

基于差分麦克风阵列的自适应语音增强算法研究及 DSP 实现

宋辉¹ 刘加¹

摘要 自适应滤波是语音增强算法中的常用技术,而算法复杂度与收敛速度是设计各种自适应算法需要首要考虑的问题.本文提出一种用于片上的语音增强自适应滤波新算法.该算法分两步实现,首先,利用一阶差分麦克风阵列,获得噪声的实时估计;其次,对传统的仿射投影算法(Affine projection algorithm, APA)加以改进,得到计算误差向量的快速算法,并根据估计误差动态调整搜索步长以及仿射投影维数,对带噪语音进行自适应滤波消噪.在 TMS320VC5509 DSP 芯片上实现该算法.实验表明,算法的自适应滤波过程具有接近递推最小二乘算法(Recursive least squares, RLS)的快速收敛速度,以及类似最小均方误差算法(Least mean squares, LMS)的低算法复杂度.

关键词 语音增强, 差分麦克风阵列, 仿射投影算法, 自适应滤波, 噪声消除

中图分类号 TN912

Research on Adaptive Speech Enhancement Algorithm Based on Differential Microphone Array and Its Implementation with DSP

SONG Hui¹ LIU Jia¹

Abstract Adaptive filtering is in common use in most speech enhancement algorithms, while complexity and convergence speed should be considered first when an adaptive filtering method is designed. This paper presents a speech enhancement adaptive filtering method implemented on chip. This method is implemented in two steps. In the first step, first-order differential microphone array is utilized to obtain real-time noise estimation. In the second step, the traditional affine projection algorithm (APA) is modified so that a fast error vector calculation method is obtained, and the search step and projection dimension can be altered dynamically according to the estimation error. Thus, noise can be removed adaptively. This method is realized on a TMS320VC5509 DSP chip. Experiments show that the proposed method has fast convergence speed like RLS (Recursive least squares) and low computational complexity like LMS (Least mean squares).

Key words Speech enhancement, differential microphone array, affine projection algorithm (APA), adaptive filter, noise reduction

随着语音信号数字处理技术的发展,各种语音处理算法的片上实现已经成为人们研究的热点.然而,实际中复杂的

环境噪声会对各种语音芯片的处理性能产生影响.因此,将语音增强作为预处理手段集成在语音芯片中是很有必要的.

麦克风阵列具有空间指向性,能够从声源方向提取高品质的语音信号.目前已经形成了众多成熟的语音增强技术,如波束形成算法(Beamforming)^[1],广义旁瓣对消(Generalized sidelobe canceller)^[2],广义奇异值分解(Generalized singular value decomposition)^[3]等.这些方法虽然可以达到比较好的增强效果,但往往需要较多数目的麦克风.一阶差分麦克风阵列由两个全向麦克风组成,麦克风间距较小,具有成本低廉,结构简单,易于片上实现等优点,在很多小型化、便携式语音通信产品中有着广泛的应用.本文的自适应语音增强算法就是以一阶差分麦克风阵列为系统平台设计的.

自适应滤波仍然是应用最广泛、最有效的语音增强手段之一,目前大多数双麦克风语音增强系统也都是采用自适应的方法.最小均方误差(Least mean squares, LMS)算法是最常见的自适应滤波方法,该算法原理简单,便于实时实现,缺点是收敛速度慢,动态跟踪性能差;递推最小二乘(Recursive least squares, RLS)算法对估计误差采用指数加权,收敛速度和跟踪性能大大优于 LMS 算法,所付出的代价是要采用矩阵求逆递推,计算量太大^[4];仿射投影算法(Affine projection algorithm, APA)是近几年来发展起来的一种性能和复杂度介于 LMS 和 RLS 之间的自适应滤波算法,多用于语音增强、回声对消等场合,目前也产生了相应的快速算法^[5],但这些算法只专注于计算量的节约,而忽略了算法收敛性能的改善,仍然没有达到片上实用的标准.本文提出一种变步长、变投影维数的仿射投影(Multi-step multi-dimension affine projection, MMAP)算法,该算法随着估计误差的变化动态调整搜索步长和投影维数.实验证明,这种算法能够在算法复杂度和收敛速度之间实现很好的折中.

