

# 公共场所异常声源定位中时延估计方法研究<sup>\*</sup>

李伟红, 汤海兵, 龚卫国

(重庆大学光电技术及系统教育部重点实验室 重庆 400044)

**摘要:** HB(Hassab-Boucher)加权广义互相关(generalized cross correlation based on HB weighted function, GCC-HB)是常用的时延估计方法,在环境为弱高斯噪声情况下,可获得较为精确的时延估计值用于声源定位。通过分析认为,通常公共场所异常声音是一种短时信号,背景噪声主要为粉红噪声与脉冲噪声,符合分数低阶 $\alpha$ 稳定分布(fractional lower order alpha-stable, FLOA)。在此背景噪声的低信噪比环境下, GCC-HB方法的时延估计性能急剧下降。为此,提出基于反正切变换的改进GCC-HB的时延估计方法(improved GCC-HB method based on arc tangent transform, ATAN-IHB)。该方法首先对加噪信号采用反正切变换抑制噪声中尖峰脉冲的影响,然后结合每帧的信噪比对HB加权函数进行改变,并由多帧HB加权后的峰值确定出时延估计值。理论分析和计算机仿真结果表明,所提出的方法即使在低信噪比的环境下,也可以获得比较满意的时延估计值,具有一定的实用性价值。

**关键词:** 声源定位;脉冲噪声;时延估计;广义互相关;HB加权

**中图分类号:** TP912.16 **文献标识码:** A **国家标准学科分类代码:** 510.40

## Time delay estimation method of abnormal sound source localization in public places

Li Weihong, Tang Haibing, Gong Weiguo

(Laboratory of Optoelectronic Technology and Systems of the Education Ministry of China,  
Chongqing University, Chongqing 400044, China)

**Abstract:** Generalized cross correlation based on Hassab-Boucher(HB) weighted function(GCC-HB) is a well known time delay estimation method. Relatively precise time delay estimation can be achieved using this method in the environment of weak Gaussian noise. This paper considers that abnormal sound signals in public places usually are short-time signals; the background noise is mainly pink noise and impulsive noise, and follows the fractional lower order alpha-stable(FLOA) distribution. The performance of GCC-HB degrades sharply under background noise with low signal to noise ratio(SNR). So, to improve the performance of GCC-HB, this paper proposes a modified GCC-HB method based on arc tangent transform(ATAN-IHB). Firstly, the proposed method makes use of arc tangent transform to pre-process the abnormal sound signal with noise; and then, the HB weighted function is changed according to the SNR of each frame; at last, the time delay estimation is determined from the peak of several frame HB weighted functions. The results of theoretical analysis and computer simulation indicate that the proposed method can get satisfactory value of time delay estimation even in the environment of low SNR with the noise of FLOA, and can be applied in practice.

**Key words:** sound source localization; impulsive noise; time delay estimation; generalized cross correlation; HB(Hassab-Boucher) weighted

收稿日期:2011-03 Received Date:2011-03

<sup>\*</sup> 基金项目:重庆市科技攻关重点项目(CSTC,2009AB0175)、中央高校基本科研业务费(CDJXS10122218, CDJXS11122216)、重庆市科委自然科学基金(CSTC,2010BB2230)资助项目

## 1 引言

公共场所的安全防范是国家政治和社会稳定、人民群众安居乐业的重要保证。目前,视频监控已成为获取异常事件发生的各类信息的主要手段。由于大多数公共场所发生的异常事件一般都伴随有异常声音的产生,如枪声、爆炸声、玻璃破碎声、尖叫声等,异常声音的产生表明异常事件已发生或预示异常事件将发生。为此,对异常声源分析及定位技术的研究具有重要的社会意义及学术价值。

基于传声器阵列的声源定位问题一直是非常热门的研究课题。目前的声源定位研究多数致力于解决视频会议、车载免提电话、语音识别和助听设备等中的声源定位问题<sup>[1-2]</sup>。在真实的声场环境中,由于混响与噪声的存在,现有的声源定位方法不同程度存在一定的局限性。目前基于传声器阵列的声源定位技术主要分为:基于最大输出功率的可控波束形成定位方法<sup>[3]</sup>、基于高分辨率谱估计的定位方法<sup>[4]</sup>和基于时延估计(time delay estimation, TDE)的定位方法<sup>[5]</sup>。其中,基于TDE的定位方法计算量较小,易于实时实现,已成为声源定位领域很有特色的研究方向<sup>[6-9]</sup>。所谓TDE是指利用一定的算法获得传声器阵列中不同传声器接收到的同源信号之间由于传输距离不同而引起的时间差的估计值。该方法将声源定位分为2步:第1步,估计出声源到每对麦克风之间的时延差(time delay of arrival, TDOA),第2步,根据这些时延差和传声器阵列的几何位置估计出声源的方位。典型的时延估计方法有广义互相关(generalized cross correlation, GCC)时延估计方法<sup>[10]</sup>、自适应滤波最小均方(least mean square, LMS)时延估计方法<sup>[11]</sup>等。

