

doi: 10.11830/ISSN.1000-5013.201702013



采用预测残差的回声消除 单双端检测算法

谢维波, 王正腾

(华侨大学 计算机科学与技术学院, 福建 厦门 361021)

摘要: 在单端与双端状态下,自适应滤波器的预测残差数值迥异,对预测残差数值进行归一化处理,拉开单端、双端状态的数值差距,扩大阈值范围,提出基于预测残差的回声消除单双端判断方法,并提出在单双端交界处使用自适应滤波器权值复用与更新机制,以及算法的实现步骤.实验结果表明:该算法具有宽广的阈值选择范围,可以有效防止因单双端误判引起的滤波器发散,减小总的回声消除预测误差.

关键词: 回声消除;自适应滤波预测残差;阈值选择;权值复用与更新机制

中图分类号: TP 311 **文献标志码:** A **文章编号:** 1000-5013(2017)02-0201-06

Echo Cancellation Single and Double-Talk Detection Algorithm Using Predictive Residuals

XIE Weibo, WANG Zhengteng

(College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China)

Abstract: In the condition of single-talk and double-talk, the predictive residual data of self-adapting filter is different. this article set normalization processing in predictive residual data, opened numerical gap in single-talk and double-talk, and expanded the scope of the threshold value, then this article put forward the echo cancellation single and double-talk judgment method based on predictive residuals. And put forward that it can use the adaptive filter working mechanism of weight reuse and update in single-talk and double-talk junction. At last, the implementation steps of the algorithm are also given. The experimental results showed that this algorithm has a wide range of threshold selection, it can effectively prevent the filter divergence caused by inaccurate double-talk detection and reduce the total echo cancellation prediction error.

Keywords: echo canceller; prediction error of adaptive filter; threshold selection; working mechanism of weight reuse and update

随着通信技术的迅速发展,涌现出各种不同的通信方式,如国际互连网协议(IP)电话、视频会议等.由于 IP 网络数据的延时和语音编解码的耗时,导致 IP 语音通信系统的通信质量受到极大的影响和干扰^[1-2].因此,学者相继提出了最小均方(LMS)、归一化最小均方(NLMS)及部分归一化最小均方(PN-LMS)等基于 LMS 改进的自适应滤波算法来模拟回声,并在麦克风采集信号中减去模拟回声信号^[3-5].在自适应滤波算法的训练滤波器权值过程中,若近端环境下的麦克风采集信号包含有近端用户语音,则破坏远端语音信号与近端语音信号的相关性,导致滤波器难以收敛,甚至发散,无法模拟出较准确的回

收稿日期: 2015-11-24

通信作者: 谢维波(1964-),男,教授,博士,主要从事现代信号处理理论研究与实践、嵌入式应用的研究. E-mail: xw-blxf@hqu.edu.cn.

基金项目: 国家自然科学基金资助项目(61271383);福建省战略性新兴产业发展专项([2013]866)

声路径. 因此,单双端判断技术和自适应回声消除技术必须相辅相成、协同合作,共同完成回声消除,是声学回声消除系统采用的主要技术. 本文提出一种基于预测残差的回声消除单双端判断算法.

1 回声生成原理及回声消除流程

1.1 回声生成原理分析

远端语音信号 $x(k)$ 在多种回声信道 $\hat{\mathbf{w}}=[\hat{w}_0,\hat{w}_1,\cdots,\hat{w}_{L-1}]^T$ 的作用下,产生回声语音信号 $v(k)$,近端用户语音信号 $r(k)$ 与回声语音信号 $v(k)$ 被麦克风收集采样后,形成近端语音信号 $d(k)$,被网络发送给另一用户^[6-7]. 因此,在网络的另一端,接收者会播放掺杂有回声的语音信号 $d(k)$,降低用户体验度.

在回声消除系统中,需判断近端语音信号、远端语音信号的存在情况. 为叙述简便,定义以下名词. 1) 单端情况. 远端语音信号 $x(k)\neq 0$,近端用户语音信号 $r(k)=0$. 2) 双端情况. 远端语音信号 $x(k)\neq 0$,近端用户语音信号 $r(k)\neq 0$. 因此,准确地判断出单双端状态是回声消除的一个要点^[8-12].