本文的结构如下:第 1 节介绍一阶差分麦克风阵列结构,提出一种基于“零点输入-低频补偿”的噪声估计方法;第 2 节详细描述 MMAP 算法,并给出具体实现框架;第 3 节简要介绍硬件实验平台结构;第 4 节为软、硬件实验结果,验证了 MMAP 算法相对于传统自适应滤波算法的优越性;第 5 节给出结论.

1 一阶差分麦克风阵列

1.1 差分麦克风阵列结构

标准的一阶差分麦克风阵列如图 1 所示.两麦克风间距为 d .可以证明,该系统的频率-波数响应函数可以表示为^[6]:

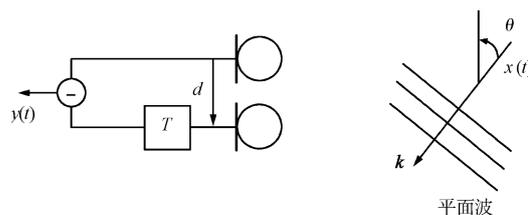


图 1 一阶差分麦克风阵列系统框图

Fig. 1 Diagram of first-order differential microphone array

收稿日期 2008-06-10 收修稿日期 2009-01-19
Received June 10, 2008; in revised form January 19, 2009
国家高技术研究发展计划(863 计划)(2006AA010101, 2007AA04Z223)
和自然科学基金委员会与微软亚洲研究院联合资助项目(60776800)资助
Supported by National High Technology Research and Development Program of China (863 Program) (2006AA010101, 2007AA04Z223) and National Natural Science Foundation of China and Microsoft Research Asia (60776800)

1. 清华大学电子工程系清华大学科学与技术国家实验室(筹)北京 100084
1. Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084

DOI: 10.3724/SP.J.1004.2009.01240

$$|H(\omega, \theta)| = \left| \frac{Y(\omega, \theta)}{X(\omega)} \right| = \left| 1 - e^{-j(\omega T + kd)} \right| = 2 \sin \frac{\omega \left[T + \frac{(d \cos \theta)}{c} \right]}{2} \quad (1)$$

其中, c 表示声速, d 是间距变量, k 为波数变量, 其幅值 $k = \omega/c$ 为通常意义下的波数值, θ 为方向角, $Y(\omega, \theta)$ 和 $X(\omega)$ 分别为输出信号 $y(t)$ 和输入信号 $x(t)$ 的傅立叶变换. 对于差分麦克风阵列, 麦克风的间距较小, 满足近似关系 $kd \ll \pi$ 以及 $\omega T \ll \pi$, 因此式 (1) 可以简化为

$$|H(\omega, \theta)| = \omega \left[T + \frac{(d \cos \theta)}{c} \right] \quad (2)$$

其归一化波束方向性图如图 2 所示.

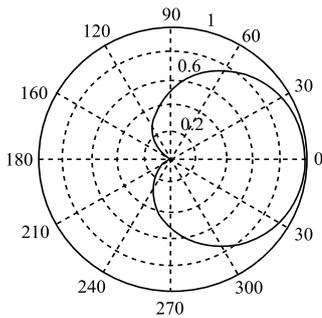


图 2 一阶差分麦克风阵列归一化方向性图

Fig. 2 Normalized amplitude response of first-order differential microphone array

1.2 噪声估计

通过估计背景噪声, 作为自适应滤波的参考信号, 是自适应增强算法的实质. 从波束方向性图中可以看出, 由于声音到达两个麦克风之间存在声程差, 使得阵列在 180° 方向出现一个零点 (Null), 这个零点是由两路信号相减得到的. 一般而言, 如果预先知道噪声的方向, 可以将该零点对准噪声方向, 此时阵列输出即为增强后的语音. 然而实际应用中, 噪声的方向是无法预知的, 而且通常是充满整个空间的. 因此, 本文采用“零点输入-低频补偿”的噪声估计方法, 即语音从零点方向输入, 阵列输出则不含语音信号, 只有纯噪声信号, 并以此作为噪声估计的雏形, 这就是“零点输入”的含义.

