

基于麦克风阵列的语音增强研究

周述畅, 宋亚男, 吴光波

(广东工业大学 自动化学院, 广东 广州 510006)

摘 要: 麦克风阵列的自适应算法通过迭代运算获取波束形成的最优权矢量时, 噪声模型的估计是一个非常关键的因素。它的好坏直接影响着系统波束形成的性能。系统地分析了最小均方(LMS)自适应语音增强算法, 并针对阻塞矩阵在估计噪声时存在的缺陷, 在该算法的基础上提出了一种利用最小值控制递归平均(MCRA)来估计噪声的方法。将此方法应用于波束形成, 并用 Matlab 软件进行仿真。仿真实验结果表明, MCRA 估计出的噪声使 LMS 自适应语音增强的效果更好和抗噪性更强。

关键词: 麦克风阵列; 语音增强; LMS 自适应; 最小值控制递归平均

中图分类号: TP312

文献标识码: A

文章编号: 1673-629X(2012)07-0204-03

Speech Enhancement Research Based on Microphone Array

ZHOU Shu-chang, SONG Ya-nan, WU Guang-bo

(School of Automation, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: When adaptive algorithm of the microphone array uses iteration to obtain the optimal weight vector about beamforming, the noise model estimation is a critical factor. It will have a direct impact on the performance of beamforming system. Analyse systematically the least-mean-square (LMS) adaptive speech enhancement algorithms, and propose a method using a minima controlled recursive averaging (MCRA) to estimate the noise in allusion to the defects of the blocking matrix. This method is applied to beamforming, and using Matlab software to simulate. Simulation results show that, LMS-MCRA adaptive has robust speech enhancement.

Key words: microphone array; speech enhancement; adaptive least mean square algorithm; MCRA

0 引言

语音增强是指当语音信号被各种各样的噪声干扰, 甚至淹没后, 从噪声背景中提取有用的语音信号, 抑制、降低噪声干扰的技术^[1]。在各种电子设备高度智能化的今天, 语音增强与声源定位技术成为语音通信领域中两种不可缺少的技术。在视频会议、智能机器人、助听器、人机接口、通信和语音识别等领域, 都需要利用麦克风阵列进行声源定位, 定位之后再对语音进行增强。

人们针对加性宽带噪声研究了各种语音增强方法, 分别适用于不同情况。参数方法对语音的模型参数依赖性强, 在低信噪比条件下不容易得到正确的模型参数; 非参数方法由于频谱相减会产生具有一定节

奏的残余噪声^[2]; 统计方法需要大量的数据进行训练以得到统计信息; 小波变换^[3]以及离散余弦变换的阈值获取困难, 运算量大。实际应用中常常根据具体的环境噪声和语音特性将不同方法结合起来应用, 文中采用非参数方法、自适应噪声抵消法^[4], 同时在噪声估计时用最小值控制递归平均来估计^[5]。

1 自适应波束形成

基于麦克风阵列^[6]波束进行语音增强的方法大致上可分成三类: 延迟-累加波束法^[7]、自适应波束法^[8]和后置自适应滤波法。

目前语音增强研究的热点方法是基于模型参数自适应的噪声补偿算法。这类方法引入语音和噪声的统计知识, 提出具有一定环境稳健性的处理算法。自适应波束形成算法结构框图如图1所示。

图1中虚线部分为对广义旁瓣抵消器^[9](GSC)的改进, 加入了维纳滤波权矢量更新, 实现自适应地对空间变化的干扰环境作出反应。上支路形成目标检测通道, 含期望信号和干扰, 下支路形成辅助通道, 通过阻塞矩阵^[10] B 阻塞掉期望信号, 用其加权求和去预测检

收稿日期: 2011-11-14; 修回日期: 2012-02-21

基金项目: 广东省自然科学基金资助项目(06300261); 广东省大学生创新实验项目(1184510035); 广东省高等教育学会实验室专业委员会2010研究项目(2010033); 广东工业大学合生珠江大学生创新实验项目(HSZJ2011019)

