

文章编号:1001-9081(2006)11-2696-02

一种改进的声测定位时延估计算法

孙韶杰¹, 孙绍俊², 李国辉¹, 李红梅²

(1. 国防科学技术大学 信息系统与管理学院, 湖南 长沙 410073; 2. 山东省图书馆 网络部, 山东 济南 250100)
(sshj_mil@126.com)

摘要:研究了时延估计算法在被动声测定位中的应用,提出了一种改进的基于最大似然(ML)权函数的广义互相关时延估计算法。改进的算法采用加窗法和最小均方差(LMS)滤波法,弥补了原算法计算量大及无法消除回响干扰的不足。仿真结果表明,改进的算法计算复杂度明显降低,能够有效地消除回响干扰,具有较高的时延估计精度和鲁棒性。

关键词:时延估计;广义互相关法;最大似然权函数;声测定向;最小均方差滤波

中图分类号: TN911.72 **文献标识码:** A

An improved time-delay estimation algorithm for acoustic detection and location

SUN Shao-jie¹, SUN Shao-jun², LI Guo-hui¹, LI Hong-mei²

(1. College of Information System and Management, National University of Defense Technology, Changsha Hunan 410073, China;
2. Department of Network, Library of Shandong Province, Jinan Shandong 250100, China)

Abstract: The application of time delay estimation in passive acoustic detection and location was studied. An improved maximum likelihood estimator as the generalized correlation method for time-delay estimation, which adopted the methods of windowing and least mean square (LMS) filter, was proposed. The simulation results show that the proposed algorithm can reduce the complexity of computation obviously, eliminate the influence of reecho efficiently, and has better estimation accuracy and robustness.

Key words: time delay estimation; generalized correlation method; maximum likelihood (ML); acoustic detection and location; least mean square filter(LMS)

声测定位^[1]是一种利用目标的音频信号来实现定位的无源探测技术。声测定位系统采用被动方式工作,不辐射电磁波,具有很强的隐蔽性。同时,声音信号是典型的机械波,与微波、可见光、红外等电磁波相比,它能够绕过障碍物进行传播,不存在线性遮蔽效应。因此,声测定位技术在军事领域和民用领域都得到了广泛的应用,如在军事领域,声测定位系统可以有效地弥补雷达系统无法探测低空飞行器的缺陷;在民用领域,视频会议系统中可以采用传声器阵列对说话人进行定向,从而控制摄像机自动转向发言人。

常用的声测定位法按其估计方法的不同大致可以分为三类^[2]:基于声束形成的定位法、基于高分辨率谱估计的定位法以及基于时延估计的定位法。基于时延的定位算法定位精度较高,而且在计算量上远远小于前两种方法,是目前声测定位系统中常用的方法。本文提出的基于最大似然权函数的时延估计算法,在抗噪声和回响干扰性方面进行了相应改进,工程上易实现,同时具有较强的鲁棒性和实时性。

1 基于最大似然权函数的广义互相关法原理

1.1 基本的广义互相关法原理

时延估计是基于时延估计的声测定位算法的第一步,也是最关键的一步,因为时延估计的精度直接影响着定位的精度。目前时延估计的算法很多,如广义互相关函数法^[3]、互功率谱相位法、自适应滤波法和参量模型法等。互功率谱相位法由于目标信号频率较低、实时性要求高,最小二乘拟合次数不能太高,故估计精度一般不高。自适应滤波法运算速度快,

但点声源干扰对其估计精度影响较大。广义互相关法具有较好的综合性能,是目前人们关注较多的时延估计方法。

广义互相关法采用的信号模型可用式(1)表示:

$$\begin{cases} x_1(t) = s(t) + n_1(t) \\ x_2(t) = s(t - D) + n_2(t) \end{cases} \quad (1)$$

其中, $x_1(t)$ 和 $x_2(t)$ 分别为两个声传感器的接收信号, $s(t)$ 为原信号, D 为两信号的时延, $n_1(t)$ 和 $n_2(t)$ 分别为两个声传感器的加性噪声。

