

汽车鸣笛声定位系统仿真

孙懋珩, 俞莹婷

(同济大学电子与信息学院, 上海 201804)

摘要: 提出了一种汽车鸣笛声定位系统用以协助执法部门锁定违章鸣笛车辆。该系统采用分布在立体空间的五元麦克风阵列, 整个定位系统由三个模块组成。首先根据鸣笛声的频域特性检测出信号, 然后基于广义互相关理论计算到达各麦克风信号的时延差, 并通过对多次估计求均值提高时延估计的准确性。最后利用闭式球形插值算法得到声源位置信息。仿真结果表明该系统可以有效地定位汽车鸣笛声, 为实际的系统实现提供了一次初步探讨。

关键词: 声音定位; 麦克风阵列; 时延估计; 球形插值

中图分类号: TB556

文献标识码: A

文章编号: 1000-3630(2009)-05-0640-05

DOI 编码: 10.3969/j.issn1000-3630.2009.05.016

Simulation of car horn localization system

SUN Mao-heng, YU Ying-ting

(Institute of Electronics and Information, Tongji University, Shanghai 201804, China)

Abstract: A system for car horn location which aims to helping agency to supervise the illegal car is proposed. The system consists of a 5-element microphone array and its algorithm is implemented in three modules. The first step is to detect the signal by the traits of the car horn in frequency domain. Then the time delay estimation (TDE) for microphone arrays is determined by generalized cross correlation (GCC), and the average of the estimations is calculated to improve the accuracy. Finally closed Spherical Interpolation (SI) is implemented to obtain the location of car horn. The simulation results indicate that the system effectively gets the location of car horn, thus provides a theoretical foundation for future implementation.

Key words: sound localization; microphone arrays; time delay estimation; spherical interpolation

1 引言

近年来, 随着经济的快速发展, 各大城市机动车数量逐年增加, 截至 2007 年底全国机动车数量已达 1.5 亿辆之多。这些穿梭于城市道路上的汽车在给人们创造便利的同时, 也带来了严重的噪声污染。研究发现, 噪声污染不仅对正常生活和工作产生干扰, 而且还会诱发多种疾病。因此我国将噪声控制列为环境保护的重点, 公安部门也颁布了相应的法律法规来禁止或者限制车辆的鸣笛声。

鉴于这种现状, 为了协助公安以及相关执法部门更加准确有效地锁定违章鸣笛的车辆方位, 本文提出了将基于麦克风阵列的声源定位技术^[1]应用到城市道路禁止鸣笛的执法系统中, 对系统的具体模型及各部分的主要算法进行了详细的探讨和说明。为了验证系统设计方案的有效性, 对整个方案进行仿真, 仿真结果表明, 本文所提出的系统性能

良好, 具有很高的实用性。

2 系统模型

如图 1 所示, 基于麦克风阵列的汽车鸣笛声定位系统由鸣笛声识别和鸣笛声定位两部分组成。其中鸣笛声识别主要通过其在频域的特征来实现。

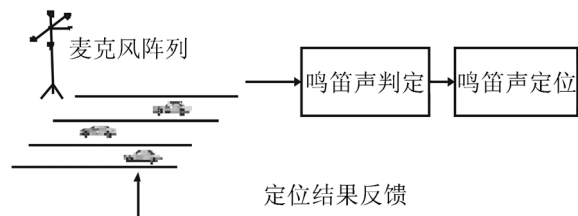


图 1 鸣笛声定位系统框图

Fig.1 Design of car horn localization system

通常的声源定位算法可以分为三类^[2]。第一类是基于波束形成的方法。该方法可以用于多声源的定位, 但存在着对初值敏感和需要声源和背景噪声等先验知识的问题。第二类是基于高分辨率谱估计的方法, 该方法可以随声源方向进行有效的估计, 但是计算量比较大。第三类是基于时间到达差的方