由式 (2) 可以看出, 阵列的频率-幅度响应与输入频率近似呈线性关系. 事实上, 阵列输出与输入的关系可以表示为

$$y(t) = x(t) - x(t - T) \quad (3)$$

这相当于输入信号通过高通滤波器 $H(\omega) = 1 - e^{-j\omega T}$ 的结果. 因此, 输出噪声信号的低频分量有较大幅度的失真, 无法反映真实环境噪声的特性. 引入低频补偿技术, 将阵列输出信号通过逆滤波器:

$$H_{LFC}(\omega) = \frac{1}{H(\omega)} = \frac{1}{1 - \alpha e^{-j\omega T}} \quad (4)$$

得到真实环境噪声估计. 其中, 因子 $\alpha > 0$ 的引入是为了避免逆滤波器在 $\omega = 0$ 处出现极点而产生的低频失稳问题.

2 MMAP 自适应滤波算法

2.1 仿射投影算法 APA

带有松弛项的标准仿射投影算法可以表述为

$$\mathbf{e}_n = \mathbf{y}_n - X_n^T \boldsymbol{\omega}_{n-1}^* \quad (5)$$

$$\Delta \boldsymbol{\omega}_n = X_n \left[X_n^T X_n + \delta I \right]^{-1} \mathbf{e}_n^* \quad (6)$$

$$\boldsymbol{\omega}_n = \boldsymbol{\omega}_{n-1} + \mu \Delta \boldsymbol{\omega}_n \quad (7)$$

其中, $\boldsymbol{\omega} = [\omega_0, \omega_1, \dots, \omega_{M-1}]^T$ 为 M 维滤波器系数向量, $\mathbf{x}_n = [\mathbf{x}_n, \mathbf{x}_{n-1}, \dots, \mathbf{x}_{n-M+1}]^T$ 为 n 时刻 M 维输入信号, $\mathbf{y}_n = [\mathbf{y}_n, \mathbf{y}_{n-1}, \dots, \mathbf{y}_{n-L+1}]^T$ 为 L 维输出信号, L 是投影维数, $X_n = [\mathbf{x}_n, \mathbf{x}_{n-1}, \dots, \mathbf{x}_{n-L+1}]^T$ 为与之对应的 $M \times L$ 维输入信号矩阵, \mathbf{e}_n 为 L 维残差向量, 本文中为增强后的语音信号, $*$ 号表示共轭, δI 为松弛项, 其中, δ 是一个很小的正数, 用以避免矩阵求逆时发生奇异, μ 为搜索步长. 通常要求 $L < M$. 如果 $L = 1$, 则 APA 算法退化为 LMS 算法; 如果 $L \geq M$ 则 APA 算法与 RLS 算法等价.

2.2 多重投影维数算法

LMS 算法收敛性能差有两个主要原因: 步长 μ 不容易选择得很适当; 滤波器输入信号可能是相关信号. 针对后者, 目前已有很多时频域解相关的 LMS 算法能够降低输入信号的相关程度, 加快收敛速度^[4], 而仿射投影算法也是解相关的一种方式. 不过, 通常 APA 算法采用一致的投影维数, 在收敛速度和算法复杂度的权衡上往往顾此失彼. 本文采用具有多重投影维数的仿射投影算法, 在初始阶段, 为了加速收敛, 采用较大投影维数; 在跟踪阶段, 选择较小的投影维数, 在保证实时跟踪的前提下降低运算量. 该算法根据估计误差动态调整投影维数 L , 而估计误差可由 APA 算法的定义式快速递推得到.