1.2 回声消除系统流程分析

一般回声消除算法通过自适应滤波器完成^[6-7],通过调整自适应滤波器权值模拟回声信道,最终使所得 $\hat{d}(k)$ 逼近回声信号 $v(k)$,使用近端语音信号 $d(k)$ 减去模拟的回声语音信号 $\hat{d}(k)$,进而得到消除回声的语音信号 $e(k)$. 采用使用滤波器权值复用与更新机制的 NLMS 自适应滤波算法,计算得到消除回声的语音信号 $e(k)$,依据 $e(k)$ 的大小状态判定单双端,并将判断结果反馈于自适应滤波算法.

2 单双端判断算法的研究现状

2.1 Geigel 算法原理分析

Geigel 算法^[2,6] 是基于能量的单双端检测算法^[13],通过远端语音信号 $x(k)$ 和麦克风信号 $d(k)$ 的幅度构造检测统计量,即

$$\xi_{\text{Geigel}} = \frac{d(k)}{\max\{|x(k)|, |x(k-1)|, \cdots, |x(k-L+1)|\}} > T. \tag{1}$$

式(1)中: L 为窗口长度; T 为阈值门限. 当近端语音信号 $d(k)$ 与远端语音信号 $x(k)$ 的 L 个采样点内最大值的比值大于阈值时,就判断为双端对讲情况;否则,判断为单端状态. 此方法简单、易于实现,但是难以设定合适的阈值 T ,且存在较高的误判率.

2.2 基于互相关单双端算法原理分析

Ye 等^[8] 提出使用回声消除系统中 2 种语音信号的相关性构造统计量. 以远端语音信号与麦克风采集信号为例,其公式为

$$c_{\text{xd}} = \frac{E(x(k)d(k))}{\sqrt{E[x^2(k)]E[d^2(k)]}}, \tag{2}$$

且

$$\zeta_{\text{xdcc}} = \max_i |c_{\text{xd},i}|, \quad i = 0, 1, \cdots, L-1. \tag{3}$$

该算法利用统计量 ζ_{xdcc} 与 T 进行比较,判断双端对讲情况,增加了计算复杂度,难以设置合适的阈值. 针对互相关算法难以设置合适阈值的问题,Benesty 等^[10] 提出归一化互相关双端检测算法,即

$$\zeta_{\text{xdcc}} = \sqrt{r_{\text{xd}}^T (\sigma_d^2 R_{\text{xx}})^{-1} r_{\text{xd}}}. \tag{4}$$

式(4)中: σ_d^2 为麦克风采集信号的方差, $r_{\text{xd}}=R_{\text{xx}}\mathbf{w}$, $R_{\text{xx}}=E[\mathbf{x}(k)\mathbf{x}^T(k)]$. 考虑到 $\sigma_d^2, R_{\text{xx}}$ 等的计算量,势必引起检测延时,不利于实时检测.

通过以上分析可知:Geigel 方法存在难以设置合适阈值和误判率高的问题;基于互相关的单双端算法仍然存在难以设置合适阈值的问题,同时,增加了计算复杂度;归一化的互相关单双端判断算法在增加计算量的基础上解决了阈值设置问题,但是存在检测延时,不利于实时检测. 文中的研究服务基于 IP 网络的楼宇对讲系统,对实时检测有一定的要求. 基于预测残差的回声消除单双端检测算法,与 Geigel 算法都是基于能量的判断算法,具有结构简单、易于实现的优点,且阈值可选择范围广,同时,在单双端交界处提出滤波器权值复用更新机制,有效地抑制了滤波器发散.