作者简介: 周述畅(1984-), 男, 广东河源人, 硕士研究生, 研究方向为信号处理; 宋亚男, 副教授, 硕士生导师, 研究方向为噪声抵消法。

测通道中的干扰信号进而对消掉。要求下支路中不含目标信号。 \mathbf{B} 为信号阻塞矩阵,它不是唯一的,利用均匀线阵列^[11]各阵元之间存在着均匀相位差,构造阻塞矩阵。

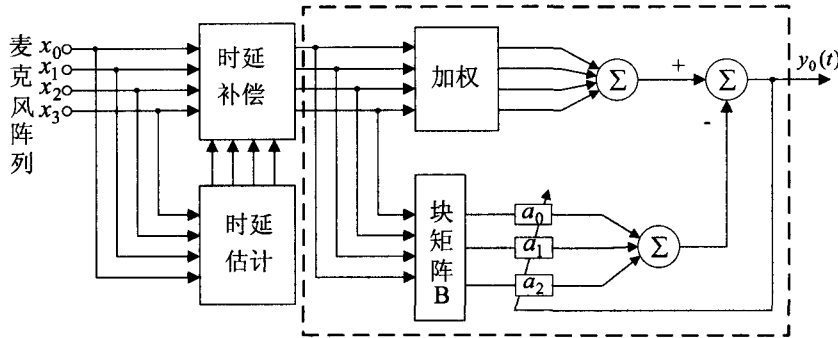


图1 自适应波束形成算法结构框图

$$\mathbf{B} = \begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & 1 & -1 & 0 \\ 0 & \cdots & 0 & 0 & 1 & -1 \end{bmatrix}_{(N-1) \times N}$$

在广义旁瓣抵消器中采用连续 LMS 算法计算下支路的自适应滤波权矢量 \mathbf{W}_x 。

GSC 的等效权矢量可写为

$$\mathbf{W}_{\text{GSC}} = \mathbf{a}(\theta_0) - \mathbf{B}_0^H \mathbf{W}_x = \mathbf{W}_q - \mathbf{B}_0^H \mathbf{W}_x \quad (1)$$

$\mathbf{W}_q = \mathbf{a}(\theta_0)$, 称为静态权矢量。满足 $\mathbf{B}\mathbf{a}(\theta_0) = 0$ 。

GSC 输出的均方误差

$$\begin{aligned} \text{MSE} &= E[|\varepsilon(n)|^2] \\ &= E[|d(n) - \mathbf{W}_x^H \mathbf{X}(n)|^2] \\ &= \sigma_d^2 - \mathbf{W}_x^H \mathbf{r}_{xd} - \mathbf{r}_{xd}^H \mathbf{W}_x + \mathbf{W}_x^H \mathbf{R}_x \mathbf{W}_x \end{aligned} \quad (2)$$

$$\sigma_d^2 = E[|d(n)|^2], \mathbf{r}_{xd} = E[\mathbf{X}(n)d^*(n)],$$

$$\mathbf{R}_x = E[\mathbf{X}(n)\mathbf{X}^H(n)]$$

对式(2)取对 \mathbf{W}_x 的梯度得

$$\nabla_{\mathbf{W}_x} = -2\mathbf{r}_{xd} + 2\mathbf{R}_x \mathbf{W}_x \quad (3)$$

要使得均方误差(MSE)下降最快,则

$$\mathbf{W}_x(n) = \mathbf{W}_x(n-1) + \frac{1}{2}\mu[-\nabla_{\mathbf{W}_x(n-1)}] \quad (4)$$

μ 为步长因子; $\nabla_{\mathbf{W}_x(n-1)}$ 可采用瞬时梯度估计代替:

$$\nabla_{\mathbf{W}_x(n-1)} = -2[\mathbf{X}(n-1)d^*(n-1) - \mathbf{X}(n-1)\mathbf{X}^H(n-1)\mathbf{W}_x(n-1)] \quad (5)$$

将(5)式代入(4)式得

$$\begin{aligned} \mathbf{W}_x(n) &= \mathbf{W}_x(n-1) + \mu\mathbf{X}(n-1)[d^*(n-1) - \mathbf{X}^H(n-1)\mathbf{W}_x(n-1)] \\ &= \mathbf{W}_x(n-1) + \mu\mathbf{X}(n-1)\varepsilon^*(n-1) \end{aligned} \quad (6)$$