式 (5) 所示的误差也称为先验误差, 写成矩阵的形式:

$$\mathbf{e}_n = \mathbf{y}_n - X_n^T \boldsymbol{\omega}_{n-1}^* = \begin{bmatrix} \mathbf{y}_n - \mathbf{x}_n^T \boldsymbol{\omega}_{n-1}^* \\ \tilde{\mathbf{y}}_{n-1} - \tilde{X}_{n-1}^T \boldsymbol{\omega}_{n-1}^* \end{bmatrix} \quad (8)$$

其中, \tilde{X}_{n-1} 由 X_{n-1} 的最左边 $L-1$ 列, 即 X_n 的最右边 $L-1$ 列组成, $\tilde{\mathbf{y}}_{n-1}$ 由列向量 \mathbf{y}_{n-1} 的最上面的 $L-1$ 个元素组成. 定义后验误差^[7]:

$$\mathbf{e}_{p,n-1} = \mathbf{y}_{n-1} - X_{n-1}^T \boldsymbol{\omega}_{n-1}^* = \left[I - \mu X_{n-1}^T X_{n-1} (X_{n-1}^T X_{n-1} + \delta I)^{-1} \right] \mathbf{e}_{n-1} \quad (9)$$

其中, 下标 p 表示“后验”的含义. 由式 (8) 和 (9) 可以看出, 先验误差的后 $L-1$ 个元素恰好对应后验误差的前 $L-1$ 个元素. 考虑到 δ 是一个很小的正数, 忽略松弛项的影响, 得到先验误差的快速更新公式:

$$\mathbf{e}_n = \begin{bmatrix} \mathbf{e}_n \\ (1 - \mu) \tilde{\mathbf{e}}_{n-1} \end{bmatrix} \quad (10)$$

事实上, 利用矩阵的相似变换可以得到与式 (10) 相同的结论^[7]. 误差的大小关系着投影维数是否被更新, 而每一步迭代的新的误差向量又和上一步的误差向量紧密相关, 因此大大节省了计算量 and 处理时间. 为了衡量误差向量的“大小”,

定义误差向量的 L_p 范数:

$$\|e_n\|_p = \left[\sum_{i=1}^L |e_{n,i}|^p \right]^{\frac{1}{p}} \quad (11)$$

不同的 L_p 范数具有不同的代数意义,但在有限维代数空间中具有等价性. 滤波初始阶段,残留误差较大;进入跟踪阶段,误差逐渐变小. 因此, $\|e_n\|_p$ 可以用来标识滤波进行的程度. 为了使算法针对不同特性、不同幅度的噪声具有普适性,本文采用归一化 L_p 范数 $\|e_n\|_p/\|e_n\|_1$ 作为衡量指标. 选取门限向量 $t_l = [t_{l1}, t_{l2}, \dots]^T$, $|t_{li}| < 1$, $t_{li} < t_{lj} \forall i, j, i > j$. 门限向量元素按绝对值降序排列,将 $(0, 1)$ 区间划分为若干子区间,根据误差范数落入的区间不同,更新投影维数. 门限向量的维数即为可变投影维数的个数,一般选取 3 个投影维数就足够了.

2.3 可变步长投影算法

搜索步长是影响算法收敛性能的另一主要因素. 为了保证仿射投影算法的收敛性能,搜索步长 μ 须满足一定的范围,通常选取 $0 < \mu < 2$. 本文中,步长的动态选择与投影维数的选择相似,同样是基于估计误差准则. 由于搜索步长与误差变化速率直接相关,因此定义误差变化率:

$$g_{m,n} = \frac{\|e_m\|_p - \|e_n\|_p}{m - n} \quad (12)$$

其中, m 和 n 为语音帧的标号. 误差变化率体现了误差包络的变化速率,可用以指示搜索步长选择是否恰当. 采用归一化误差变化率 $g_{n-c,n}/g_{1,2}$ 作为步长选择的指标,其中, c 是大于 1 的整数. 式 (12) 体现出滤波误差的实时变化率,且随着滤波过程的进行逐渐变缓. 对 c -邻域内的误差变化率序列加以平滑,避免因个别点出现奇异而引起算法的不稳定. 选取门限向量 $t_g = [t_{g1}, t_{g2}, \dots]^T$, $g_{n-c,n} < 0$, 门限向量元素同样按降序排列,满足 $t_{gj} < 1$, $t_{gi} < t_{gj}, \forall i, j, i > j$. 根据实际误差变化率与门限的关系,选择合适的搜索步长.