CCC是一种基本的、常用的时延估计方法,仅需要1帧声音数据(如采用512个采样点为1帧的长度)就可以作出一个TDOA估计。但在实际场所中,由于脉冲噪声以及混响的影响,CCC方法的时延估计性能会急剧衰减。为此,出现了对互相关函数进行加权、锐化互相关函数峰值的各种改进方法。根据其加权形式和准则的不同,有ROTH(roth filter)、SCOT(smoothed coherence transform)、ECKART(eckart filter)、ML(maximum likelihood)、PHAT(phase transform)、WP(Wiener filters)和HB(Hassab-Boucher)等广义相关时延估计方法<sup>[5]</sup>。在这些加权方法中,ECKART、ML、WP和HB加权的广义相关时间延迟估计可以达到克拉美罗界(CRLB)。其中HB加权的广义相关时延方法(GCC-HB)在提高TDE精度及在未知信号噪声先验知识前提下,动态跟踪时延变化等方面表现出较好的性能<sup>[12]</sup>。

所谓异常声音是指某种特定的正常环境下所不应该

发生的声音,公共场所的异常声音通常包括枪声、爆炸声、玻璃破碎声、尖叫声等<sup>[13]</sup>。公共场所异常声音的基频一般都在中低频段,时间长度在350~2500 ms,其中大部分异常声音长度低于1000 ms<sup>[13]</sup>。而在广场、车站、码头、校园操场等公共场合,常出现汽车鸣笛声、咳嗽声、脚步声以及低频大气噪声等干扰。本文分析认为,这类噪声符合分数低阶 $\alpha$ 稳定分布(fractional lower order alpha-stable distribution, FLOA)。由于FLOA分布只有不超过 $\alpha(0 < \alpha \leq 2)$ 的阶矩,这使GCC-HB方法的相关函数峰值会出现伪峰,有时还会呈现出类似异常声音的特征,甚至难以得到正确的时延估计结果<sup>[14-15]</sup>。为此,本文提出一种基于反正切变换的改进GCC-HB的时延估计方法(ATAN-IHB),以增强GCC-HB方法处理公共场所异常声源定位的抗噪能力。该方法首先对接收到的2路声源信号进行反正切变换,抑制信号噪声中尖峰脉冲的影响。并将原有固定的HB加权函数改进为随信噪比而变的加权函数;并将原来由1帧GCC-HB的峰值决定的TDE值,改进为由多帧GCC-HB加权后的峰值来获得TDE值,从而进一步锐化相关函数的峰值,提高算法的抗噪能力。实验结果显示,本文提出的ATAN-IHB改进方法可以很好地提高GCC-HB方法的性能,具有一定的实用价值。

## 2 公共场所信号噪声模型

### 2.1 加噪声源信号模型

传声器接收到的声音信号会受到噪声的干扰,如图1所示。

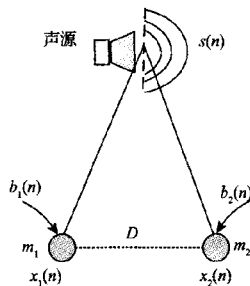


图1 传声器接收加噪声源信号  
Fig. 1 Microphones receive sound source signal with noise

间隔为 $D$ 的2个传声器 $m_1$ 和 $m_2$ 接收到的信号 $x_1(n)$ 、 $x_2(n)$ 可以表示为:

$$x_1(n) = a_1 s(n) + b_1(n) \quad (1)$$

$$x_2(n) = a_2 s(n - \tau_{12}) + b_2(n) \quad (2)$$

式中: $s(n)$ 为声源信号, $a_1$ 、 $a_2$ 是声波传播衰减因子( $a_i < 1, i = 1, 2$ ),即由多路径反射噪声引起的声波在传

输过程中的信号损失。 $\tau_{12}$ 是传声器 $m_2$ 相对于 $m_1$ 接收信号的时间延迟, $b_1(n)$ 、 $b_2(n)$ 为环境噪声(加性噪声),且 $s(n)$ 、 $b_1(n)$ 和 $b_2(n)$ 之间彼此不相关。 $\tau_{12} = \tau_1 - \tau_2$ 为2个传声器接收信号之间的时延。

## 2.2 公共场所噪声模型

在广场、车站、校园操场等公共场所环境下的噪声以粉红噪声为主,往往存在一些大气噪声、人为噪声如汽车鸣笛声、脚步声等脉冲噪声的干扰。这类噪声具有非常显著的脉冲特性,并且其统计密度函数具有长拖尾<sup>[14]</sup>。实际上,较多的脉冲状噪声和长拖尾正是分数低阶 $\alpha$ 稳定分布过程的显著特性。