3 回声消除应用中的 NLMS 算法

通常回声消除算法通过自适应滤波器完成，回声消除的关键是使自适应滤波器的权值 $\mathbf{w} = [\omega_0, \omega_1, \dots, \omega_{L-1}]^T$ 逼近回声路径信道 $\hat{\mathbf{w}} = [\hat{\omega}_0, \hat{\omega}_1, \dots, \hat{\omega}_{L-1}]^T$ ，自适应滤波器权值 \mathbf{w} 与回声路径信道 $\hat{\mathbf{w}}$ 的误差越小，回声消除效果越好。采用 NLMS 算法^[3,14-16]。回声消除原理，如图 1 所示。

回声信号模型表达式为

$$\hat{d}(k) = \mathbf{w}^T(k) \times \mathbf{x}(k). \tag{5}$$

式(5)中： $\mathbf{x}(k) = [x(k), x(k-1), \dots, x(k-L+1)]^T$ 为远端语音信号，是自适应滤波器的输入信号； L 为自适应滤波器的长度。

滤波器输出的先验误差信号为

$$e(k) = d(k) - \hat{d}(k) = d(k) - \mathbf{w}^T(k) \mathbf{x}(k). \tag{6}$$

结合图 1 与式(5)，(6)可知：通过初始误差 $e(k)$ 以某种机制更新滤波器的权值，从而使回声估计值 $\hat{d}(k)$ 不断逼近实际的回声 $d(k)$ 。

NLMS 算法采用滤波器权向量最小更新方式逼近回声信道，使稳态误差最小， $n+1$ 次迭代中滤波器抽头权向量的修正项被 n 次迭代时，输入向量 $\mathbf{x}(k)$ 的平方欧式范数归一化。因此，使式(7)的欧式范数最小化，即

$$\delta \mathbf{w}(k+1) = \mathbf{w}(k+1) - \mathbf{w}(k). \tag{7}$$

同时，受式(8)约束，有

$$\mathbf{w}^T(k+1) \mathbf{x}(k) = d(k). \tag{8}$$

NLMS 算法的代价函数表示为

$$\mathbf{J}(k) = \|\delta \mathbf{w}(k+1)\|^2 + \lambda [d(k) - \mathbf{w}^T(k+1) \mathbf{x}(k)]. \tag{9}$$

式(9)中： $d(k) - \mathbf{w}^T(k+1) \mathbf{x}(k)$ 为权向量更新后滤波器输出与期望响应的瞬时误差。将代价函数 $\mathbf{J}(k)$ 对 $\mathbf{w}(k+1)$ 求导置零，有

$$\frac{\partial \mathbf{J}(k)}{\partial \mathbf{w}(k+1)} = 2[\mathbf{w}(k+1) - \mathbf{w}(k)] - \lambda \mathbf{x}(k), \tag{10}$$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + 1/2\lambda \mathbf{x}(k). \tag{11}$$

将式(10)代入权值更新函数，可得

$$d(k) = [\mathbf{w}(k) + 1/2\lambda \mathbf{x}(k)]^T \mathbf{x}(k) = \mathbf{w}(k)^T \mathbf{x}(k) + 1/2\lambda \|\mathbf{x}(k)\|^2. \tag{12}$$

由式(12)可得

$$\lambda = \frac{2[d(k) - \mathbf{w}^T(k) \mathbf{x}(k)]}{\|\mathbf{x}(k)\|^2} = \frac{2e(k)}{\|\mathbf{x}(k)\|^2}. \tag{13}$$

将式(13)代入式(11)，可得

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\mu}{\gamma + \|\mathbf{x}(k)\|^2} \mathbf{x}(k) e(k). \tag{14}$$

式(14)中： μ 为收敛因子； γ 为修正值，可避免 $\|\mathbf{x}(k)\|^2$ 过小时步长太大。

以上公式分析了 NLMS 的原理，结合语音数据回声消除的应用环境。NLMS 算法在回声处理中的细节流程，如图 2 所示。当判断为单端时，NLMS 算法不仅滤波，同时更新权值；当判断为双端时，NLMS 算法只采取滤波的处理方式，同时使用单端时的最新权值。