收稿日期: 2009-05-18; 修回日期: 2009-08-30

作者简介: 孙懋珩(1957-), 男, 山东人, 博士, 副教授, 研究方向为信号处理。

通讯作者: 俞莹婷, E-mail: yingting1222@hotmail.com

法, 由于该方法原理简单, 计算量小, 易于实现, 故本文所提出的系统采用基于时间到达差的方法对鸣笛声进行定位。

基于时间到达差的定位算法具体又分为时延估计和定位两个部分, 其中时延估计可以选择广义互相关(GCC)算法^[3], 而定位则选用闭式球形插值算法^[4,5], 以较小的运算量来满足系统的实时性需求。

3 系统关键算法

3.1 鸣笛声识别算法

本系统的目标声源是城市禁鸣路段出现的违章汽车鸣笛声, 所以可靠地检测出汽车鸣笛声是后期声源定位算法能否实现的关键所在。车用喇叭主要分为电喇叭和气喇叭, 当前汽车喇叭以电喇叭为主, 它的音量级别通常在 70~120dB, 有一到三种音调, 且以双音调为主, 时频特性如图 2 所示。本系统的定位对象为城市道路上最为常见的双音电喇叭。可以看到, 喇叭声在频域上主要分布在 3.2kHz 和 4.1kHz 两个频点上。

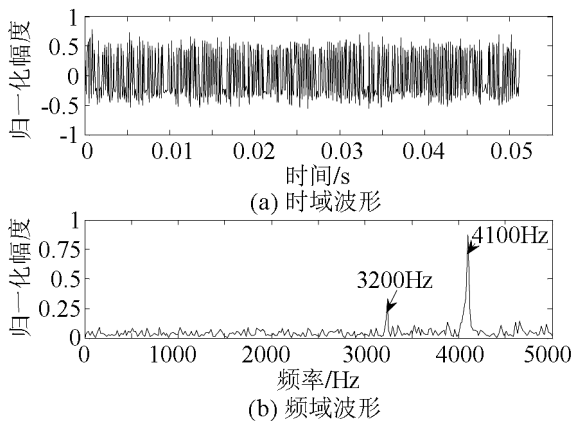


图 2 喇叭声时域与频谱特性
Fig.2 Car horn traits in time and frequency domain

根据汽车鸣笛声在频域的特性, 采用一种相对于基于模板库的识别系统更为简便的识别算法, 从而可以减少系统的运算量, 提高实时性。识别过程框图如图 3 所示, 鸣笛声识别模块由带通滤波器、分帧加窗、端点检测以及鸣笛声识别四部分组成。

首先根据汽车鸣笛声和其他城市道路声音的不同频谱特性, 本文采用 5kHz 的低通滤波器。在保留汽车鸣笛声的绝大部分能量的前提下, 滤除大部分干扰的噪声。

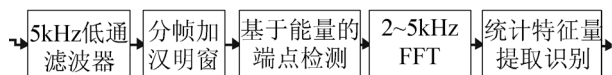


图 3 鸣笛声识别过程框图
Fig3 The process of car horn recognition

鉴于信号的时变特性, 可以利用窗函数将信号分成一定长度的帧, 并且认为在这些短时间内信号的特征基本不变。加窗时, 不同的窗长度和窗函数形状会影响到后面的分析结果, 在本系统中用主瓣较宽的汉明窗得到较为平滑的频谱。

端点检测的目的是检测出语音信号的位置, 目前已经有了很多成熟准确的端点识别算法。本系统采用基于时域短时能量的端点检测法, 见式(1)。

$$E_n = \sum_{m=-\infty}^{\infty} [x(m)w(n-m)]^2 = \sum_{m=n}^{n+N-1} [x(m)w(n-m)]^2 \quad (1)$$