$\alpha$ 稳定分布模型是一类适用范围很宽并得到广泛应用的随机信号模型,包括高斯分布和非高斯分布2种情况。 $\alpha$ 稳定分布与其他统计模型不同,除少数几个特定情况外, $\alpha$ 稳定分布的概率密度函数没有统一的解析表达式。通常用特征函数和4个参数( $\alpha$ 、 $\beta$ 、 $\gamma$ 、 $\mu$ )来描述:

$$\begin{cases} \varphi(t) = \exp\{j\mu t - \gamma|t|^\alpha[1 + j\beta \operatorname{sgn}(t)\omega(t, \alpha)]\} \\ \omega(t, \alpha) = \begin{cases} \tan \frac{\alpha\pi}{2}, & \alpha \neq 1 \\ \frac{\pi}{2} \log|t|, & \alpha = 1 \end{cases} \end{cases} \quad (3)$$

式中: $\operatorname{sgn}(\cdot)$ 为符号函数, $\alpha$ 为特征指数( $0 < \alpha \leq 2$ ), $\alpha$ 控制着概率密度函数的拖尾情况, $\alpha$ 越小拖尾越厚,脉冲性越强; $\beta$ 为对称系数, $\beta = 0$ 时表示对称分布,记为SaS分布。 $\gamma$ 为分散系数,类似于高斯分布的方差。 $\mu$ 为位置参数( $-\infty < \mu < +\infty$ ),对应于均值或中值。

图2显示了操场环境下的背景噪声与 $\alpha$ 值不同时的SaS分布过程的离散实现。

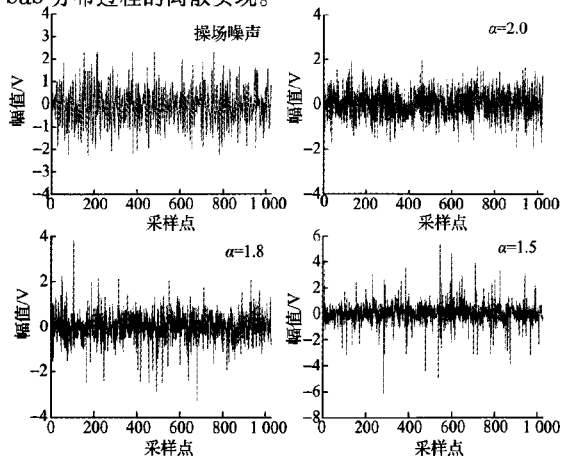


图2 操场背景噪声与 $\alpha = 2.0$ 、1.8、1.5时的SaS分布过程的离散实现

Fig. 2 Noise waveform under playground condition, and the discrete realization of SaS when  $\alpha = 2.0$ , 1.8 and 1.5

当 $\alpha = 2$ 时, $\alpha$ 稳定分布的特征函数与高斯分布的特征函数完全相同,这表明高斯分布是 $\alpha$ 稳定分布当 $\alpha = 2$ 时的特例。 $0 < \alpha < 2$ 时的SaS分布保持了高斯分布的一些特性,但又有明显不同。其显著特点是在统计分布上有长拖尾,即远离均值或中值的样本数较多,从而造成了其时间域波形上较多的尖峰脉冲。本文定义 $0 < \alpha < 2$ 的 $\alpha$ 稳定分布为分数低阶 $\alpha$ 稳定分布以区别 $\alpha = 2$ 的高斯分布。公共场所噪声分布应服从SaS分布,不同场所的环境噪声在时间域波形上的尖峰脉冲出现的频率不同,在统计分布上拖尾的厚度就不同, $\alpha$ 值的大小也就不同。由于分数低阶 $\alpha$ 稳定分布过程没有稳定有限的二阶统计量<sup>[14-15]</sup>,因此,基于二阶统计量和高斯噪声假定所设计的信号处理器在分数低阶 $\alpha$ 稳定分布噪声条件下,其性能显著退化。

## 3 HB加权广义互相关时延估计法

广义互相关(GCC)法的基本原理如下:根据2.1节的信号模型,2个传声器( $m_1$ 和 $m_2$ )接收到的信号的相关函数可表示为:

$$R_{x_1x_2}(\tau) = E[x_1(n)x_2(n-\tau)] \approx a_1a_2R_{ss}(\tau-\tau_{12}) + R_{b_1b_2}(\tau) \quad (4)$$

式中: $E[\cdot]$ 为数学期望, $R_{ss}$ 为信号源的自相关函数, $R_{b_1b_2}(\tau)$ 为两个传声器加性噪声的互相关函数。当 $\tau_{12} = \tau_1 - \tau_2$ 时 $R_{x_1x_2}(\tau)$ 取最大值。即 $R_{x_1x_2}(\tau)$ 最大值对应的 $\tau_{12}$ 为两个传声器之间的时延。