4 单双端状态判断算法

NLMS 算法通过模拟回声信道预测回声语音信号，且输出误差是消除了模拟回声信号的理想语音，分别分析单端和双端状态下的 NLMS 自适应滤波算法的

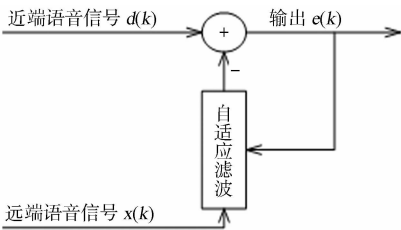


图 1 回声消除滤波算法
Fig. 1 Filtering algorithm of echo cancellation

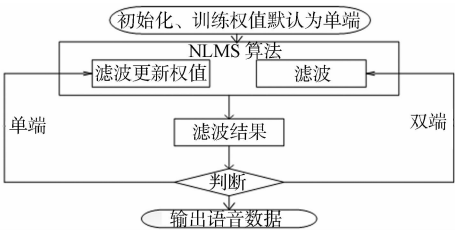


图 2 NLMS 滤波器的工作机制

Fig. 2 Working mechanism of NLMS filter

输出结果.

1) 在单端情况下,近端语音数据中混合的大部分回声数据已被去除.此时,输出的语音信号趋近于 0,NLMS 滤波器稳态误差小,且 $e(k)$ 趋近于 0,即

$$\lim_{k \rightarrow +\infty} \hat{d}(k) = d(k), \quad e(k) = d(k) - \hat{d}(k) \approx 0. \tag{15}$$

2) 在双端情况下,近端语音数据中混合的大部分回声数据虽然已去除,但近端用户语音信号 $r(k)$ 仍保留.此时,输出的语音数据信号 $e(k)$ 抖动大,NLMS 滤波器稳态误差大,即

$$e(k) = r(k) + \Delta e(k), \quad \lim_{k \rightarrow +\infty} \Delta e(k) = 0. \tag{16}$$

不同状态下,NLMS 的输出误差(预测残差) $e(k)$ 不同,当输出误差 $e(k)$ 大于阈值 T 时,为双端状态;输出误差 $e(k)$ 小于阈值 T 时,为单端状态.为了统一衡量预测残差比率,现取预测残差 $e(k)$ 与原始值 $d(k)$ 的比值.同时,为了提高判断结果的抗干扰能力,采用加窗平均估计, L 为窗口长度,即

$$\xi = \left[\frac{1}{L} \sum_{l=0}^{L-1} \text{abs}(e(k-l)) \right] / \left[\frac{1}{L} \sum_{l=0}^{L-1} \text{abs}(d(k-l)) \right] \geq T, \quad k = 1, 2, \dots, n. \tag{17}$$

因为单双端状态下滤波器输出误差 $e(k)$ 差异大,因此,阈值 T 的可选择范围宽广,可有效地区分单端状态与双端状态,使算法的表现更佳.文中阈值 T 可为系统训练值.

针对单双端检测误判导致滤波器发散的问题,算法在单双端交界处引入滤波器权值复用与更新机制,降低因误判而引起滤波器发散的机率,维持系统的鲁棒性,减少回声消除的总误差,在固定环境中,回声消除效果具有良好的表现.

在系统初始运行时,默认为单端状态训练滤波器,得到较佳的滤波器权值.然后,将此滤波器权值 w 保存到权值备份 w_backup 中.在单双端状态检测过程中,若检测到从双端状态进入单端状态时,NLMS 自适应滤波器复用已保存的滤波器权值;若检测到从单端状态转变到双端状态时,首先,衡量此时的滤波器是否收敛,若自适应滤波器收敛,则将当前的滤波器权值更新到权值备份 w_backup 中,否则,不做任何处理.其中,滤波器权值更新条件为

$$e_{win}(k) = [e(k-l), e(k-l+1), \dots, e(k)], \tag{18}$$

$$D(e_{win}(k)) \leq d_{last} \& \& \overline{e_{win}(k)} \leq \overline{e_{last}}. \tag{19}$$

式(19)中: $e_{win}(k)$ 为交界处前的 l 个已参加回声消除运算的语音数据样本; $d_{last}, \overline{e_{last}(k)}$ 分别为上一单双端交界处状态的滤波器输出误差 $e(k)$ 的方差与均值.当符合式(19)时,满足滤波器权值更新条件.

系统初始时间段内默认为单端状态(根据楼宇对讲应用情景)进行训练跟踪达到初步的稳态,同时,保存初始滤波器权值;然后,依据滤波器预测残差判断系统的单双端状态并反馈给系统,并在单双端状态转变时使用权值复用与更新机制.算法的整体流程,如图 3 所示.