为了求得时延 D ,需要对信号 $x_1(t)$ 和 $x_2(t)$ 进行预处理,然后再送入相关器求互相关函数,这种方法称之为广义互相关法。

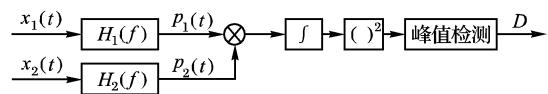


图1 广义互相关时延估计方法模型

广义互相关时延估计方法模型由一对加权滤波器 H 、乘法器、积分器、平方器和峰值检测器组成。信号 $x_1(t)$ 和 $x_2(t)$ 经滤波后得到的信号的互功率谱 $G_{x_1x_2}(f)$ 可表示为:

$$G_{p_1p_2}(f) = H_1(f) \overline{H_2(f)} G_{x_1x_2}(f) \quad (2)$$

又因为互相关函数与互功率谱之间是一对傅立叶变换的关系,因此相关函数 $R_{12}(\tau)$ 可表示为:

$$R_{12}(\tau) = \int_{-\infty}^{\infty} W(f) \overline{G_{x_1x_2}(f)} e^{j2\pi f\tau} df \quad (3)$$

其中, $W(f) = H_1(f) \overline{H_2(f)}$ 称为权函数。

可以看出,当权函数 $W(f) = 1$ 时,就是一般的互相关法。一般互相关法只适应于谱线连续的宽带信号。而对于谱线随

收稿日期:2006-05-25;修订日期:2006-07-05

作者简介:孙韶杰(1981-),男,山东即墨人,硕士研究生,主要研究方向:多媒体信号处理、多媒体信息系统;孙绍俊(1975-),男,山东即墨人,主要研究方向:计算机网络、多媒体信息系统;李国辉(1963-),男,湖南长沙人,教授,博士生导师,主要研究方向:多媒体信息系统、多媒体信息安全、图像压缩、数据挖掘;李红梅(1974-),女,山东菏泽人,主要研究方向:多媒体处理。

频率跃变的信号,互相关函数将只由少数很强的频率成分决定,从而会引起较大的时延估计误差,因此要选用权函数以提高时延估计的精度。

1.2 最大似然加权函数

基于最大似然加权的时延估计法是由 G. C. Carter^[4]等人于 1976 年提出的,并处于不断地发展中^[5,6],至今仍在被动声测定位中发挥着重要的作用。该方法的基本思想是按照输入信号的信噪比进行加权预处理,其估计方差可以达到克拉美罗下界(CRLB)。不过,该方法的最优性能依赖于输入信号和噪声统计先验知识的了解程度。在实际应用中,可以将自适用信号处理技术与最大似然时延估计相结合,来优化算法的性能。

在广义互相关时延估计框架中,最大似然加权函数 $W(f)$ 为:

$$W(f) = [|\gamma_{12}(f)|^2 / (1 - |\gamma_{12}(f)|^2)] / |G_{x_{1x_2}}(f)| \quad (4)$$

其中, $|\gamma_{12}(f)|^2$ 为两接收信号的幅度平方相干函数。

最大似然加权的广义互相关函数 $R_{x_{1x_2}}(\tau)$ 为:

$$\begin{aligned} R_{x_{1x_2}}(\tau) &= \text{FFT}^{-1} \{ W(f) G_{x_{1x_2}}(f) \} \\ &= \text{FFT}^{-1} \left\{ \left[\frac{G_{x_{1x_2}}(f)}{|G_{x_{1x_2}}(f)|} \right] \right. \\ &\quad \left. \left[\frac{|\gamma_{12}(f)|^2}{(1 - |\gamma_{12}(f)|^2)} \right] \right\} \quad (5) \end{aligned}$$