由于之前选择投影维数时已经计算出了误差的 L_p 范数,因此此处不会增加额外的计算复杂度,只需多开辟一些存储空间用来存放 n 时刻之前的误差范数序列, c 平滑邻域大小即为新开辟空间大小.

2.4 自适应语音增强系统框架

在自适应滤波过程中,确定滤波器系数更新的时刻通常需要引入语音端点检测 (Voice activity detection, VAD) 算法,由于输入信号含有噪声,基于“能量-过零率”的端点检测算法不再适用. 本文采用频域软判决^[8]的方法,利用语音和噪声在频域中微分瓣的差异,消除噪声对 VAD 判决性能的影响. 当只有噪声信号存在时,滤波器系数才被更新. 实验表明, VAD 的引入对自适应算法的增强性能具有积极作用. 自适应增强流程如图 3 所示.

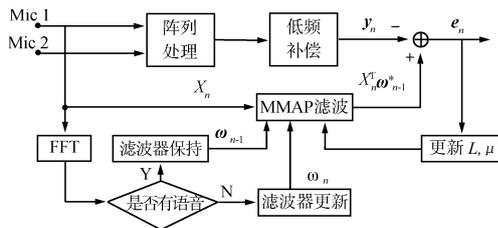


图 3 自适应语音增强系统框图

Fig. 3 Diagram of adaptive speech enhancement system

3 硬件平台

DSP 选型需考虑运算速度、成本、功耗、硬件资源和程序可移植性等因素. 本文采用美国德州仪器 (TI) 的 TMS320 VC5509 定点 DSP 作为核心处理器,配合使用 PLL 时钟发生器、JTEG 标准测试接口、异步通信串口、DMA 控制器、通用输入输出 GPIO 端口以及多通道缓冲串口 (McBSPs) 等主要片内外设. 5509DSP 芯片采用多总线结构,内含 $64 \text{ K} \times 16 \text{ bit}$ 的片上 RAM 和 64 KB 的 ROM;片内可屏蔽 ROM 固化有引导装载程序 (Bootloader) 和中断向量表;为扩充存储空间,外接三菱公司 M5M29GB/T320VP 系列 Flash 芯片. 全片容量 2 MW ,分为 128 个扇区,通过外部存储器接口 (EMIF) 方式与读写时序接入 DSP;采用流水线结构提高指令执行的整体速度;包含统一的数据、程序空间和 I/O 空间,寻址空间可达 16 MB ;片内包含两个算术逻辑单元 (ALU),在最高时钟频率 200 MHz 下,指令周期可达 5 ns ,最高速度为 400 MIPS . DSP 硬件仿真流程如图 4 所示.

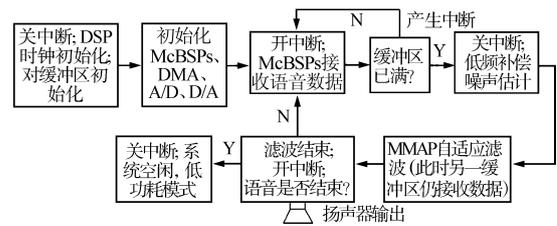


图 4 MMAP 算法的 DSP 硬件实现原理框图

Fig. 4 Diagram of hardware realization of MMAP

4 实验结果

采用微分麦克风阵列作为输入. 麦克风间距 $d = 3 \text{ cm}$, 信号采样频率为 8 kHz , 处理帧长为 256, 帧移为 128. 噪声为实际环境下录制的汽车噪声. 图 5 为增强前后语音波形和语谱图, 表 1 列出了该系统对客观输入信噪比以及主观 Mos (Mean opinion score) 分数的改善. 可以看出, 在 SNR 改善方面, RLS 改善效果最明显, 原因在于 RLS 算法收敛速度最快, 相比之下, LMS 算法的改善能力在 4 种算法当中最弱. 而 MMAP 最接近 RLS 的结果: 在 Mos 分数上, 几种自适应算法的改善结果相差不多.

