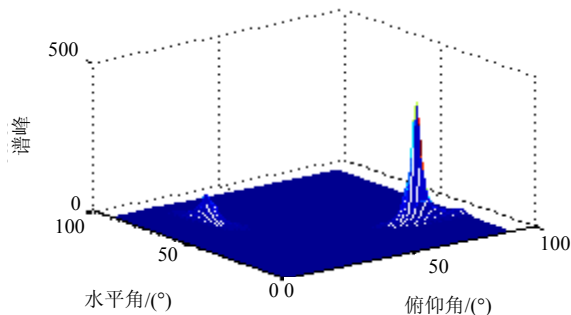


图8 TCT算法误差为1°时仿真结果

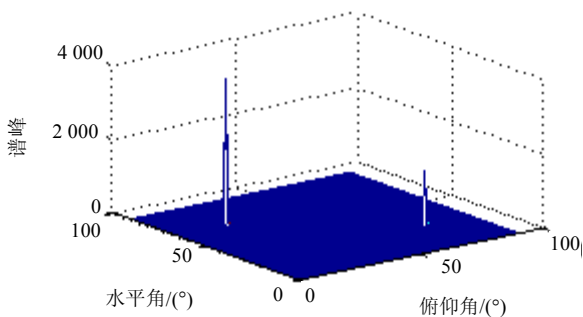


图9 TCT算法误差为0.5°时仿真结果

4 算法复杂度分析

从上面的讨论可以看出影响本算法计算复杂度的主要因素为子带数 P 的选取,当 P 取很大时,与2D-MUSIC算法复杂度相当。在实际系统中,该算法的计算量是2D-MUSIC算法的十分之一到百分之一不等(根据 P 的选择)。对于TCT算法,由于采取的预估计的算法和迭代次数的不确定,当直接输入信号的预估信息(不进行预估)以及不进行迭代时,TCT算法的复杂度小于本文提出的算法,但是当进行预

估和迭代时算法复杂度大于本文提出的算法,并且实时性大大降低。

5 结束语

在许多关于麦克风阵列的实际应用中,声源有可能位于近场,这将增加模型误差,使估计性能下降。同时该算法只能估计声源的方向,无法估计声源的距离。因此在计算复杂度不高的情况下,如何实现声源的三维位置参数估计,还需要进一步的研究。

参 考 文 献

- [1] JU Tai-liang, XU Yi-liang, PENG Qi-cong, Speech source localization in near field[C]//IEEE ICCAS'04, [S.l.]: IEEE Press, 2004, 2: 769-772.
- [2] ASANO F, ASOH H, MATSUI T, Sound source localization and separation in near field[J].IEICE Trans Fundamentals, 2000, E83-A(11): 2286-2294.
- [3] 居太亮, 彭启琮, 邵怀宗, 等. 基于任意麦克风阵列的声源二维DOA估计算法研究[J]. 通信学报, 2005, 26 (8): 129-133.
- [4] JU Tai-liang, PENG Qi-cong, SHAO Huai-zong, et al. Speech source localization sub-space algorithm research based on microphone array in near field[C]//ICEMI'2005. Beijing: Internal Academic Publisher/Beijing World Publishing Corporation, 2005, 3: 695-701.
- [5] VALEE S, KABAL P. Wideband array processing using two-sided correlation transformation [J]. IEEE Trans on SP, 1995, 43(1): 160-172.
- [6] HUNG H, KAVEH M. Focusing matrices for coherent signal subspace processing [J]. IEEE Trans on ASSP, 1988, 36(8):1272-1281.
- [7] DORON M A, WEISS A J. On focusing matrices for wideband array processing [J]. IEEE Trans on SP, 1992, 40(6): 1295-1302.
- [8] VALEE S, CHAMPAGNE B. Localization of wideband signals using least-squares and total least-squares approaches[J]. IEEE Trans on SP, 1999, 47(5): 1213-1222.

编辑 张俊