其中, $x(m)$ 为信号值, $w(n-m)$ 为所加的窗函数, E_n 为第 n 帧的短时能量。

环境噪声的能量必须经过预先实验获得。然后根据式(2)算出能量门限, 最后根据计算出的门限得到语音端的开始。

$$E_L = \min[0.03(E_{\max} - E_{\min}) + E_r, 4E_r] \quad (2)$$

其中 E_{\max} 为能量最大值, E_{\min} 为能量最小值, E_r 为所有帧的能量平均值。

最后, 将短时能量方法检测出的信号变换到频域, 利用鸣笛声和噪声频谱的统计特性对鸣笛声进行识别。在统计学中, 期望值表示了数值的平均值, 方差描述了数值的离散程度, 用四阶中心距表示的峰度, 反映频数分布曲线顶端尖峭或扁平程度。本系统正是借助某个频段(2~5kHz)的这些统计学指标来区分出鸣笛声。

3.2 鸣笛声定位算法

如图 4 所示, 基于到达时间差的鸣笛声定位算法主要由两部分组成, 第一部分为时延估计, 用于计算来自同一声源的信号到达具有配对关系的两个麦克风的时间差。第二部分为定位估计, 根据所得到的时延差, 利用麦克风阵元相互之间的几何关系确定声源位置。

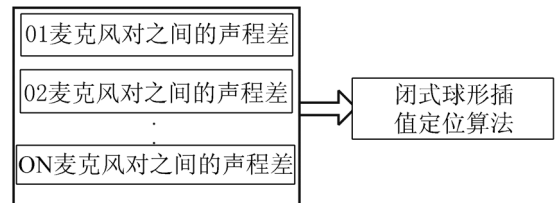


图 4 鸣笛声定位算法框图
Fig.4 The algorithm for horn localization

3.2.1 时延估计算法

在理想情况下, 第 i 个麦克风的所接收到的信号为:

$$X_i(t) = \alpha_i s(t - \tau_i) + n_i(t) \quad (3)$$

其中 $s(t)$ 为源信号; τ_i 、 α_i 、 $n_i(t)$ 分别表示从声源到第 i 个麦克风的延迟、幅度衰减和加性噪声。



图 5 用 GCC 方法进行时延估计框图
Fig.5 Time delay estimation by using GCC method

GCC 方法是一种传统的时延估计的方法,如图 5 所示。该方法借助于来自同一声源的信号存在一定的相关性这一特性,通过计算麦克风阵列接收信号之间的互相关函数,从而估计出 TDOA 值。就两个麦克风而言,接收到的信号 $x_1(t)$ 和 $x_2(t)$ 的互相关函数可以表示为:

$$R_{x12}(\tau) = E(x_1(t)x_2(t-\tau)) \approx \alpha_1\alpha_2R_s(\tau-\tau_{12}) + R_n(\tau) \quad (4)$$

式(4)中 E 为数学期望。鉴于本系统实际运用在宽阔的交通道路,可以忽略掉混响的因素,近似为信号和噪声互相关函数的叠加。

相关函数的傅立叶变换给出了两通道信号的互相关谱:

$$G_{x12}(f) = \alpha_1\alpha_2G_s(f)e^{-i2\pi f\tau_{12}} + G_n(f) \quad (5)$$

为了使 TDOA 估计不受信号本身特性的影响,并尽可能地抑制混响和噪声,需要利用不同的加权函数对观察信号的频谱做特殊处理。表 1 列举了几种通用的加权函数^[6]。

表 1 典型加权函数
Table 1 Typical weighting function

加权	表达式	特性
CC	1	对外围噪声,反射和有限观测数据敏感。
Roth	$\frac{1}{G_{x11}(f)}$	等价于维纳滤波,可以有效地抑制噪声的频带,但会展宽相关函数的峰。
SCOT	$\frac{1}{\sqrt{G_{x11}(f)G_{x22}(f)}}$	同时考虑两个通道的影响,但也会会展宽相关函数的峰。
PHAT	$\frac{1}{ G_{x12}(f) }$	相当于白化滤波,但是在信号能量较小的时候会使分母趋于零,增加误差。