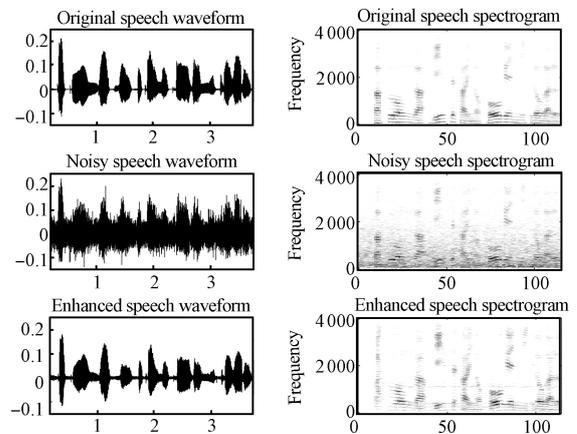
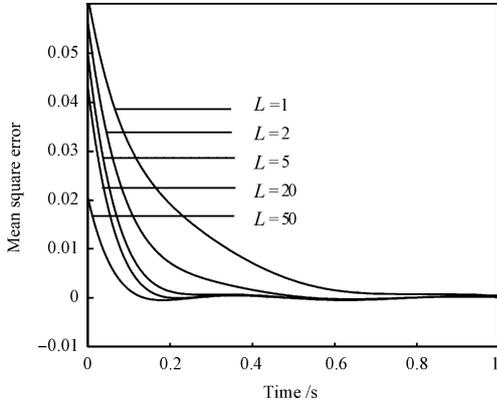


图 5 MMAP 增强前后语音波形与语谱图对比结果

Fig. 5 Comparison of speech waveforms and spectrograms before and after MMAP

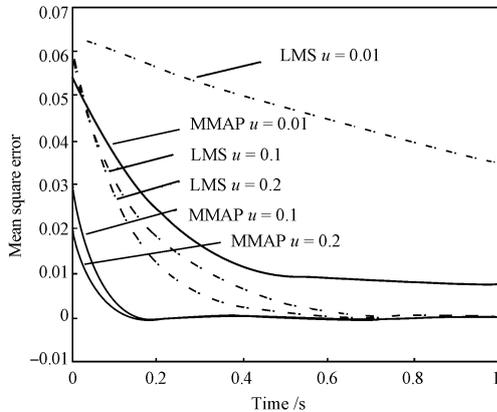
表 1 信噪比 (dB) 及 Mos 分数 (/5) 改善结果
Table 1 Improvement of SNR (dB) and Mos (/5)

原始 SNR /Mos	MMAP 后 SNR/Mos	LMS 后 SNR/Mos	APA 后 SNR/Mos	RLS 后 SNR/Mos
-7.77/2.4	8.60/3.7	7.92/3.7	8.51/3.5	9.21/3.6
-2.40/2.9	13.98/3.9	13.50/3.9	13.87/3.8	14.37/3.9
3.91/3.2	19.58/3.9	18.88/3.8	19.38/4.0	20.65/4.0
9.97/3.7	24.20/4.1	23.78/4.1	24.13/4.1	25.03/4.1
15.42/4.0	27.46/4.2	27.30/4.1	27.41/4.2	27.93/4.1



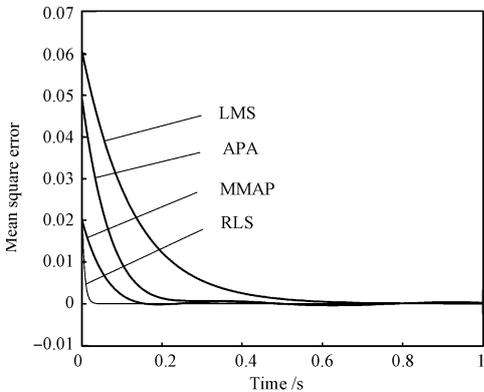
(a) 投影维数对收敛性能的影响 ($\mu = 0.2$)