其中

$$\begin{aligned} |\gamma_{12}(f)|^2 &= [G_{x_{1x_2}}(f)/G_{x_{2x_2}}(f)] [G_{x_{2x_1}}(f)/G_{x_{1x_1}}(f)] \\ &= H_{12}(f)H_{21}(f) \quad (6) \end{aligned}$$

$$\begin{aligned} &G_{x_{1x_2}}(f) / |G_{x_{1x_2}}(f)| \\ &= [G_{x_{1x_2}}(f)/G_{x_{2x_2}}(f)] / [G_{x_{1x_2}}(f)/G_{x_{2x_2}}(f)] \\ &= H_{12}(f) / |H_{12}(f)| \quad (7) \end{aligned}$$

由式(6), (7)可知,最大似然加权的广义互相关函数可以由两个 Roth 处理器 $\text{FFT}^{-1}(H_{12}(f))$ 和 $\text{FFT}^{-1}(H_{21}(f))$ 的组合来实现(详细推导过程参考文献[7]),该方法的优点是估计精度较高,收敛快,但是其明显的缺点是计算量大,而且难以消除回响对信号的干扰。本文将针对上述问题做进一步改进。

2 算法改进

2.1 采用加窗法降低计算复杂度

在实现 Roth 处理器的自适应迭代过程中,我们发现权函数的峰值点在远未到达正确时延估计值之前,就已经收敛到其稳态值的一个较小的邻域中。通常,这个小邻域的宽度仅为几个采样间隔,一般不会超出主瓣的范围。也就是说,绝大部分有意义的迭代过程仅在此小邻域中进行,而该小邻域之外的权系数的迭代更新,对权矢量峰值位置的确定基本没有影响。所以我们用一个矩形窗口与滤波器权矢量相乘,并仅对窗口内的权系数进行迭代更新。本文采用的窗口函数 $c(n)$ 可表示为:

$$c(n) = \begin{cases} \delta(n), & \hat{D} - N \leq n \leq \hat{D} + N \\ 0, & \text{其他} \end{cases} \quad (8)$$

其中, \hat{D} 表示加窗时的时延估计值, N 表示窗半宽。

实验证明,通过该方法仅对上述邻域内的权系数进行迭代计算,可以大大减少自适应处理过程中的迭代更新所花费的冗余的计算量,便于实时应用。

2.2 利用 LMS 滤波法消除回响影响

最大似然加权函数^[4]是一个频域信噪比函数,它对信噪

比大的频段赋予大权值,而对信噪比小的频段赋予小权值,从而较好地抑制了噪声的影响。

从上述对最大似然加权函数的分析中,可以看到它虽然可以降低噪声的影响,但是无法消除回响对信号的干扰。针对回响干扰问题,本文采用 LMS 滤波法^[8]对信号进行预处理。该算法描述如下:设 $x(n)$ 是原始信号, $y(n)$ 是滤波后的信号, $w(n)$ 是自适应迭代权系数($w(1) = 1$)。

$$y(n) = \sum_{i=-p}^p w(n+i) x(n+i) \quad (9)$$

$$\sigma_x^2 = \sum_{n=0}^{N-1} (x(n) - \mu_x)^2, \mu_x = 1/N \sum_{n=0}^{N-1} x(n) \quad (10)$$

$$z(n) = w(n)x(n), w(n+1) = w(n) + z(n)x(n)/\sigma_x^2 \quad (11)$$

实验证明,这一算法可以有效地消除回响对信号的干扰。图 2 是在密闭房间内采集到的回响信号,图 3 为经过 LMS 滤波后的信号。

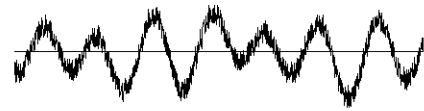


图 2 未经处理的回响信号

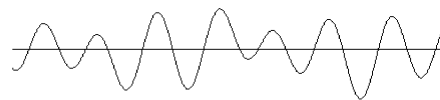


图 3 经过 LMS 滤波后的信号

2.3 改进算法的框架流程

本文实现的改进算法的流程如图 4 所示,可分为以下四个步骤:

- 1) 利用一对麦克风采集两路音频信号,然后分别经过模数转换(A/D)及数字采样,得到待处理的信号。
- 2) 将待处理信号进行预处理,然后进行 LMS 滤波,以消除回响的干扰。
- 3) 使用加窗的基于最大似然权函数的广义互相关法求两路信号的相关函数。
- 4) 对求得的相关函数进行峰值检测,得到的峰值即为时延的估计值。

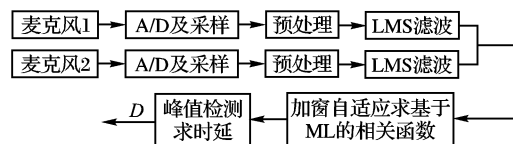


图 4 广义互相关时延估计算法流程

3 实验仿真

为了检验本文提出的改进算法的性能,我们在具备回响条件的密闭房间内进行了计算机仿真,并在时延估计精度和计算复杂度这两个方面与改进前算法进行了对比。

实验中,利用麦克风 1 获得实际的音频信号,通过设定不同的采样率得到待处理的数字音频数据 x_1 ; 然后根据指定的信噪比,将信号 x_1 加噪,使 x_1 可以更好地模拟复杂环境下的信号;接着将信号 x_1 平移指定的采样周期数后得到数字音频信号 x_2 ,用 x_2 模拟麦克风 2 的输入信号,这样处理的目的是为了能够获得精确的理论时延值,例如,若指定的平移周期数为 n ,信号的采样频率为 f ,则理论时延值为 $n * f$ 单位时间。

(下转第 2723 页)

$$S_{CIS}(x) = \frac{\sum_{i=1}^{n \times m-1} S_{CIS}(x, y_i) \cdot f(x, y_i)}{\sum_{i=1}^{n \times m-1} f(x, y_i)}$$

其中, m, n 分别为每层图像的长度和宽度。

3 基于阈值的显著区域

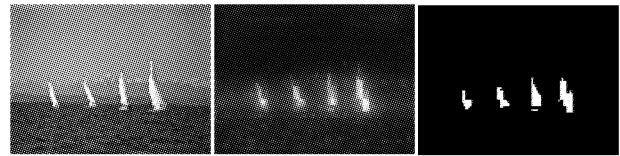
通过对每层计算显著性,可以得到显著区域的轮廓。要想得到清晰准确的显著区域,需要把各个层的结果叠加,使得各层的轮廓互相填充。如果想进一步获得明显的显著区域(显著区域用白色表示,其余区域用黑色表示),就需要根据上一步叠加的结果计算出一个阈值。这里,我们提出了一个设置动态阈值的方法。

图 2 所示的是各层叠加后的灰度图像(如图 3(b)所示)的直方图。其中, A 为直方图的数学期望, B 为数学期望与标准差的和, C 为数学期望与标准差的差, D 为 256 个直方图数据中最接近第 256 个分量的一个不为零的点。根据概率论和数理统计可知,直方图中大部分数据都集中在 C 和 B 之间。根据这一特点,我们可以把阈值设置为 B 和 D 之间的某一个值,一般设为 B 和 D 的中点或者直接把 B 点设为阈值。这样,无论原始图像如何,都可以找到一个适合它的阈值。这个阈值可以有效地消除叠加后图像中的噪声点,达到检测显著区域的目的。

4 实验结果

对于原始图像(如图 3(a)所示),把原始图像分为 3 层,对其中的每一层,应用上述方法计算其显著性,然后把各层的显著性值叠加,结果如图 3(b)所示。对于叠加后的值,按照上述方法计算阈值。大于此阈值的像素点用白色表示,其余

像素点用黑色表示,形成最终的显著区域(如图 3(c)所示)。可以看出,采用正态分布函数,很好地解决了像素间的关联问题,并产生了预期的结果。由于对每一层计算所得的值只反映了物体的轮廓,而叠加后的图像恰恰弥补了每一层的不足。此外,由于采用了动态阈值的方法,很好地消除了叠加后的图像(如图 3(b)所示)中的噪声点。



(a) 原始图像 (b) 各层显著性值叠加后的图像 (c) 显著区域
图 3 实验图像

参考文献:

- [1] 沈庭之, 方子文. 数字图像处理及模式识别[M]. 北京: 北京理工大学出版社, 1998.
- [2] MA Y-F, ZHANG H-J. Contrast-based Image Attention Analysis by Using Fuzzy Growing[J]. Proceedings of the 11th ACM International Conference on Multimedia[C], 2003. 374 - 381.
- [3] KO B, KWAK SY, BYUN H. SVM-based Salient Region(s) Extraction Method for Image Retrieval[J]. Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004[C], 2004. 977 - 980.
- [4] LTTI L, KOCH C, NIEBUR E. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254 - 1259.
- [5] 章毓晋. 图像处理和分析基础[M]. 北京: 高等教育出版社, 2002.
- [6] 章毓晋. 基于内容的视觉信息检索[M]. 北京: 科学出版社, 2003.
- [7] SUN Y. Hierarchical Object-Based Visual Attention for Machine Vision[J]. Artificial Intelligence, 2003, 146(1): 77 - 123.

(上接第 2697 页)

仿真实验中,指定的平移周期数为 16,在特定的信号采样频率条件下,均进行 50 次实验,然后分别计算两种算法的平均时延实际值和平均所耗时间。表 1 是不同采样频率条件下的仿真结果。

表 1 不同采样频率条件下的仿真结果

采样频率 /KHz	时延值/ms		算法耗时/ms		
	理论	改进前	改进后	改进前	改进后
1	15.6250	6.6437	13.8049	1.131	0.362
2	7.8125	4.4563	6.6125	1.141	0.371
4	3.9062	2.8049	3.7975	1.184	0.423
8	1.9531	1.4234	1.8523	1.521	0.635
16	0.9766	0.6542	0.7542	2.036	0.752
32	0.4883	0.2546	0.3542	2.852	0.954
64	0.2441	0.0825	0.1923	3.215	1.235

从仿真结果可以看出,在同信噪比不同采样频率的情况下,由于改进算法能够有效地消除回响的干扰,所以其时延估计精度相对于改进前算法的时延估计精度,平均提高了 20% ~ 30%。另外,从表 1 可知改进算法所耗时间平均仅为改进前算法的 30% 左右,计算复杂度大大降低,便于在工程中实时应用。

4 结语

本文针对基于最大似然权函数的广义互相关时延估计算法的计算量大以及无法消除回响干扰的不足,提出了一种改

进的时延估计算法,并利用实测噪声数据进行了仿真实验。仿真结果表明,改进算法能够有效消除噪声和回响的干扰,同时具有计算复杂度低、对硬件要求不高、实时性好等优点,具有潜在的应用前景。

参考文献:

- [1] SROUR N, ROBERTSON J. ARL-TR-706, Remote Netted Acoustic Detection System: Final Report[R]. US Army Research Laboratory, 1995.
- [2] BRANDSTEIN M. Microphone Arrays[M]. New York: Springer-Verlag, 2001.
- [3] RUI Y, FLORENCIO D. Time delay estimation in the presence of correlated noise and reverberation[J]. IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004, 2(2): 133 - 136.
- [4] KNAPP CH, CARTER GC. The generalized correlation method for estimation of time delay[J]. IEEE Transaction on Acoustics, Speech and Signal Processing, 1976, 24(4): 320 - 327.
- [5] CARTER GC. Coherence and time delay estimation[J]. Proceedings of the IEEE, 1987, 75(2): 236 - 255.
- [6] 陈华伟, 赵俊渭, 郭业才. 二次加权频域自适应时延估计算法与应用[J]. 声学学报, 2003, 28(1): 61 - 65.
- [7] 邱天爽, 王宏禹. 加窗 LMS-ML 自适应时间延迟估计[J]. 大连理工大学学报, 1994, 34(4): 470 - 476.
- [8] GU Y, TANG K, CUI H. Convergence analysis of a deficient-length LMS filter and optimal-length sequence to model exponential decay impulse response[J]. Signal Processing Letters, 2003, 10(1): 4 - 7.