对于本系统研究的双音频喇叭,信号自身在频谱就具有很明显的离散周期性,上面的加权函数都能较好地抑制噪声,得到较为准确的时延估计。

除此以外,在实际的环境中,麦克风收到的汽车鸣笛声,至少会存在直达声和路面反射声两个分量。反射分量经历了地面反射,一部分能量被吸收,至少会有 3~5dB 的衰减。但是由于麦克风之间距离比较近,它们收到的信号反射分量情况基本相似。虽然不同麦克风接收信号的时延估计函数受此影响会出现其他小的干扰相关峰,但还是可以将其能量最大的相关峰作为时延差。

3.2.2 定位算法

图 6 显示了鸣笛声定位系统的远场模型。当声源与麦克风阵列的距离较远时,即 $|s_j| \gg |p_i|$, 矢量 $s_j - p_i$ 与声源 s_j 位置矢量可以近似看成平行矢量。此时信号到达两个麦克风之间的幅度衰减差异很小,可以认为近似相等。

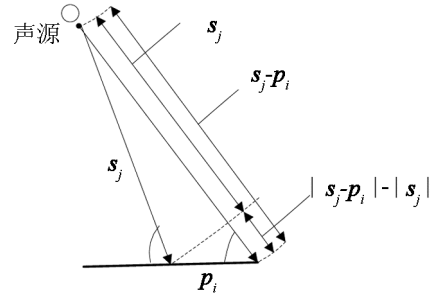


图 6 定位系统的远场模型
Fig.6 The far-field model for localization system

系统所采用的球形插值法根据多个麦克风的时延求得一组方程,并在满足最小均方误差准则下解方程组。

麦克风阵列由 $N+1$ 个位于 $r_i = (x_i, y_i, z_i)^T$ 的麦克风阵列组成。麦克风 $r_0 = (0, 0, 0)^T$ 为参考坐标。声源位于 $r_s = (x_s, y_s, z_s)^T$ 。各麦克风和声源到原点的距离为 R_i 和 R_s , 各个麦克风 R_i 与参考麦克风 r_0 的声程差用 d_i 来表示, 见图 7。

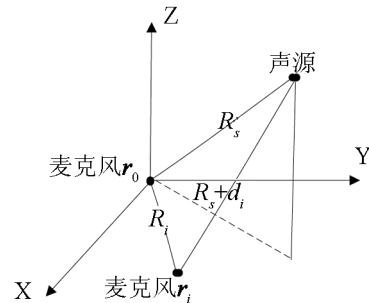


图 7 麦克风与声源几何位置关系图
Fig.7 The geometry location between microphones and source

由图 7 的几何关系可得:

$$\|r_i - r_s\| - \|r_s\| = d_i \quad (6)$$

将 $R_i = \|r_i\|, R_s = \|r_s\|$ 代入式(6), 两边平方可得:

$$(R_s - d_i)^2 = R_i^2 - 2r_i^T r_s + R_s^2 \quad (7)$$

整理并展开得到:

$$R_i^2 - d_i^2 - 2r_i^T r_s - 2R_s d_i = 0 \quad (8)$$

由于 d_i 是通过时延估计得到的, d_i 与实际值有偏差, 因此式(8)存在误差 ε 。

令 $\delta = (R_i^2 - d_i^2)$, 并改成矩阵形式, 则:

$$\varepsilon = \delta - 2R_s d - 2rr_s \quad (9)$$

其中:

$$\delta = \begin{bmatrix} R_1^2 - d_1^2 \\ R_2^2 - d_2^2 \\ \vdots \\ R_N^2 - d_N^2 \end{bmatrix}, d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_3 \end{bmatrix}, r = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{bmatrix}$$

