

MATLAB加窗滤波在 广义互相关时延估计中的应用

王森, 宫玉琳, 王春阳, 李明秋, 李健建

(长春理工大学 电子信息工程学院, 长春 130022)

摘要: 针对麦克风阵列广义互相关声源定位方法中, 麦克风接收到的人声信号掺杂混响以及噪声, 导致算法求解过程中准确性不高的缺点, 通过MATLAB分窗滤波的方法, 在算法中对实时接收的声音信号进行预处理, 针对人声频率进行改进, 提升低信噪比状况下估计时延的准确性, 完成人声声源定位的目的, 并通过仿真实验验证了该方法的有效性。

关键词: 声源定位; 分窗滤波; 广义互相关

中图分类号: TN912.3

文献标识码: A

文章编号: 1672-9870(2016)06-0048-05

Application of MATLAB Windowing and Filtering in Generalized Cross-correlation Time Delay Estimation

WANG Sen, GONG Yulin, WANG Chunyang, LI Mingqiu, LI Jianjian

(School of Electronics and Information Engineering, Changchun University of Science and Technology, Changchun 130022)

Abstract: Aiming at the shortcomings of the microphone array signal de-noising and the reverberation of the human voice signal received by the microphone in the generalized cross-correlation sound source localization method, this paper uses MATLAB to divide the windowing and filtering method, And the accuracy of estimating the time delay under the condition of low SNR is improved and to complete the task of estimating the localization of the human voice source. Simulation experiments have been carried out to verify the effectiveness of the proposed method.

Key words: sound source localization; windowing and filtering; generalized cross correlation

麦克风阵列对语音定位的方法, 常常应用在智能机器人、通信、火警救援等众多领域^[1], 关于定位算法, 当今使用较多的有高分辨率谱估计的定位方法, 基于可控波束形成器的定位方法和基于到达时间差(TDOA)的定位方法等。

几种方法各有优劣:(1)基于高分辨率谱估计的定位方法是先求取麦克风阵列采集信号间的相关矩阵的空间谱, 然后通过求取相关矩阵, 确定出声源的方向角。由于它要求信号是平稳的, 而且其不适合做近场声源定位, 导致定位效果下降;(2)基于可控波束形成器的方法是将 M 个麦克风收到的信号进行加权、延迟后相加, 然后通过角度扫描, 计算相加后的结果, 不同角度时结果不同, 当角度和目标方向一致时, 会使相加和达到最大值。不同的加权方法,

可以得到不同的波束形成方法, 但该方法需要声源和环境噪声的先验知识, 这是其重大的缺陷;(3)基于到达时间差(TDOA)的定位方法是通过计算出两路信号的到达时间差然后通过算法计算出语音的角度和距离^[2]。该方法计算量小, 定位精度高, 不需要太多的先验知识, 因此应用广泛。但是在实际应用中由于环境噪声、房间混响的原因, 麦克风收集的声音信号受到了很大的干扰, 尤其在复杂的环境中, 定位的精确度受到了很大的影响。

综合比较并结合实际, 本文使用基于到达时间差(TDOA)的广义互相关定位方法, 针对其缺点, 使用MATLAB分窗滤波方法对算法效果进行了改进, 并针对实际声音进行仿真实验, 结果表明, 该方法对去除噪声和定位的性能有很大的提高。

收稿日期: 2016-09-23

作者简介: 王森(1991-), 男, 硕士研究生, E-mail: wangsen321@126.com

通讯作者: 宫玉琳(1983-), 男, 博士, 讲师, E-mail: garrygong1983@126.com

1 基于到达时间差(TDOA)的时延估计方法

TDOA定位方法研究历史已经很长,在实际中得到了广泛的应用,该方法主要有两个步骤,第一步为时延估计,用该方法计算被测声音信号到达不同麦克风的时间差;第二步为位置估计,利用第一步中得到的值结合麦克风阵列中麦克风具体的几何位置关系来确定声源的位置,步骤表示如图1所示。