当信噪比足够大时,TDOA可以根据两个传声器接收到信号的互相关函数的峰值位置来获得。值得注意的是上述时延估计方法都是基于二阶统计量的,当环境噪声为高斯噪声时才具有较好的性能。但是,实际环境中存在背景噪声和混响的影响。其背景噪声不具有稳定的有限二阶统计量时,互相关运算就不能正常工作。并且互相关函数包含了直达波与反射波的峰值,这些峰值称为伪峰,伪峰会造成TDOA检测困难。为了解决这一问题,GCC方法在频域利用加权函数来加强信源信号直达部分,抑制噪声及混响信号,从而突出相应的峰值。

两个传声器接收到信号互功率谱密度为:

$$\Phi_{x_1x_2}(\omega) = a_1a_2\Phi_{ss}(\omega)e^{-j\omega\tau_{12}} + \Phi_{b_1b_2}(\omega) \quad (5)$$

式中: $\Phi_{ss}(\omega)$ 为信号源的自功率谱密度, $\Phi_{b_1b_2}(\omega)$ 为两个传声器加性噪声的互功率谱密度。

广义互相关函数定义为:

$$R_{GCC}(\tau) = \int_{-\infty}^{+\infty} H_{12}(\omega) \Phi_{x_1x_2}(\omega) e^{-j\omega\tau} d\omega \quad (6)$$

式中: $H_{12}(\omega)$ 为加权函数。Knapp在文献[5]中总结了常用的6种加权函数,并对各种加权函数的性能做了相应的比较。并在期望信号峰与输出噪声之比为最大的准

则下,Hassab 和 Boucher 导出了 HB 加权函数<sup>[12]</sup>。HB 加权的广义相关时间延迟估计可以达到克拉美罗界(CRLB),其中 HB 加权在实际应用中具有较好的性能。GCC-HB 方法时延估计流程图如图 3 所示,其中 $(\cdot)$ 表示共轭运算, $\otimes$ 表示相关运算。

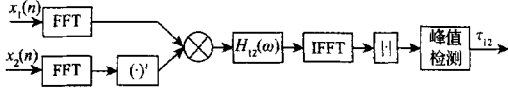


图3 GCC-HB 方法时延估计流程框图

Fig. 3 Framework for time delay estimation of GCC-HB

其中,HB 加权函数  $H_{12}(\omega)$  为:

$$H_{12}(\omega) = \frac{\Phi_{ss}(\omega)}{\Phi_{x_1x_1}(\omega)\Phi_{x_2x_2}(\omega)} \quad (7)$$

式中:  $\Phi_{x_1x_1}(\omega)$  为传声器  $m_1$  接收信号的自功率谱密度,  $\Phi_{x_2x_2}(\omega)$  为传声器  $m_2$  接收信号的自功率谱密度。并且,在信号噪声相互统计独立的假定下,有:

$$|\Phi_{x_1x_2}(\omega)| = \Phi_{ss}(\omega) \quad (8)$$

由式(6)的峰值点可以得到相应的时延估计值为:

$$\hat{\tau}_{12} = \arg \max_{\tau} [R_{GCC}(\tau)] \quad (9)$$

HB 加权对互功率谱密度起到 2 个作用:一是对互功率谱密度乘上一个源信号功率谱密度,这相当于加重了输入信号功率谱密度中的源信号成分;二是对互功率谱密度除以两输入信号的自功率谱密度,这实质上是对 Roth 处理器仅对一个输入信号进行归一化处理的一种改进,与 ROTH 加权类似,起到对输入信号白化的作用。

HB 加权时间延迟估计具有较强的抑制噪声的能力,但是这种方法的正确估计却依赖于对信号和噪声统计知识的了解。在实际应用中,由于很难得到这些先验知识,在存在脉冲噪声的低信噪比的公共场所异常声源定位中,式(8)并不成立<sup>[14]</sup>,可见 GCC-HB 时延估计方法性能将大大下降。

#### 4 基于反正切变换的改进 GCC-HB 时延估计方法 (ATAN-IHB)

公共场所中存在分数低阶  $\alpha$  稳定分布噪声 ( $0 < \alpha \leq 2$ ) 以及混响,随着  $\alpha$  值的减小,其二阶统计量和一阶统计量将不再存在,用  $|\Phi_{x_1x_2}(\omega)|$  来近似  $\Phi_{ss}(\omega)$  会有较大的误差,使得相关函数  $R_{GCC}(\tau)$  的峰值不再明显,不但会出现伪峰,并且可能将不再正常工作,从而影响时延估计值的准确性。

##### 4.1 反正切变换

根据最小分散系数准则(minimum dispersion, MD),采用非线性变换可以获得比较理想的效果。当变换函数

$\psi[\cdot]$  为奇对称单调增函数时,能对变量的正值和负值进行同样的比例变换,对于较小的自变量,变换函数近似线性,对数据没有太大的影响。而对于较大的脉冲数值,通过变换函数  $\psi[\cdot]$  将输入信号映射到某一范围,以达到抑制尖峰脉冲的目的。基于这种思想,本文考虑采用反正切函数作为非线性映射函数来抑制脉冲噪声,变换后的输入信号在统计量上存在二阶矩,从而保证算法在分数低阶  $\alpha$  稳定分布噪声下的韧性<sup>[14]</sup>。