假定给定 R_s , 则在最小均方准则下, 当满足式(10)时, 式(9)的均方误差最小。

$$r_s = 1/2 S_w^* (\delta - 2R_s d), S_w^* = (r^T r)^{-1} r^T \quad (10)$$

将式(10)代入 $R_s^2 = r_s^T r_s$, 并且化简可以得到:

$$aR_s^2 + bR_s + c = 0 \quad (11)$$

其中

$$\begin{aligned} a &= 4 - 4d^T (S_w^*)^T S_w^* d, \\ b &= 4d^T (S_w^*)^T S_w^* \delta, \\ c &= -\delta^T (S_w^*)^T S_w^* \delta \end{aligned} \quad (12)$$

解方程并将正根代入就可以得到声源位置为:

$$R_s = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, r_s = \frac{1}{2} S_w^* (\delta - \frac{d\sqrt{b^2 - 4ac} - b}{a}) \quad (13)$$

4 仿真结果分析

在本定位系统中, 采用如图 8 所示的五元麦克风阵列, 以麦克风 0 为中心建立直角坐标系, 其余四个麦克风坐标(单位: cm)分别为 $m_1(40,0,0)$ 、 $m_2(-40,0,0)$ 、 $m_3(0,40,0)$ 、 $m_4(0,0,40)$ 。该系统由鸣笛声识别、时延估计和几何定位三个模块组成。下面对这三个模块进行仿真以及结果分析。

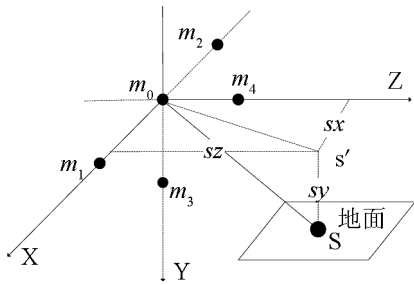


图 8 麦克风阵列分布
Fig.8 Layout of microphone array

4.1 鸣笛声识别

本系统的信号源为一段通过同步采样芯片得到的实际路面道路上混有噪声的汽车鸣笛声。该信号的采样率为 10kHz, 窗函数采用汉明窗, 帧长为 512 点, 帧移为 256 点。采用 3.1 节的方法对信号进行预处理, 结果如图 9 所示。

由短时能量检测出的信号, 对其鸣笛声特性集中的频谱区域(2~5kHz)做傅立叶变换提取统计量特征, 见表 2。

由表 2 的统计特性可以看出, 鸣笛声相对于一般的噪声有更高的方差、峰值比和峰度, 借此可以

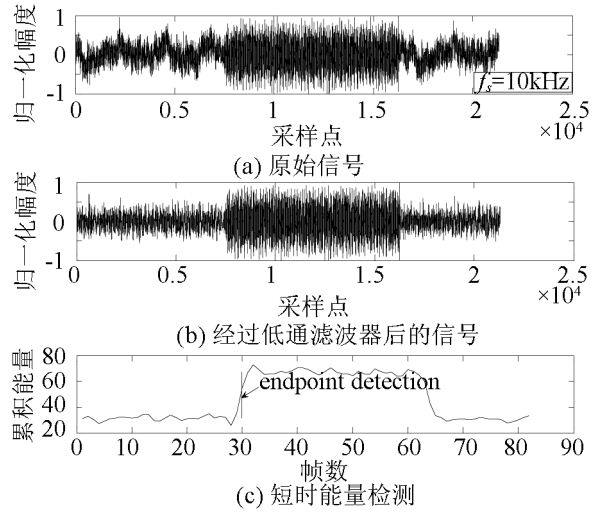


图 9 汽车鸣笛声识别
Fig.9 Car horn recognition

表 2 鸣笛声和噪声统计特性比较

Table 2 Statistical trait comparison between horn and noise

	均值	峰值	方差	峰值比	峰度
鸣笛声	0.4626	4.2768	0.5533	9.24	16.5755
噪声	0.1568	0.3656	0.0067	2.33	2.7807

设定一个阈值用以识别出鸣笛声。

4.2 时延估计

时延估计采用了 3.2.1 中所述的不同加权函数。在实际的系统仿真中发现, 在采用上述算法的基础上, 如果能够对多次不同段的数据的估计结果做平均, 则可以得到更高的增益。如图 10 所示, 以 PHAT 算法为例, 对六次时延估计的结果做平均之后, 估计结果的峰值和噪声之间的比值明显增大, 更容易得到正确的时延结果。