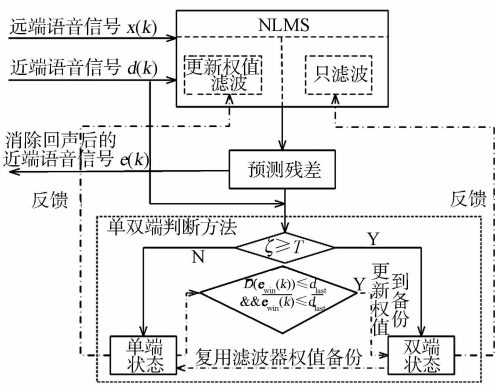


图 3 权值复用与更新机制
Fig. 3 Working mechanism of weight reuse and update

5 仿真实验及结果分析

在网络中随机下载语音信号 a、语音信号 b,如图 4 所示.语音信号采样率均为 8 K,远端语音信号 $x(k)$ 为语音信号 a 加信噪比为 30 的随机高斯白噪声,近端用户语音信号 $r(k)$ 为语音信号 b 加信噪比为 30 的随机高斯白噪声.回声语音信号 $v(k)$ 为采用 ITU-T G. 168^[17] 标准的回声模型 2 与远端语音信号 $x(k)$ 卷积合成,如图 5 所示.近端语音信号 $d(k)$ 为回声语音信号 $v(k)$ 与近端用户语音信号 $r(k)$ 的合成,且外加信噪比为 40 的随机高斯白噪声.合成后的语音信号,如图 6 所示.

当阈值 T 设置过大时,双端语音部分误判为单端语音部分,导致滤波器更新权值致使滤波器发散,引起误差.而当阈值 T 设置过小时,单端语音部分被误判为双端,存在大量的回声未被抑制,引起误差.因此,单双端判断算法直接影响到回声消除效果,通过回声消除的误差,可以反映出单双端算法的性能,

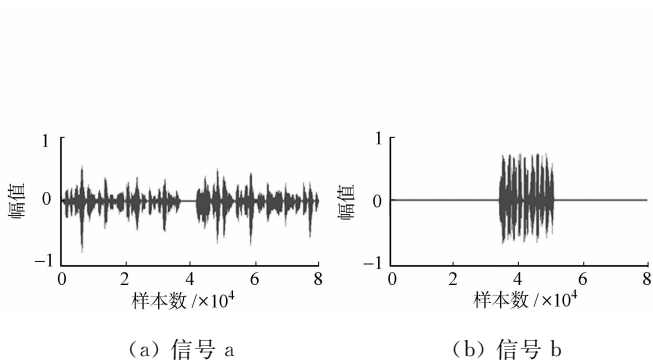


图 4 原始语音信号

Fig. 4 Original speech signal

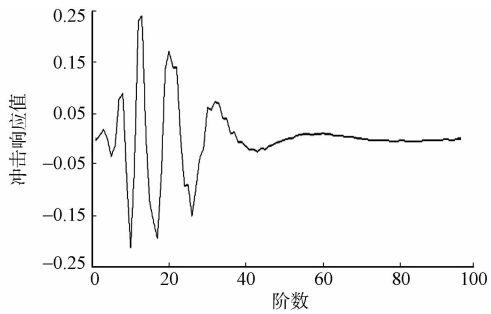


图 5 G.168 标准的回声模型 2

Fig. 5 Echo model 2 of ITU-T G.168

定义回声消除应用中双端部分的误差为

$$\left. \begin{aligned} \text{Error}'_{\text{double}}(k) &= [e(k) - r(k)]^2, \\ \text{Error}_{\text{double}} &= \sum_{k=1} \text{Error}'_{\text{double}}(k). \end{aligned} \right\} \quad (20)$$

回声消除应用中的总误差为

$$\left. \begin{aligned} \text{Error}_{\text{single}}(k) &= e(k)^2, \\ \text{Error}_{\text{total}} &= \text{Error}_{\text{double}} + \sum_{k=n_{\text{train}}} \text{Error}_{\text{single}}(k). \end{aligned} \right\} \quad (21)$$

在实验中,自适应滤波器的阶数设置均为 96,步长因子设置均为 0.075. 首先,默认前 6 000 个数据(0.75 s)为单端情况,且训练自适应滤波算法的权值. 然后,对语音数据采用文中的回声消除算法及单双端判断算法. 随着阈值变化,文中算法回声消除的双端部分误差变化(η),如图 7 所示. 随着阈值变化的回声消除总误差变化(η_{tot}),如图 8 所示. 由图 7,8 可知:当阈值在 0.25~0.70 时,文中算法回声消除的总误差与双端部分的回声消除误差平稳并且在 200 内.