(a) Influence of projection dimension on convergence ($\mu = 0.2$)



(b) 搜索步长对收敛性能的影响 ($L = [50 \ 20 \ 2]^T$)

(b) Influence of search step on convergence ($L = [50 \ 20 \ 2]^T$)



(c) 各种自适应算法收敛性能对比 ($\mu = 0.2$)

(c) Convergence comparison of different methods ($\mu = 0.2$)

图 6 MMAP 算法收敛性能分析

Fig. 6 Analysis of convergence performance of MMAP

图 6 列出了各种自适应算法的收敛性能对比, 以均方误差来衡量. 可以看出, APA 针对 LMS 收敛速度慢的问题有了明显改善; 对单纯的 APA 算法, 投影维数越高, 收敛速度越快, 但同时计算复杂度也越高; 对比各种自适应算法, MMAP 算法收敛性能最接近理论极限 (RLS 算法), 而算法复杂度较 RLS 算法要低得多. APA 算法要想达到与 MMAP 相同的收敛速率, 投影维数需要很高, 计算复杂度加大, 不易实时处理.

此外, 由于自适应过程涉及到矩阵的数值运算, 有必要衡量 MMAP 算法的数值稳定性, 保证其抗干扰能力. 考虑式 (6), 算法的主体在于矩阵求逆运算. 图 7 所示为 MMAP 和 APA 算法中矩阵 $X_n^T X_n$ 的对数条件数分布. 可以看出, MMAP 较之 APA 条件数明显下降, 证明了在算法稳定性方面, MMAP 要优于 APA. 条件数表示误差传播的程度, 条件数越小, 算法越有效, 数值稳定性越强.

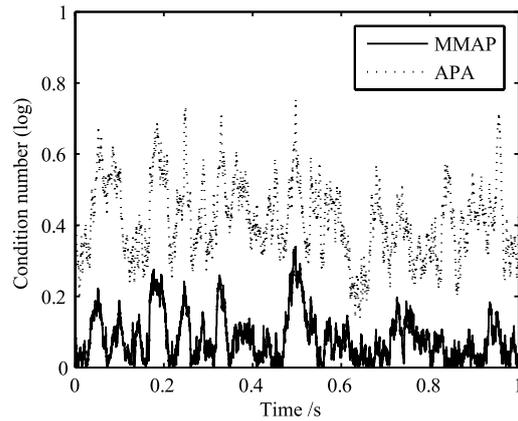


图 7 $X_n^T X_n$ 对数条件数比较

Fig. 7 Comparison of log condition numbers of $X_n^T X_n$

下面对各种自适应滤波算法加以比较. 定义算法复杂度为处理一帧数据所需要的乘法次数. 对于传统的 APA 算法, 误差计算需要 ML 次乘法运算, 矩阵求逆运算近似需要 $7L^2$ 次乘法运算^[7], 求 $\Delta\omega_n$ 需要 ML 次乘法运算, 因此 APA 算法总体复杂度为 $2ML + 7L^2$; 对于 MMAP 算法, 计算复杂度可以达到 $(M + 1)\bar{L} + 7\bar{L}^2$, 其中 \bar{L} 为平均投影维数; RLS 算法复杂度则为 $4M^2 + 4M + 2$. 表 2 列出了三种自适应滤波算法的平均复杂度, 以及在运算速度为 288 MIPS, 工作时钟为 144 MHz 条件下的 DSP 硬件仿真结果对比. LMS 算法复杂度虽然也比较低, 但由于其跟踪性能较差, 故此处没有列出. MMAP 算法复杂度取决于投影维数向量 L 和门限向量 t_i 的选取. 在收敛速度相同的前提下, MMAP 算法复杂度比 APA 算法低一个数量级, 比 RLS 算法低两个数量级. 在内存占用方面, MMAP 算法占用的程序空间也是最少的.