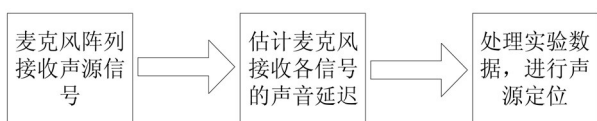


图1 TDOA实验估计步骤

时延估计方法有许多种,这些方法各有优劣,针对实际中不同的状况有着不同的定位效果。本论文针对实际应用中噪声干扰强,混响严重的情况下的人声定位,使用了广义互相关的方法,流程如图2所示。

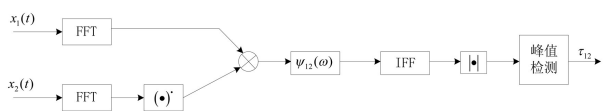


图2 广义互相关法进行时延估计框图

图中 (\cdot) 为共轭运算; $\psi_{12}(\omega)$ 为广义互相关方法的加权函数;FFT为傅里叶变换;IFF为反傅里叶变换。

1.1 基本互相关方法

假设两个麦克风 M_1 、 M_2 ,其接收到的声音信号的模型为式(1):

$$\begin{aligned} x_1(t) &= s(t - \tau_1) + n_1(t) \\ x_2(t) &= s(t - \tau_2) + n_2(t) \end{aligned} \quad (1)$$

其中,接收到的信号用 $s(t)$ 表示,两麦克风的高斯白噪声用 $n(t)$ 表示,声音信号到达麦克风的时间用 τ 来表示,这样, $\tau_{12} = \tau_1 - \tau_2$ 就是声音信号的到达时间差值,这里我们假设模型中 $s(t)$, $n_1(t)$, $n_2(t)$ 是互不相关的平稳随机变化过程,则 $x_1(t)$, $x_2(t)$ 的互相关函数可以表示为式(2):

$$R_{12}(\tau) = E[x_1(t)x_2(t - \tau)] \quad (2)$$

将两信号模型带入上式中,得式(3):

$$\begin{aligned} R_{12}(\tau) &= E[s(t - \tau_1)s(t - \tau_1 - \tau)] + \\ &E[s(t - \tau_1)n_2(t - \tau)] + E[s(t - \tau_2 - \tau)n_1(t)] \\ &+ E[n_1(t)n_2(t - \tau)] \end{aligned} \quad (3)$$

前面已经假设说明, $s(t)$, $n_1(t)$, $n_2(t)$ 之间两两

互不相关,简化上式得到式(4):

$$R_{12}(\tau) = E[s(t - \tau_1)s(t - \tau_1 - \tau)] = R_{ss}(\tau - (\tau_1 - \tau_2)) \quad (4)$$

由自相关的性质得知:当 $\tau - (\tau_1 - \tau_2) = 0$ 时, $R_{12}(\tau)$ 达到最大值,这时,最大值对应的 τ 值就是两麦克风的时间延时值。所以, M_1 、 M_2 两麦克风之间的到达时间延迟可以表示为式(5):

$$\hat{\tau} = \arg \max_{\tau} R_{12}(\tau) \quad (5)$$

式中, $\tau \in [-\tau_{\max}, \tau_{\max}]$, τ_{\max} 可能为最大的时延值。

由此,两个不同位置的麦克风之间的时延值就可以通过求两个信号的互相关函数 $R_{12}(\tau)$ 的极大值的方法来获得。当互相关函数最大峰值比较尖锐明显时就可以准确地求出极大点所在的位置,最大峰值越尖锐,越明显,那么计算出的时延值就越精确。

基本互相关方法是通过互相关函数的峰值来得到延迟时间差值,这种方法计算量小,而且容易实现,但是该方法要求噪声和噪声之间都是两两不相关的函数,这个前提在实际应用中并不一定能够实现。并且,该互相关函数在数学角度上可看成严格的统计平均。当条件为平均遍历时可以用无穷时间平均值来替代严格统计平均,而在实际应用中,一般是用有限定的时间平均值来代替严格统计平均,该方法不可以忽略噪声对相关函数的影响,会导致互相关函数峰值变得不明显,最终降低延迟估计的精度。为降低噪声的影响,产生了广义互相关算法。