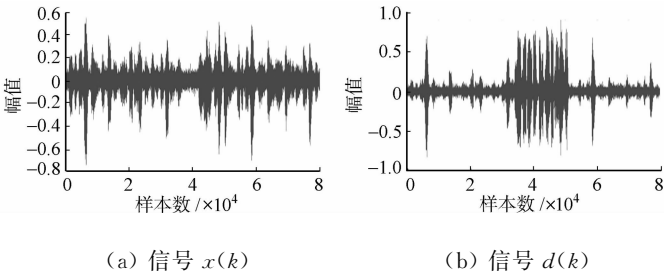


图 6 合成的语音数据

Fig. 6 Synthesized voice data

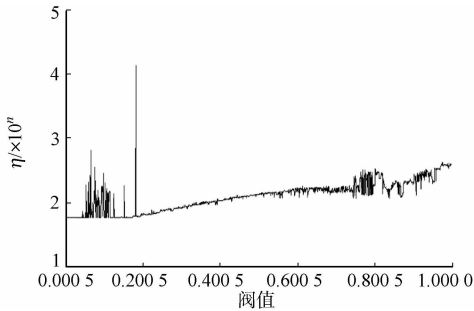


图 7 回声消除应用中双端部分误差

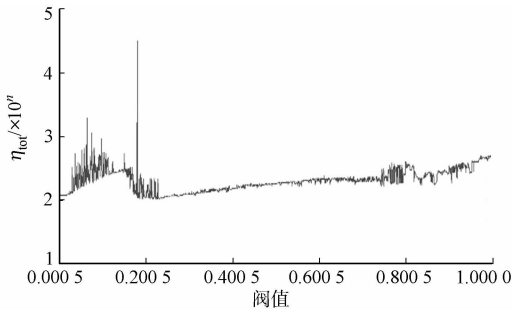


图 8 回声消除应用中总误差

Fig. 7 Error of double-talk part in echo cancellation system Fig. 8 Overall error of echo cancellation system

Geigel 算法仍使用同样配置的 NLMS 算法后的回声消除系统随阈值变化的双端部分误差变化,如图 9 所示. 随着阈值变化的回声消除总误差变化,如图 10 所示.

由图 9,10 可知:当 Geigel 算法的阈值范围在 0.34~0.38 时,回声消除的总误差与双端部分的回声消除误差平稳并且在 1 000 内,其他范围的回声消除误差极大. 通过以上对比可知,文中算法的可选阈值范围比 Geigel 算法更宽广. 文中算法在合适阈值范围内,回声消除的总误差在 200 内;而 Geigel 算法在合适的阈值范围内,回声消除的总误差在 1 000 内. 由此可知,文中算法提出的滤波器权值复用与更新机制,有效抑制因单双端误判引起的滤波器发散,减小了回声消除中的误差.