将(6)式代入(1)得

$$\begin{aligned} \mathbf{W}_1 &= \mathbf{W}_q - \mathbf{B}^H \mathbf{W}_x(n) \\ &= \mathbf{W}_q - \mathbf{B}^H [\mathbf{W}_x(n-1) + \mu\mathbf{X}(n-1)\varepsilon^*(n-1)] \end{aligned}$$

$$\begin{aligned} &= \mathbf{W}_{x_1}(n-1) - \mu\mathbf{B}^H \mathbf{X}(n-1)\varepsilon^*(n-1) \\ &= \mathbf{W}_{x_1}(n-1) - \mu\mathbf{B}^H \mathbf{B} \mathbf{X}_1(n-1)\varepsilon^*(n-1) \end{aligned} \quad (7)$$

$$\varepsilon(n) = \mathbf{W}_{x_1}^H(n)\mathbf{X}_1(n) \quad (8)$$

初始状态为

$$\mathbf{W}_{x_1}(0) = \mathbf{W}_q \quad (9)$$

$$\varepsilon(0) = \mathbf{W}_q^H \mathbf{X}_1(0) \quad (10)$$

2 噪声估计

阻塞矩阵在估计噪声的时候存在缺陷,即语音中含有噪声的幅值较小时,就会使通过信号阻塞矩阵后含有大量的语音。这些语音接下来消弱期望语音信号,则增强

性能降低。解决的办法有,在无语音的信号段才估计噪声。然而这个方法不能有效地估计周期性噪声。下面介绍一种噪声估计方法,即极小值控制递归平均。

观测信号为:

$$y(n) = x(n) + d(n) \quad (11)$$

其中, $x(n)$ 和 $d(n)$ 分别是语音和非相关加性噪声。对式(11)做短时傅里叶变换:

$$Y(k, l) = \sum_{n=0}^{N-1} y(n + lM)h(n)e^{-j(2\pi/N)nk} \quad (12)$$

式中 k 是频率, l 是帧序列, h 是长度为 N 的窗函数。假设 $H_0(k, l)$ 和 $H_1(k, l)$ 分别表示语音缺失和语音存在,则:

$$\begin{aligned} H_0(k, l): Y(k, l) &= D(k, l) \\ H_1(k, l): Y(k, l) &= X(k, l) + D(k, l) \end{aligned} \quad (13)$$

式中 $X(k, l)$ 和 $D(k, l)$ 分别表示干净信号和噪声信号。让 $\lambda_d(k, l) = E[|D(k, l)|^2]$ 表示噪声变量。噪声测量时,用瞬时递归平滑得到它的估计值。则估计值:

$$\begin{aligned} H_0(k, l): \hat{\lambda}_d(k, l+1) &= \hat{\lambda}_d(k, l) \\ H_0(k, l): \hat{\lambda}_d(k, l+1) &= \alpha_d \hat{\lambda}_d(k, l) + (1 - \alpha_d) |Y(k, l)|^2 \end{aligned} \quad (14)$$

其中, $\alpha_d (0 < \alpha_d < 1)$ 是平滑系数。

$p'(k, l) = \Delta P(H_1(k, l) | Y(k, l))$ 表示信号存在的条件概率。

$$\begin{aligned} \hat{\lambda}_d(k, l+1) &= \hat{\lambda}_d(k, l)p'(k, l) + [\alpha_d \hat{\lambda}_d(k, l) + (1 - \alpha_d) |Y(k, l)|^2] \times (1 - p'(k, l)) \\ &= \tilde{\alpha}_d(k, l) \hat{\lambda}_d(k, l) + [1 - \tilde{\alpha}_d(k, l)] |Y(k, l)|^2 \end{aligned} \quad (15)$$

$$\tilde{\alpha}_d(k, l) = \Delta \alpha_d + (1 - \alpha_d)p'(k, l) \quad (16)$$

式(16)是时变平滑系数。

在一个频带的给定帧中,语音的存在由一个比值决定,该比值是含噪语音的能量与它的最小值之比。

含噪语音的能量为:

$$S_f(k, l) = \sum_{i=-\omega}^{\omega} b(i) |Y(k-i, l)|^2 \quad (17)$$

由一阶递归平均实现平滑化, 表示为

$$S(k, l) = \alpha_s S(k, l-1) + (1 - \alpha_s) S_f(k, l) \quad (18)$$

其中, $\alpha_s (0 < \alpha_s < 1)$ 是一个系数。能量的最小值 $S_{\min}(k, l)$ 和临时变量 $S_{\text{tmp}}(k, l)$ 由 $S_{\min}(k, 0) = S(k, 0)$ 和 $S_{\text{tmp}}(k, 0) = S(k, 0)$ 初始化。