对 2 个输入信号的反正切变换为:

$$r_1(n) = \psi[x_1(n)] = \arctan[kx_1(n)] \quad (10)$$

$$r_2(n) = \psi[x_2(n)] = \arctan[kx_2(n)] \quad (11)$$

式中:  $\arctan[\cdot]$  为反正切函数,  $k$  为控制非线性变换程度的参数 ( $k > 0$ )。

##### 4.2 HB 加权函数 $H_{12}(\omega)$ 的改进

在 HB 加权方法中,由于用  $|\Phi_{x_1x_2}(\omega)|$  来代替  $\Phi_{ss}(\omega)$ ,因此,在高信噪比的情况下,两者之间的差值较小,估计出的时延精度较高;而在低信噪比的情况下,两者之间将会有较大的差值,估计出的时延精度较低。

随着信噪比的下降,  $\Phi_{ss}(\omega)$  在  $|\Phi_{x_1x_2}(\omega)|$  中所占的比例也下降,为此,本文将 HB 的加权函数  $H_{12}(\omega)$  修改为:

$$H_{12}^*(\omega) = \left| \frac{\Phi_{x_1x_2}(\omega)}{\Phi_{x_1x_1}(\omega)\Phi_{x_2x_2}(\omega)} \right|^\lambda \quad (12)$$

式中:  $0.5 \leq \lambda \leq 1$ ,  $\lambda$  随信噪比的改变而变化。

$$\lambda = \begin{cases} \lambda_0, & \sigma < \sigma_0 \\ \frac{\lambda_1 - \lambda_0}{\sigma_1 - \sigma_0}(\sigma - \sigma_1) + \lambda_1, & \sigma_0 \leq \sigma < \sigma_1 \\ \lambda_1, & \sigma > \sigma_1 \end{cases} \quad (13)$$

式中:  $\sigma$  表示现场环境下的信噪比,信噪比按照  $SNR = 10\log_{10}(v_s^2/v_n^2)$  设定,其中  $v_s^2$  表示信号项的方差,  $v_n^2$  表示现场环境下背景噪声项的方差。  $\sigma_0$ 、 $\sigma_1$ 、 $\lambda_0$  和  $\lambda_1$  是按实际情况确定的常数,且  $\lambda_1 > \lambda_0$ 、 $\sigma_1 \geq \sigma_0$ 。

加权以后的互功率谱密度函数  $\Phi_{x_1x_2}^*(\omega)$  改变为:

$$\Phi_{x_1x_2}^*(\omega) = H_{12}^*(\omega)\Phi_{x_1x_2}(\omega) \quad (14)$$

在式(12)中,若取  $\lambda = 1$ ,则为 HB 法。

##### 4.3 多帧 HB 加权

加长输入声源信号的长度可以提高算法的抗噪能力,因此为了进一步突出互相关函数的峰值,对两个传声器信号之间的互功率谱密度函数做累积加权。互相关累积函数为:

$$S_{x_1x_2}^m(\omega) = \begin{cases} \Phi_{x_1x_2}^{*1}(\omega), & m = 1 \\ 0.5(\Phi_{x_1x_2}^{*1}(\omega) + \Phi_{x_1x_2}^{*2}(\omega)), & m = 2 \\ (\Phi_{x_1x_2}^{*(m-2)}(\omega) + \Phi_{x_1x_2}^{*(m-1)}(\omega) + \Phi_{x_1x_2}^{*m}(\omega))/3, & m = 3 \end{cases} \quad (15)$$

式中:  $m$  为帧数。对  $S_{x_1x_2}^m(\omega)$  求反傅里叶变换, 即将式(15)代入式(6)可得:

$$R_{GCC}^m(\tau) = \int_{-\infty}^{\infty} S_{x_1x_2}^m(\omega) e^{-j\omega\tau} d\omega \quad (16)$$

在估计时延值时不仅可以利用当前的信号帧, 同时也充分利用了历史结果, 可以防止脉冲噪声影响时延估计值。  $R_{GCC}^m(\tau)$  的峰值位置即为传声器 1 和传声器 2 间的时延。图 4 为 ATAN-IHB 方法的时延估计流程框图, 其中  $(\cdot)^*$  表示共轭运算,  $\otimes$  表示相关运算。