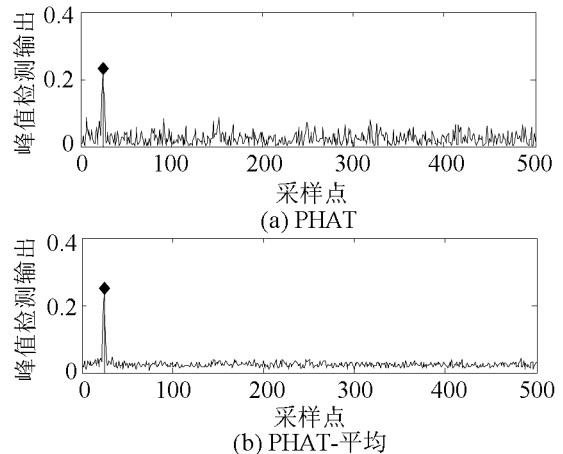


图 10 以 PHAT 为例, 平均六次的时延估计结果
Fig.10 Time delay estimation results for 6 times' average (PHAT)

下面是对不同的 GCC 加权函数进行的仿真结果, 各种结果都经过了平均以提高性能。由图 11 可以看到, PHAT 和 SCOT 性能相当, 优于 CC 和 ROTH。

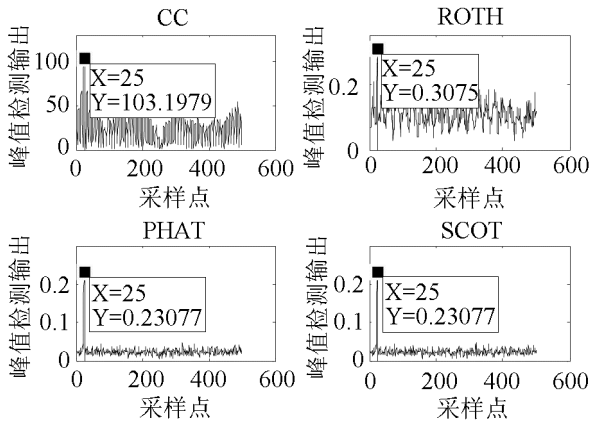


图 11 各种算法平均六次的时延估计结果
Fig.11 Time delay estimation results for 6 times' average (different algorithms)

4.3 声源定位算法

在考虑声源定位算法时，选择了球形插值法。实际上，除了球形插值法之外，还有其他各种定位方法。因此，在仿真中同时给出了另一种常用的一步最小二乘方法的定位结果作为比较。

图 12 分别是当声源在同一位置(300,1000,600)而时延估计的误差不同时，球形插值算法产生的相应误差。当时延偏差越大时，产生的误差越大。图 12 中时延偏差的单位是 $1/f_s$ (f_s 为采样频率)，对应前面时延估计算法的一个点的估计偏差。从图 12 可以看出，球形插值法在声源在此位置的情况下，具有更好的估计效果。