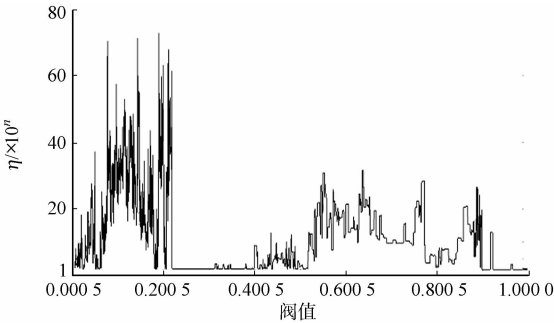


图 9 回声消除应用中双端部分误差

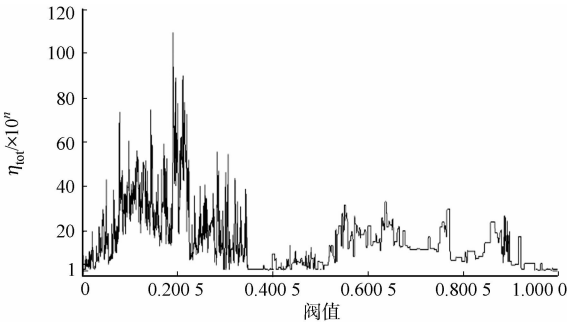


图 10 回声消除应用中总误差

Fig. 9 Error of double-talk part in echo cancellation system Fig. 10 Overall error of echo cancellation system

6 结束语

分析了回声产生的原理与回声消除系统的整体流程,提出基于预测残差的回声消除单双端判断方法,同时,提出在单双端交界处使用滤波器权值复用与更新机制,有效抑制因单双端误判引起的滤波器发散.文中算法与 Geigel 算法都是基于能量的判断算法,同样结构简单、易于实现,但阈值可选择范围更广,适应性更强.从仿真实验结果看,文中算法的可选择阈值范围远大于 Geigel 算法的合适阈值范围.同时,文中算法的回声消除总误差明显小于 Geigel 算法,具有更强的实用性.

参考文献:

[1] 西蒙.赫金.自适应滤波器原理[M].北京:电子工业出版社,2003:251-270.

[2] 宋知用. MATLAB 在语音信号分析与合成中的应用[M].北京:北京航空航天大学出版社,2013:284-320.

[3] NEKUII M, ATARODI M. Composite PNLMS & NLMS adaptation: A new method for network echo cancellation [C]//14th International Conference on Digital Signal Processing. Santorini:IEEE Press,2002:757-760.

[4] DENG H, DOROSLOVACKI M. Improving convergence of the PNLMS algorithm for sparse impulse response identification[J]. Signal Processing Letters,2005,12(3):181-184.

[5] 刘立刚, FUKUMOTO M, 张世永. 一种变步长 Proportionate NLMS 自适应滤波算法及其在网络回声消除中的应用[J]. 电子学报,2010,38(4):973-978.

[6] 尹浩琼. 通信系统中回声消除的研究[D].北京:北京邮电大学,2006:20-35.

[7] 江冰, 叶玲, 朱成健, 等. 一种提高稳态误差及收敛速度的回声消除算法[J]. 声学技术,2009,28(3):295-299.

[8] YE Hua, WU Boxiu. A new double-talk detection algorithm based on the orthogonality theorem[J]. IEEE Transactions on Communications,1991,39(11):1542-1545.

[9] AHGREN P, JAKOBSSON A. A study of double-talk detection performance in the presence of acoustic echo path changes[J]. IEEE Transactions on Consumer Electronics,2006,52(2):515-522.

[10] BENESTY J, MORGAN D R, CHO J H. A new class of doubletalk detectors based on cross-correlation[J]. IEEE Transactions on Speech and Audio Processing,2000,8(2):168-172.

[11] 宋彦, 汪萌, 戴礼荣. 带双端发声检测的回声消除改进算法[J]. 清华大学学报(自然科学版),2008(增刊 1):760-764.

[12] 许礼武, 张伟彬, 陈波. 一种新的回音消除的双向通话检测算法[J]. 声学技术,2008,27(4):539-542.

[13] 吴超, 付强, 颜永红. 基于噪声估计和能量比的双讲检测方法[C]//第十二届全国人机语音通讯学术会议. 贵阳:[出版者不详],2013:5.

[14] 谢鹏, 刘加. 采用变步长算法的无双端检测回声消除系统[J]. 数据采集与处理,2010,25(6):736-740.

[15] 万政伟, 惠晓威, 朱婷. 基于 NLMS 算法的自适应滤波器的研究与应用[J]. 计算机系统应用,2010,19(4):46-48.

[16] 范团宝. 基于滑动窗增强 NLMS 的回声消除算法研究[C]//2011 年度全国无线及移动通信学术大会. 北京:[出版者不详],2011:7.

[17] International Telecommunication Union. ITU-T recommendation G. 168[S]. Geneva:ITU,2000:86-105.