表 2 三种自适应算法复杂度与硬件仿真结果比较

Table 2 Comparison of complexity and hardware simulation results of three different adaptive filtering algorithms

自适应算法	平均算法复杂度	硬件仿真实时率	内存占用率 (%)
MMAP	$< 7.1 \times 10^3$	0.18 倍实时	25.8
APA	1.3×10^4	0.32 倍实时	30.4
RLS	2.6×10^5	4.88 倍实时	49.0

5 结论

本文提出了一种基于一阶微分麦克风阵列的自适应语音增强算法。采用“零点输入-低频补偿”技术,利用麦克风阵列的空域滤波结构,获得环境噪声的估计值,作为自适应滤波的参考信号;在自适应滤波阶段,采用变步长,变投影维数的仿射投影(MMAP)算法,在算法复杂度和收敛速度方面找到合理的折中。在5509 DSP芯片上实现了算法的硬件仿真。实验表明,MMAP算法实时性能明显提高,硬件资源占用率低,算法稳定性好。计算复杂度比其他自适应算法低1~2个数量级,与LMS算法最为接近,而收敛速度最接近于RLS算法。

References

- 1 Wee S, Chen H W, Yu Z L. Self-calibration-based robust near-field adaptive beamforming for microphone arrays. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2007, **54**(3): 267–271
- 2 Warsitz E, Krueger A, Haeb-Umbach R. Speech enhancement with a new generalized eigenvector blocking matrix for application in a generalized sidelobe canceller. In: *Proceedings of International Conference on Acoustic, Speech, and Signal Processing*. Las Vegas, USA: IEEE, 2008. 73–76
- 3 Ju G H, Lee L S. A perceptually constrained GSVD-based approach for enhancing speech corrupted by color noise. *IEEE Transactions on Audio, Speech, and Language Processing*, 2007, **15**(1): 119–134
- 4 Zhang Xian-Da. *Matrix Analysis and Applications*. Beijing: Tsinghua University Press, 2004. 310–332
(张贤达. 矩阵分析与应用. 北京: 清华大学出版社, 2004. 310–332)
- 5 Bouchard M, Ding H P. An exact relaxed fast affine projection algorithm for multichannel active noise control. In: *Proceedings of International Conference on Digital Signal Processing*. Cardiff, UK: IEEE, 2007. 47–50
- 6 Heinz T, Gray W E. First- and second-order adaptive differential microphone arrays. In: *Proceedings of International Workshop on Acoustic Echo and Noise Control*. Darmstadt, Germany: IEEE, 2001. 35–38
- 7 Gay S L, Tavathia S. The fast affine projection algorithm. In: *Proceedings of International Conference on Acoustic, Speech, and Signal Processing*. Detroit, USA: IEEE, 1995. 3023–3026
- 8 Choi M S, Baik C H, Park Y C, Kang H G. A soft-decision adaptation mode controller for an efficient frequency-domain generalized sidelobe canceller. In: *Proceedings of International Conference on Acoustic, Speech, and Signal Processing*. Hawaii, USA: IEEE, 2007. 893–896

宋 辉 清华大学电子工程系博士研究生。主要研究方向为语音增强、回声消除、麦克风阵列信号处理技术。本文通信作者。

E-mail: song.carter@gmail.com

(**SONG Hui** Ph.D. candidate in the Department of Electronic Engineering, Tsinghua University. His research interest covers speech enhancement, echo cancelation, and microphone array signal processing technique. Corresponding author of this paper.)

刘 加 清华大学电子工程系教授。主要研究方向为语音识别、语音合成、语音编码、语音识别专用芯片设计、多传感器融合技术、多媒体数字通信系统。E-mail: liuj@tsinghua.edu.cn

(**LIU Jia** Professor in the Department of Electronic Engineering, Tsinghua University. His research interest covers speech recognition, speech synthesis, speech coding, speech recognition ASIC design, multi-sensor fusion technique, and multi-media digital communication system.)