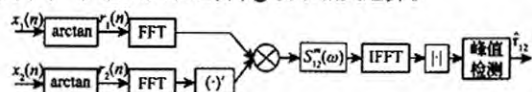


图 4 ATAN-IHB 方法时延估计流程框图

Fig. 4 Framework for time delay estimation of ATAN-IHB

## 5 实验结果及分析

实验采用异常声音库<sup>[13]</sup>中的比较典型的 3 种枪声作为目标声源。3 种枪声的总长度为 1 056 ms, 采样频率为 16 kHz。在处理过程中, 帧长为 512 点 (相当于 32 ms 的声音数据), 采用半重叠的汉宁窗 (Hanning)。选择白天操场环境下的环境噪声与分数低阶  $\alpha$  稳定分布噪声的混合噪声作为背景噪声, 其噪声包括汽车鸣笛声、行人说话声以及大气噪声等。图 5 显示了操场环境噪声、对称  $\alpha$  稳定分布噪声 ( $\alpha = 1.8$  时  $S\alpha S$  离散分布实现)、操场现场噪声环境下  $SNR = 0$  dB 时枪声的波形对比图。

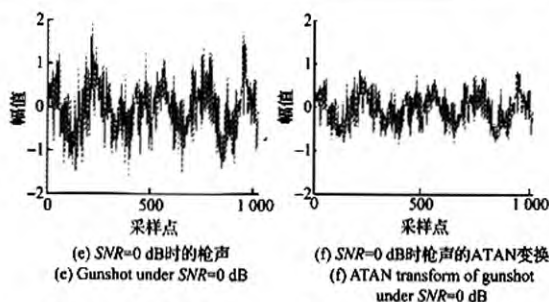
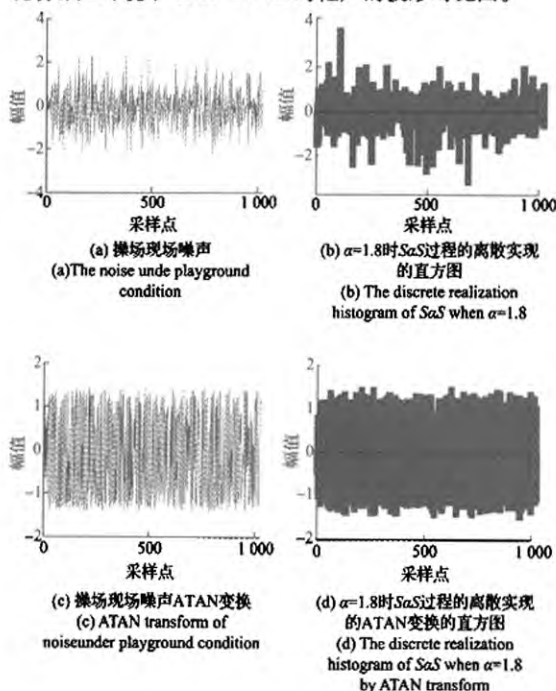
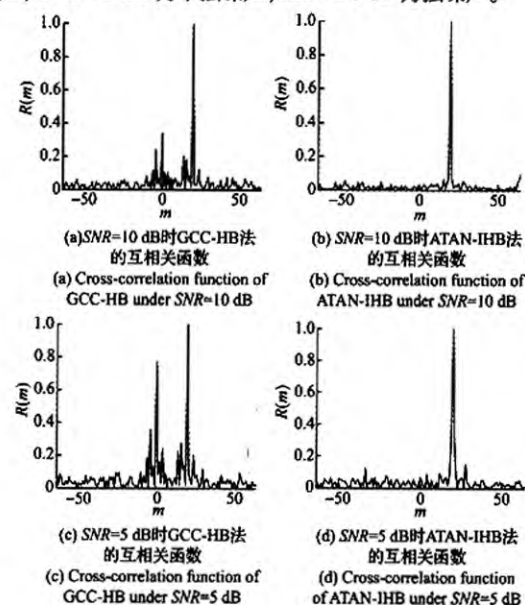


图 5 操场环境噪声下  $\alpha = 1.8$  时  $S\alpha S$  离散实现及  $SNR = 0$  dB 时枪声及其反正切变换

Fig. 5 The waveform and arc tangent transform of the noise under playground condition; the discrete realization of  $S\alpha S$  when  $\alpha = 1.8$  and the gunshot under  $SNR = 0$  dB and playground condition

从图 5(a) 和 (b) 中可以看出操场环境下的背景噪声与  $\alpha = 1.8$  时  $S\alpha S$  分布接近。图 5(c) 和 (d) 可以看出 ATAN 变换能有效的抑制噪声中的尖峰脉冲。而图 5(e) 和 (f) 在信号变化不大的情况下, ATAN 变换能有效地抑制了噪声信号中的尖峰脉冲。

为了检验本文算法的性能, 采用操场现场环境噪声作为背景噪声, 在不同的信噪比下进行计算机模拟。由于互相关函数峰值的尖锐程度以及伪峰的影响体现出该算法估计时延的精度, 所以仿真结果可以由互相关函数图来表示。图 6 中横坐标  $m$  表示采样点, 纵坐标  $R(m)$  表示互相关函数值, 其中, 正确的互相关函数峰值应该出现在第 20 个采样点处。图 6 为不同信噪比情况下, GCC-HB 法和 ATAN-IHB 法的性能比较。其中  $SNR = 10$  dB 为中度噪声,  $SNR = 5$  dB 为中强噪声,  $SNR = 0$  dB 为强噪声。