1.2 广义互相关法

广义互相关法首先求出两信号间的互功率谱,并在频域进行加权。可以白化处理信号和噪音,同时增强信噪比的频率比重,进而抑制噪音的功率,再反变换到时域,得到广义互相关函数,使互相关函数的峰值更加尖锐,提高准确性。具体流程如下:

对式(1)做傅里叶变换,得到 $x_1(t)$ 和 $x_2(t)$ 信号的互功率谱为:

$$G_{x_1 x_2}(\omega) = G_{ss}(\omega) e^{-j\omega \tau_{12}} \quad (6)$$

傅里叶变换后对上式进行加权,加权后反傅里叶变换到时域^[3],如式(7):

$$R_{x_1 x_2}^g(\tau) = \int_{-\infty}^{\infty} G_{x_1 x_2}(\omega) \varphi_{12}(\omega) e^{j\omega \tau} d\omega \quad (7)$$

φ_{12} 是广义加权函数,随实际状况中噪声和混响的不同而有不同的选择,使 $R_{12}(\tau)$ 具有明显的峰值。表1是几种常见的加权函数。

表1 几种常见的加权函数

名称	加权函数 $\psi_{12}(\omega)$
基本互相关(CC)	$\psi_{12}(\omega) = 1$
PHAT	$\psi_{12}(\omega) = \frac{1}{ G_{x_1 x_2}(\omega) }$
Roth处理器	$\psi_{12}(\omega) = \frac{1}{G_{x_1 x_2}(\omega)}$
平滑相干变换(SCOT)	$\psi_{12}(\omega) = \frac{1}{\sqrt{G_{x_1 x_1}(\omega) G_{x_2 x_2}(\omega)}}$
最大似然加权(ML)	$\psi_{12}(\omega) = \frac{ \gamma(\omega) ^2}{ G_{x_1 x_2}(\omega) (1 - \gamma(\omega) ^2)}$
WP加权	$\psi_{12}(\omega) = \frac{ G_{x_1 x_2} ^2}{G_{x_1 x_1}(\omega) G_{x_2 x_2}(\omega)}$

针对这些加权函数,选用了前面三种加权方法,对其定位的效果做了仿真实验,声音的波形,是以横坐标轴为对称轴上下波动的,所以在MATLAB中使用余弦波形模拟了一定长度的声音信号并对声音信号加入了随机白噪声,然后对定位的效果做了仿真,在仿真过程中设置延迟时间为0.01秒,仿真结果如图3~5所示:

(1) PHAT-GCC方法

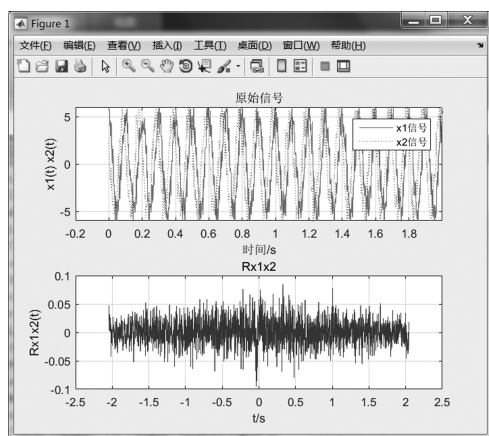


图3 PHAT-GCC方法仿真

(2) ROTH-GCC方法

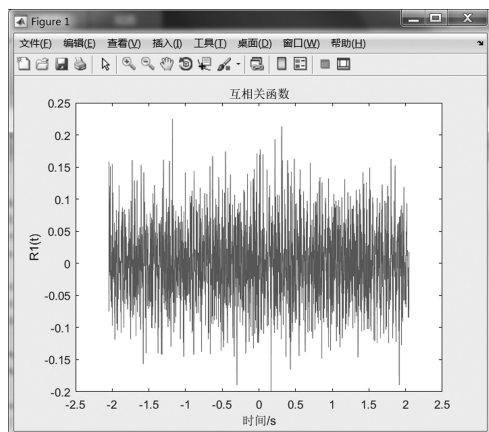


图4 ROTH-GCC方法仿真

(3) SCOT-GCC方法

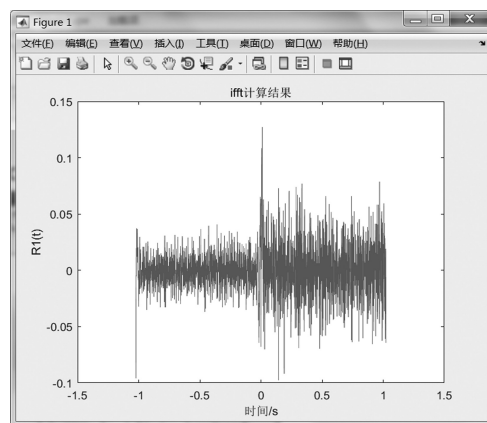


图5 SCOT-GCC方法仿真

在仿真实验中发现,PHAT-GCC方法估计的延迟为0.0060s,ROTH-GCC方法估计的延迟为0.0360s,SCOT-GCC方法估计的延迟为0.0110s。

可见,这几种方法在加入了噪声信号之后,声音延迟的时间跟加入的延迟时间0.01s相比,均有很大的偏差,会导致定位不准确,并且随噪声的情况计算精度也不同,甚至出现了错误的声音延迟计算。为此对声音信号处理如下。

2 MATLAB分帧加窗

声音是随时间变化的,一般在10ms到30ms内是被视为短时平稳的。为了减少频谱能量泄露,可以采用不同的截取函数对信号进行截断,该函数就称为窗函数。分帧加窗处理可以实现语音平稳化,一般来说,每秒钟声音信号可以分成33~100帧。分帧的一般方法是连续交叠分段,交叠部分在1/3~1/2之间时,取得的语音分析效果最佳。在这个过程中主要考虑三方面:帧长、交叠长度、加窗函数选择^[4]。

声音在分帧加窗处理之后,就可以对其进行傅里叶变换研究了。对几种广义互相关不同的加权方法进行了仿真实验,发现分帧加窗可以对声音定位的效果有很大的提升。

本文采用了汉宁窗的加窗方法,汉宁窗又称升余弦窗,汉宁窗可以看作是3个矩形时间窗的频谱之和,其使旁瓣互相抵消,消去高频干扰和漏能。汉宁窗可以使主瓣加宽并降低,旁瓣则显著减小。其函数为:

$$\omega(n) = 0.5 \left[1 - \cos \left(\frac{2\pi n}{M+1} \right) \right], \quad 1 \leq n \leq M \quad (8)$$

在以上三种加权方法的定位算法中加入了汉宁窗函数,效果如图6所示:

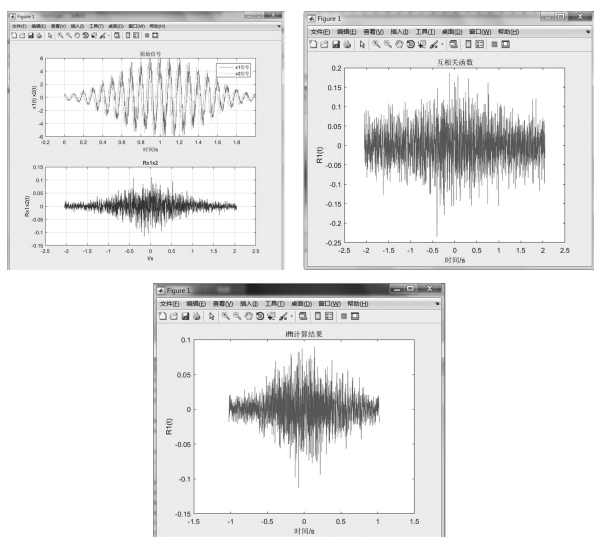


图6 三种加权方法在加窗后的互相关效果

3 MATLAB滤波

在广义互相关的计算研究中发现,实际环境中的噪声和混响严重影响了互相关的计算,得到的延迟误差仍然较大。对于这个问题,本文考虑针对人声的频率20Hz~20KHz,使用MATLAB软件进行滤波处理,去除其它非人声频率的无用声音,从而减小噪声对互相关方法定位结果的影响。