此时的能量和前一帧的最小值比较得到当前帧的最小值:

$$\begin{aligned} S_{\min}(k, l) &= \min\{S_{\min}(k, l-1), S(k, l)\} \\ S_{\text{tmp}}(k, l) &= \min\{S_{\text{tmp}}(k, l-1), S(k, l)\} \end{aligned} \quad (19)$$

当 l 达到 L 帧时, 临时变量初始化为:

$$\begin{aligned} S_{\min}(k, l) &= \min\{S_{\text{tmp}}(k, l-1), S(k, l)\} \\ S_{\text{tmp}}(k, l) &= S(k, l) \end{aligned} \quad (20)$$

让 $S_r(k, l) = \Delta S(k, l) / S_{\min}(k, l)$ 表示含噪语音能量和得到的最小值之间的比率。

式(16)中 $p'(k, l)$ 的估计值为:

$$p'(k, l) = \alpha_p p'(k, l-1) + (1 - \alpha_p) I(k, l) \quad (21)$$

其中, $\alpha_p (0 < \alpha_p < 1)$ 是平滑系数, $I(k, l)$ 是指标函数, 如果 $S_r(k, l) > \delta$, 则 $I(k, l) = 1$, 否则 $I(k, l) = 0$, δ 是设定的阈值。

最后通过式(15)求得噪声功率。

3 算法仿真及性能分析

实验工具: matlab, labview, DAQ 助手, 实验条件是: 在 13×8 的室内, 由吊扇产生的噪声。麦克风的构成阵元数目为 $M=4$ 的线性阵列, 阵元数目为 $M=4$, 阵元间距 $d=0.25\text{m}$ 。然后通过 BNC-2120 接线盒和 PCI6220 板卡, 最后用 LabVIEW 编程采集并存储。

经过自适应波束形成方法来实现语音增强, 通过改变噪声估计模型来观察实验结果, 当采用阻塞矩阵 B 来估计噪声时, 其 MATLAB 仿真结果如图 2 所示; 当采用 MCRA 来估计噪声时, 其 MATLAB 仿真结果如图 3 所示。