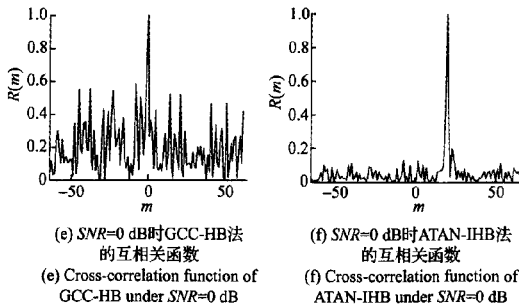


图 6 不同信噪比环境下 2 种方法的性能比较

Fig. 6 Comparison of the performances of GCC-HB and ATAN-IHB under different SNR conditions

从图 6(a)与(b)可以看出,在  $SNR = 10\text{ dB}$  情况下,2 种方法互相关函数的峰值都比较尖锐,GCC-HB 方法在冲击噪声存在的时候其峰值有一定的拓展,并且最大的伪峰幅值约为 0.35。而 ATAN-IHB 法基本没有伪峰出现。此时两者均给出了正确的时延估计值;从图 6(c)与(d)可以看出,当信噪比降低为  $SNR = 5\text{ dB}$  时,GCC-HB 法的互相关函数峰值拓展的比较明显,并且最大的伪峰幅值增大到 0.78。而 ATAN-IHB 法的互相关函数峰值仍很尖锐,伪峰幅值仅为 0.1,此时两者仍然能给出正确的时延估计值;从图 6(e)与(f)则可以看出,在  $SNR = 0\text{ dB}$ ,即信噪比很低时,GCC-HB 法的互相关函数的伪峰幅值明显增大,约为 0.58,此时 GCC-HB 法给出了错误的时延估计值。而 ATAN-IHB 法的互相关函数峰值仍然比较尖锐,伪峰幅值均在 0.2 以下,在有一定拓展的同时仍能给出正确的时延估计值。

从图 6 可以看到,公共场所环境噪声的信噪比从  $10 \sim 0\text{ dB}$  变化时,ATAN-IHB 方法得到的互相关函数的伪峰峰值的平均值为  $0.05 \sim 0.1$ ,小于 GCC-HB 法得到的互相关函数的伪峰峰值的平均值  $0.2 \sim 0.3$ 。ATAN-IHB 方法的互相关函数的伪峰峰值的平均值增长的幅度仅为 0.05,也明显小于 GCC-HB 法的增长幅度 0.2。

在反正切变换处理环境噪声有效的情况下,比较了 GCC-HB 法 ( $\lambda = 1$  及  $\lambda = 0.75$ ) 与 ATAN-IHB 法的性能,对 20 组数据(每组 65 帧)进行了实验。表 1 给出分别用 GCC、维纳加权广义互相关(GCC-WP)、GCC-HB 及 ATAN-IHB 进行 TDOA 估计获得的正确时延估计值的平均百分比,其计算方法为:正确率 = 正确样本数/样本总数  $\times 100\%$ 。其结果如表 1 所示。

由表 1 可以看出,环境噪声的信噪比不同的时候,GCC-HB 方法均比 GCC、GCC-WP 方法更好。但在中度噪声与强噪声的情况下,TDOA 估计的正确率下降的比较很明显, $\lambda = 1$  和  $\lambda = 0.75$  时分别下降 7.69% 和 18.46%。ATAN-IHB 法正确估计时延值的百分比均高于

其他 GCC-HB 法,在强噪声情况下 TDOA 估计的正确率提高了 21.54%。

表 1 不同信噪比下 TDOA 估计的正确率(%)比较  
Table 1 Comparison of TDOA estimation accuracy under different SNR conditions

方法	噪声			
	20 dB	10 dB	5 dB	0 dB
GCC	89.23	66.15	47.63	35.38
GCC-WP	92.31	72.31	55.38	41.54
GCC-HB	$\lambda = 1$	95.38	75.38	60.00
	$\lambda = 0.75$	92.31	78.46	66.15
ATAN-IHB		98.46	89.23	81.54

就计算量而言,ATAN-IHB 法比 GCC-HB 法多计算每帧信号的能量、2 次反正切变换,由于信号样本数较小,虽然计算量稍有增大,但完全可以在实时系统中实现。