为了便于观察效果,采集了一段男性声波信号,并以此来验证滤波的效果。原始信号如图7所示:

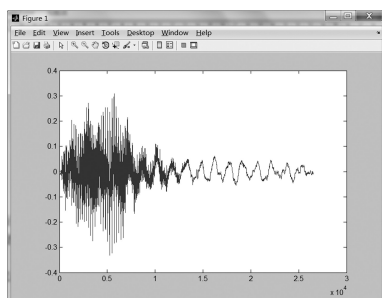


图7 实际采集的男性声音信号

两个麦克风接收到的信号如下:

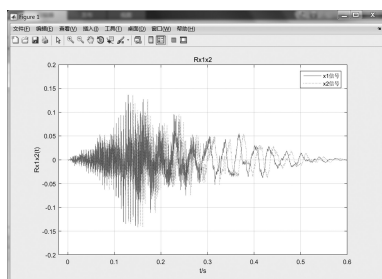


图8 两麦克风接收信号

MATLAB可以简便地构造模拟滤波器函数,常见的方法有贝塞尔(Bessel)模拟滤波器、巴特沃

思(Butterworth)模拟/数字滤波器、切比雪夫(Chebyshev)I型模拟/数字滤波器等。

利用MATLAB的构造滤波器函数的方法,构造了一个贝塞尔(Bessel)模拟带通滤波器,并且设置了其带通频率为20Hz~20KHz,然后将其应用在广义互相关定位的计算中。

以下是滤波之前与滤波之后的仿真比较:

(1)PHAT-GCC方法

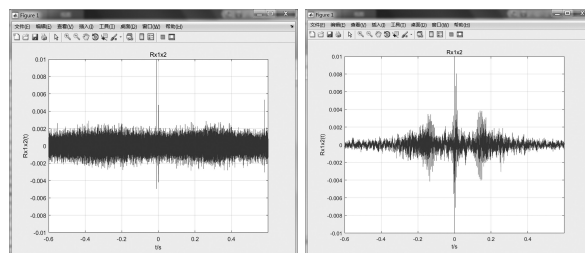


图9 PHAT-GCC滤波前后效果图

(2)ROTH-GCC方法

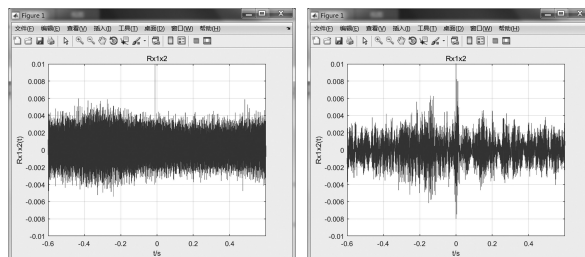


图10 ROTH-GCC滤波前后效果图

(3)SCOT-GCC方法

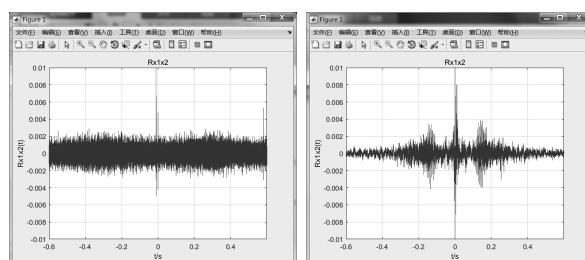


图11 SCOT-GCC滤波前后效果图

可以很明显地看出,经过加窗滤波后的声音信号在进行广义互相关的声源定位计算时,其效果比直接进行计算要准确的多,得到的互相关波形中,许多不必要的噪声干扰已经消除,并且其波峰变得突出明显,声源定位的效果得到了很大的提高,该结果显示,加窗滤波在复杂环境中的定位计算中,有很高的应用价值。