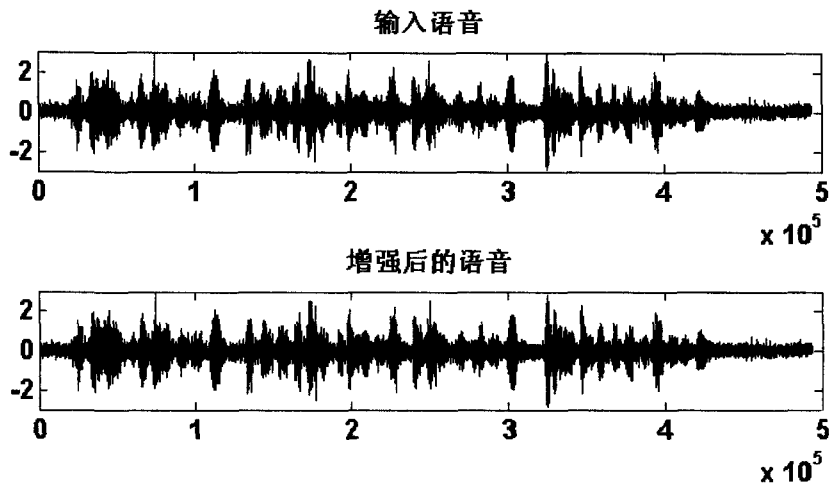


图 2 阻塞矩阵估计噪声时仿真结果

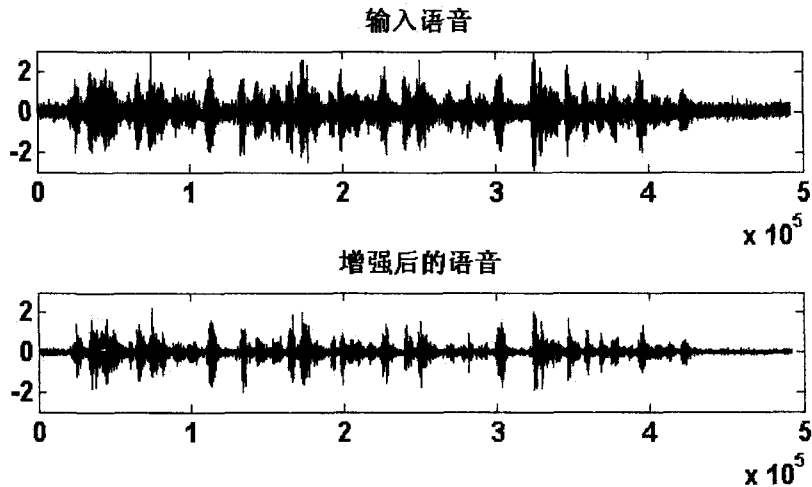


图 3 MCRA 估计噪声时仿真结果

对比图形可以看出自适应波束形成时使用 MCRA 噪声估计有非常明显的效果。

4 结束语

文中从目前常用的语音增强方法谈起, 重点讲解自适应语音增强方法。介绍了自适应滤波器的组成原理、实现自适应滤波最陡下降算法^[12]。在自适应的基础上, 针对阻塞矩阵存在的缺陷, 提出了 MCRA 噪声估计的方法。

实验结果表明, 在相同环境噪声条件下, MCRA 噪声估计的方法比阻塞矩阵更优。

参考文献:

- [1] 刘钰, 马艳丽, 董蓓蓓. 语音增强技术及算法综述[J]. 电脑编程技术与维护, 2010(16): 88-93.
- [2] 索忠伟. 语音增强技术研究及其 DSP 实现[D]. 成都: 西南交通大学, 2006.
- [3] 李如玮, 鲍长春, 樊慧晶. 基于小波变换的语音增强算法综

(下转第 210 页)

验证后模块符合设计要求,能够完成标准的 PCI 操作和错误处理,满足时序要求。

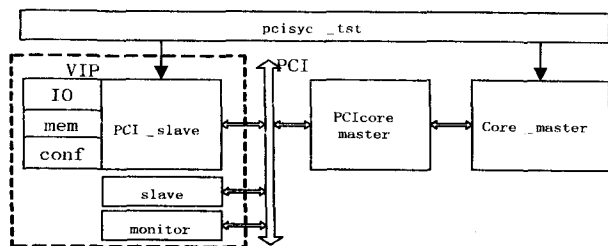


图 4 模块验证结构

2.2 FPGA 验证

将本设计连接到一款 SPARC V8 处理器后完成了 FPGA 验证工作^[10,11]。如图 5 所示,验证主要包括两个 FPGA,分别为 xilinx 公司的 v5 和 v2 芯片,其中 v5 实现处理器和 Master 模块,v2 实现 PCI Target 模块。v5 中包含 PCIMaster(挂载在处理器片内 AHB 总线上)、处理器和 PCI 仲裁。v2 中的 Target 模块包含 PCI Target 功能模块和挂载在 Target 上的 memory1 和 memory2,其中 memory1 为 PCI memory 操作中对应的空间, memory 2 为 IO 操作对应的空间。v2 和 v5 通过标准的 PCI 总线进行通信,来验证 Master 模块的正确性。