6 结 论

本文提出了一种基于反正切变换的改进 GCC-HB 时延估计方法(ATAN-IHB)。首先对传声器接收到的声音信号作反正切变换,降低环境中高斯噪声与脉冲噪声等混合噪声的影响;然后根据公共场所环境下信噪比多变的情况,提出 HB 加权函数的改进方法,提高 HB 加权时延估计在不同混合信噪比下抗噪能力;最后引入一个累积函数对 ATAN-IHB 算法进行改进,加长输入声音信号的长度,从而更加准确地估算出两个传声器间的时间延迟值。本文提出的算法能很好地抑制环境噪声中的尖峰脉冲,在公共场所环境下有良好地韧性。虽然计算量稍有增大,但仍然可以较好的满足实时性的要求。

参考文献

[1] MICHAEL S B, HARVEY S. A practical methodology for speech source localization with microphone arrays [J]. Computer Speech and Language, 1997, 11(2): 91-126.  
[2] RYAN J G, GOUBRAN R A. Application of near-field optimum microphone arrays to hands-free mobile telephony[J]. IEEE Trans. on Vehicular Technology, 2003, 52(2): 390-400.  
[3] KENNED Y R. Broadband near-field beam-forming using a radial beam-pattern transformation[J]. IEEE Trans. on Signal Process, 1998, 46(8): 2147-2156.  
[4] OH S K, UN C K. Improved MUSIC algorithm for high-resolution array processing[J]. Electronic Letters, 1989, 25(22): 1523-1525.  
[5] KNAPP C H, CARTER G C. The generalized correlation method for estimation of time delay[J]. IEEE Trans. on

- ASSP, 1976, 24(4): 320-327.
- [6] 牛新亮, 赵国庆, 张永顺, 等. 可以稳定快速收敛的无源时差定位系统[J]. 仪器仪表学报, 2010, 31(5): 1114-1119.
- NIU X L, ZHAO G Q, ZHANG Y SH, et al. Stable-fast-convergence passive TDOA location system [J]. Chinese Journal of Scientific Instrument, 2010, 31(5): 1114-1119.
- [7] 周立君, 刘宇, 郭城. 无线传感器网络声源定位中的时延判决方法研究[J]. 电子测量技术, 2010, 33(4): 35-37, 64.
- ZHOU L J, LIU Y, GUO CH. Study on time delay estimation in sound source locating base on WSN[J]. Electronic Measurement Technology, 2010, 33(4): 35-37, 64.
- [8] WAN X W, WU Z Y. Improved steered response power method for sound source localization based on principal eigenvector [J]. Applied Acoustics, 2010, 71: 1126-1131.
- [9] SWARTLING M, GRBIĆ N. Calibration errors of uniform linear sensor arrays for DOA estimation: an analysis with SRP-PHAT [J]. Signal Processing, 2011, 91(4): 1071-1075.
- [10] 崔玮玮, 曹志刚, 魏建强. 声源定位中的时延估计技术[J]. 数据采集与处理, 2007, 20(1): 90-99.
- CUI W W, CAO ZH G, WEI J Q. Time delay estimation techniques in source location[J]. Journal of Data Acquisition Processing, 2007, 20(1): 90-99.
- [11] YOUNG D H, AHMED N, CARTER G C. On using the LMS algorithm for time delay estimation [J]. IEEE Trans. on Acoustics, Speech and Signal Processing, 1982, 30(5): 798-801.
- [12] JOSEPH C H, RONALD E B. Optimum estimation of time delay by a generalized correlation[J]. IEEE Trans. on ASSP, 1979, 27: 373-380.
- [13] 栾少文, 龚卫国. 公共场所典型异常声音的特征提取[J]. 计算机工程, 2010, 36(7): 208-210.
- LUAN SH W, GONG W G. Feature extraction of typical abnormal sounds in public places [J]. Computer Engineering, 2010, 36(7): 208-210.
- [14] SUN Y M, QIU T S. The new HB weighted time delay estimation algorithm under impulsive signal environments [J]. Journal of Systems and Engineering and Electronics, 2008, 19(6): 1102-1108.
- [15] 邱天爽, 王宏禹, 张扬, 等. 一种基于非线性变换的 EP 潜伏期变化自适应检测方法[J]. 电子与信息学报, 2002, 24(12): 1842-1849.
- QIU T SH, WANG H Y, ZHANG Y, et al. A non-linear transform based adaptive EP latency change detection algorithm [J]. Journal of Electronics and Information Technology, 2002, 24(12): 1824-1849.

### 作者简介



李伟红, 2006 年于重庆大学获得博士学位, 现为重庆大学副教授, 主要研究方向为模式识别、图像处理。

E-mail: weihongli@cqu.edu.cn

Li Weihong received her Ph. D. degree from Chongqing University in 2006, and now she is an associate professor in Chongqing University. Her researches are mainly focused on pattern recognition and image processing.



汤海兵, 2011 年于重庆大学获得硕士学位, 主要研究方向为音频信号处理。

E-mail: 307tanghaibing@163.com

Tang Haibing received his M. Eng. degree from Chongqing University in 2011. His research is mainly focused on audio signal processing.