4 结论

本文简述了广义互相关的定位方法,并且对互相关定位的计算中加入了分窗、滤波的方法,针对人的声音信号,对声音中掺杂的噪音进行了去除。通过加窗、滤波后的广义互相关波形与未做任何处理的广义互相关波形的比较,表明该方法对去除噪声、平稳声音信号、声源定位计算有良好的作用。通过该方法,声源定位结果的精度有很大提高。接下来我们将会对人声信号中噪声的去除做进一步的研究。

参考文献

- [1] 夏阳,张元元.基于矩形麦克风阵列的改进的GCC-PHAT语音定位算法[J].山东科学,2011,12(24):76-79.
- [2] 宋志亚.基于麦克风阵列的室内声源定位方法研究[D].重庆:重庆大学,2012.
- [3] 梁宇,马良,纳霞,等.基于广义互相关算法的时延估计[J].计算机科学,2011,10(10A):454-456.
- [4] 段丽萍.基于TDOA算法的四麦克风阵列三维声源定位研究[D].兰州:兰州理工大学,2014.
- [7] 姜茹,艾勇,单欣,等.空间光通信中GPS辅助快速捕获对准技术[J].光通信技术,2015,2(4):47-49.
- [8] 白帅.空间二维光电转台的高稳定捕获跟踪技术研究[D].北京:中国科学院大学,2015.
- [9] 薛正燕,开波,任戈.星间光通信中快速捕获技术研究[J].光通信技术,2015,3(6):53-56.
- [10] 薛正燕.卫星光通信捕获跟踪技术研究[D].北京:中国科学院大学,2015.
- [11] 赵勇.陆军激光通信系统全光捕获技术实验研究[D].长春:长春理工大学,2012.
- [12] 朱博韬.卫星光通信双向扫描捕获技术研究[D].哈尔滨:哈尔滨工业大学,2007.
- [13] 单风华,佟首峰,吕春雷.自由空间光通信APT系统信标探测技术[J].长春理工大学学报:自然科学版,2013,36(3):53-59.
- [8] Zhang L, Qiao S, Jin Y, et al. Magnetic hollow spheres of periodic mesoporous organosilica and Fe_3O_4 nanocrystals: fabrication and structure control [J]. Advanced Materials, 2008, 20(4): 805-810.
- [9] Suzuki N, Gupta P, Sukegawa H, et al. Aerosol-assisted synthesis of thiol-functionalized mesoporous silica spheres with Fe_3O_4 nanoparticles [J]. Journal of Nanoscience & Nanotechnology, 2010, 10(10): 6612-7.
- [10] Yavuz C T, Mayo J T, Yu W W, et al. Low-field magnetic separation of monodisperse Fe_3O_4 nanocrystals [J]. Science, 2006, 314(5801): 964-7.
- [11] 李冬梅,徐光亮.制备超顺磁性 Fe_3O_4 纳米粒子的研究进展[J].中国粉体技术,2008,14(4):55-58.
- [12] 宋新峰,孙汉文,吴静,等.超顺磁性 Fe_3O_4 纳米粒子化学合成及生物医学应用进展[J].应用化工,2015(4): 711-715.
- [13] Sun Y, Zhang Z, Wong C P. Study on mono-dispersed nano-size silica by surface modification for underfill applications [J]. Journal of Colloid & Interface Science, 2005, 292(2): 436-44.
- [14] Rahman I A, Padavettan V. Synthesis of silica nanoparticles by sol-gel: size-dependent properties, surface modification, and applications in silica-polymer nanocomposites—a review [J]. Journal of Nanomaterials, 2012, 2012(1687-4110): 2817-2827.

(上接第19页)

(上接第56页)