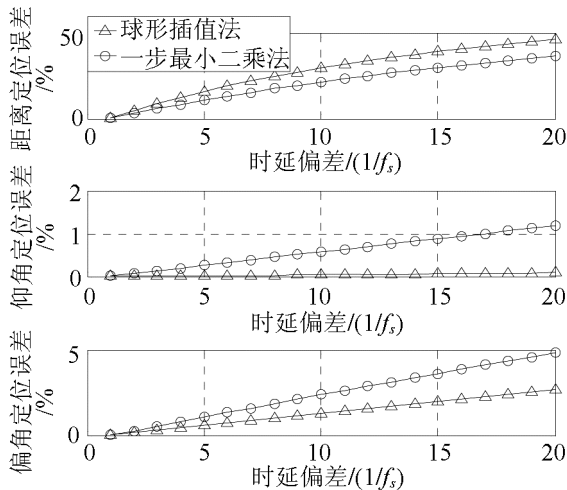


图 12 定位估计误差随时延偏差的变化
Fig.12 Location bias vs. time delay estimation error

图 13 为定位估计误差随声源距离的变化。从图 13 可以看出，球形插值法的估计误差随声源距离的变化较小，更适合移动车辆的定位。而且，采用球形插值法虽然对距离的估计误差稍大，但是所得到的仰角和偏角的误差都很小，更适合系统具体应用的场景。因为在汽车鸣笛声定位系统的设计中，定

位的主要目的是通过定位返回的结果，控制安装在监控路段的摄像头的角度，使其能够对准违章车辆，因此，对距离的估计准确度要求将会低于对角度的准确度要求。

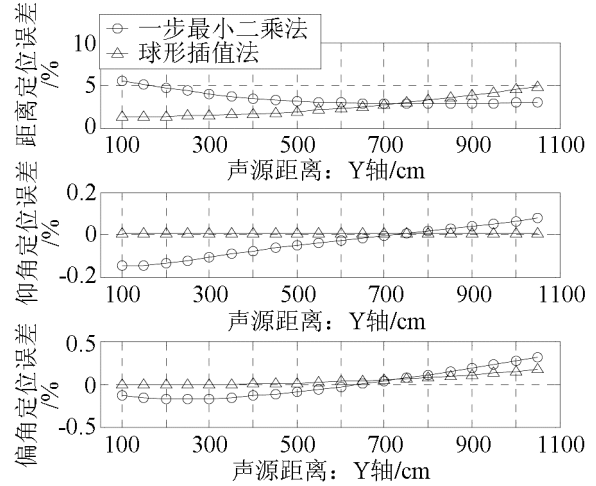


图 13 定位估计误差随声源距离的变化
Fig.13 Location bias vs. distance of car horn

5 结 论

本文提出一种基于麦克风阵列声源定位技术的城市道路汽车鸣笛声定位系统。该系统通过麦克风阵列获得汽车鸣笛声，根据鸣笛声的频谱特性对其进行识别，并通过基于时间延迟的声音定位技术，获得声源的位置信息，从而为执法提供了有效的依据。为了验证系统方案的有效性，本文对系统各个模块的算法进行了仿真，仿真结果表明该系统方案能够很好地满足计算量小和实时性需求，为以后的硬件实现提供了理论基础。

参 考 文 献

- [1] Rabinkin Daniel V, Richard J. Renomeron, Dahl A, Joseph C. French, James L. Flanagan, Michael H. Bianchi. A DSP implementation of source localization using microphone arrays[A]. In proceedings of the 131st meeting of the acoustical society of American[C]. 1996, 88-99.
- [2] Brandstein M S, Sliverman H F. A practical methodology for speech source localization with microphone arrays[J]. Computer, Speech and language, 1997, 11(2): 91-126.
- [3] Knapp C H, Carter G C. The generalized correlation method for estimation of time delay[J]. IEEE Trans. Acoustic Speech Signal Processing, 1976, 24(4): 320-327.
- [4] 杨祥清, 汪增福. 基于麦克风阵列的三维声源定位算法及其实现[A]. 声学技术, 2008, 27(2): 260-265.
YANG Xiangqing, WANG Zengfu. 3D sound source localization algorithm and its implementation based on microphone array[J]. Technical Acoustic, 2008, 27(2): 260-265.
- [5] Smith J, Abel J. Closed-form least-squares source location estimation from range-difference measurements[J]. IEEE Trans. Acoustic Speech Signal Processing, 1987, 35(12): 1661-1669.
- [6] 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, 22(1): 90-99.
CUI Weiwei, CAO Zhigang, WEI Jianqiang. Time delay estimation techniques in source location[J]. Journal of Data Acquisition & processing, 2007, 22(1): 90-99.