验证后模块能实现 CPU 对 PCI 从设备的正常读写和错误处理,时序符合要求,完成设计目标。

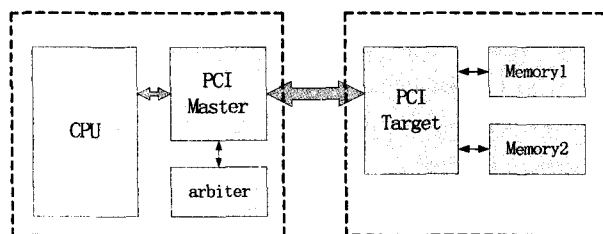


图 5 FPGA 验证结构

3 结束语

作为一种被广泛应用的总线,PCI 是在 CPU 和原来的系统总线之间插入的一级总线,具体由一个桥接

电路实现对这一层的管理,并实现上下之间的接口以协调数据的传送^[12]。

文中介绍了一种新颖的 PCI Master 接口模块设计,该设计符合标准的 PCI 2.2 协议和 AMBA 总线协议 2.0。负责将 CPU 内部 AHB 总线上的标准 AHB 信号转化成为标准 PCI 信号,并完成 PCI 总线操作和异常处理,实现 CPU 对 PCI 从设备的控制功能。通过模块验证和 FPGA 验证,保证了模块能实现设计需要的所有功能。

参考文献:

- [1] 李贵山,威德虎. PCI 局部总线开发者指南[M]. 西安:西安电子科技大学出版社,1997.
- [2] 戴紫彬. PCI 9054 局部总线设计及应用[J]. 微电子学与计算机,2003(8):122-124.
- [3] 陈露晨. PCI 9054 性能分析及外部 FIFO 的扩充[J]. 集成电路应用,2001(4):68-69.
- [4] 王 芃. 基于 PCI 总线位置控制系统的接口卡实现[J]. 实验技术与管理,2010,27(5):97-99.
- [5] PCI Special Interest Group. PCI Local Bus Specification Revision 2.2[S]. PCI Special Interest Group,1998.
- [6] PCI 9054 Data book PLX Technology Inc[S]. 1999.
- [7] 顾 超. 总线式数控系统中 PCI 接口控制器的 FPGA 实现[J]. 计算机应用,2011,31(2):565-567.
- [8] 李 攀. 基于 SOC 的 PCI 通信接口与实现[J]. 计算机技术与发展,2009,19(9):211-214.
- [9] 黄志强. Xilinx 可编程逻辑器件的应用与设计[M]. 北京:机械工业出版社,2007.
- [10] Bohm P. Incremental and Verified Modeling of the PCI Express Protocol[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2010, 29(10): 1495-1508.
- [11] 单天昌. 基于 FPGA 的 PCI 接口 DMA 设计与实现[J]. 计算机技术与发展,2010,20(4):215-219.
- [12] 师于杰. 高可用 CompactPCI 平台在串口通信上的应用[J]. 计算机技术与发展,2006,16(4):203-205.

(上接第 206 页)

- [4] Jeong J, Moir T J. A real-time kepsrum approach to speech enhancement and noise concellation[J]. Neurocomputing, 2008, 71(13-15):2635-2649.
- [5] Cohen I, Berdugo B. Noise Estimation by Minima Controlled Recursive Averaging for Robust Speech Enhancement[J]. IEEE Signal Processing Letters, 2002, 9(1):12-15.
- [6] 许佳龙,张一闻,刘建平. 基于麦克阵列的信号采集处理系统设计[J]. 电子科技,2011,24(9):124-127.
- [7] 王桂宝,林中朝,王兰美,等. LMS 自适应波束形成方法研究[J]. 电子科技,2011,24(7):103-105.
- [8] 杜 军,桑胜举. 基于麦克风阵列的语音增强技术及应用

[J]. 计算机应用与软件,2009,26(10):75-77.

- [9] Griffiths L J, Jim C W. An alternative approach to linearly constrained adaptive beamforming[J]. IEEE Trans. on Antennas Propag., 1982, 30(1):27-34.
- [10] Manolakis D G, Ingle V K, Kogon S M. Statistical and adaptive signal processing-spectral estimation, signal modeling, adaptive filtering and array processing[M]. Beijing:Tsinghua University Press, 2003.
- [11] 李 平,史小卫. 一种均匀直线阵列干扰抑制的新方法[J]. 西安电子科技大学学报,2004,31(4):523-525.
- [12] 左全生. 自适应滤波器的一种新 LMS 算法-改进的最陡下降法[J]. 常州工学院学报,2000,13(4):34-37.