

一种新的回音消除的双向通话检测算法

许礼武, 张伟彬, 陈波

(大唐移动通信设备有限公司, 上海 200233)

摘要: 在回音消除的算法中,近端语音信号的出现会引起预测回声路径的自适应滤波器发散,一个很成熟的回声消除算法一定包含双向通话检测算法。提出了一种基于滤波器抽头活跃位置的双向通话检测算法,该算法利用滤波器最活跃的抽头位置相邻的 n 个抽头 2-范数作为一个检测参数,并且区分滤波器间差异性的参数的阈值是动态的,该算法对双向通话的出现与消失能做出快速的响应,实现简单,检测准确。文章的双向通话检测新算法是基于双辅助滤波器的回音消除器的结构,通过录制了真实的环境语音仿真验证了该算法的鲁棒性。

关键词: 回音消除;双向通话检测;自适应滤波器

中图分类号: TB556

文献标识码: A

文章编号: 1000-3630(2008)-04-0539-04

A novel double talk detection of acoustic echo cancellation algorithm

XU Li-wu, ZHANG Wei-bin, CHEN Bo

(Datang Mobile Communication Equipment CO.,LTD, Shanghai 200233, China)

Abstract: In acoustic echo cancellation algorithm, the presence of near-end speech causes divergence of the adaptive filter which is used to model the echo path. A robust acoustic echo canceller must be equipped with a double talk algorithm. This paper proposes a double talk detection method based on the active filter's tap position, which utilizes continuous n tap's coefficient 2-norm as detection parameter after the most active filter's tap position and the decision threshold is dynamic. This method can quickly react to onset and cessation of double talk detection, and it is easy to implement and detect precisely. The canceller of the proposed DTD algorithm is based on dual auxiliary structure, which is verified and simulated by recorded speech under the real environment.

Key words: acoustic echo cancellation; double talk detection; adaptive filtering

1 引 言

在电话会议系统以及 3G 手机终端中,声音的播放与拾取是采用扬声器与麦克风及其相应的硬件和软件实现的。扬声器播放的声音将经过多种路径传播后被麦克风拾取到,多路径反射的结果在远端形成不同时延的回声,如图 1。远端信号 $x(n)$ 的回声 $d(n)$ 加上近端语音信号 $s(n)$ 一起被传输到远方,使远端用户感觉不舒服,声学回音消除就成为这些设备的基本组件。

回音消除的核心是使用一个自适应的滤波器来

预测回声路径的脉冲响应。当近端麦克风不仅采集了近端说话者语音也包含了远端的回音,即是所谓的双方通话状态时,由于远端信号与近端信号的不相关特性,所以自适应滤波器会大大受干扰,甚至发散。解决这个问题的共同的方法就是使用双向通话检测器(DTD)来决定是否更新滤波器的系数。目前比较流行的 DTD 方法有最简单易行的 GEIGEL 法、相关函数法和回声路径估计法,变化的脉冲响应法等^[1,2]。GEIGEL 方法是假定回声路径不变,所以这种方法很难适合声学回声消除,相关函数法假定的远端语音信号与近端语音信号不相关,实际中它们还是有点相关性,回声路径估计法和变化的脉冲响应法都有响应时间滞后的弊端,以上方法对阈值的设定很敏感,而本文的抽头活跃位置比较法对阈值要求具有不敏感的特性,实际中具有很好的应用。

收稿日期:2007-07-28;修回日期:2007-12-15

作者简介:许礼武(1975-),男,安徽巢湖人,硕士,高级音频工程师,研究方向为语音信号处理。

通讯作者:许礼武,E-mail:tulip0371@163.com

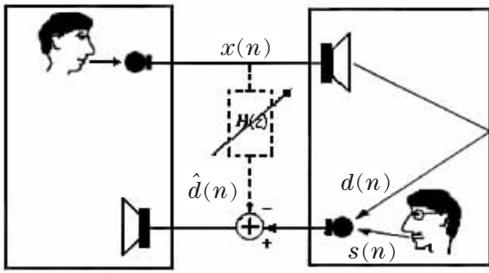


图1 回声产生的路径
Fig.1 Echo path

2 双辅助的自适应滤波器的结构

无论哪一种 DTD 算法都有个共同的特点: 他们都是将判决值和判决因子比较, 由判决结果的两种情况来划分不同的语音模式。由于回声路径的时变、延迟、非线性等特征, 所以很难有个确定的阈值来准确地检测通话的状态, 因此基于双滤波器结构回声消除器被提出^[3], 这种结构的主要的思想就是形成一个前景滤波器和一个背景滤波器, 背景滤波器被一直更新, 当背景滤波器的产生的剩余回声小于前景滤波器产生的剩余回声时, 背景滤波器系数拷贝到前景滤波器, 但是当双向通话时, 很容易造成误传而使前景滤波器不是工作在最佳状态。本文采用文献[4]提出的双辅助的滤波器结构(如图2)。 $x(k)$ 表示远端信号, $y(k)$ 表示麦克风采集到的信号, $s(k)$ 包含近端说话信号与周围噪声信号, $h(k)$ 表示回声路径的脉冲响应。图中除了一个主滤波器 h_1 外还有二个辅助滤波器, 分别是 h_0 与 h_2 , 根据它们不同的分工, h_0 被称作探测滤波器, h_2 是备份滤波器。

探测滤波器 (h_0) 的系数一直更新, 如 h_0 趋于发散则不更新主滤波器 (h_1) 系数, 否则拷贝 h_0 到 h_1 , 也就是完成了 h_1 的系数更新, 为了使回声消除器更鲁棒, 当由于 h_0 误传给 h_1 而使 h_1 也处于发散状态时, 备份滤波器 h_2 被使用, 这时 AEC 的输出来自备份滤波器的输出 $e_2(k)$ 。这样整个装置一直输出较小的回声 $\min(e_1(k), e_2(k))$ 。

3 双向通话检测算法

AEC 算法的双向通话检测 (DTD) 算法的结果将直接影响整个回声消除器的性能。许多 DTD 算法都是基于单参数的阈值比较, 这些 DTD 产生一个检测参数 $\xi(k) = f(x(k), y(k))$, 检测参数 $\xi(k)$ 同一个阈值相比较: $\xi(k) > \text{Threshold}$ 则双向通话存在, 通常来说, DTD 有个粘滞的特征, 也就是当某个

样本时刻出现了双向通话则下 N_{hold} 个样本也保持在双向通话状态而不管 $\xi(k)$ 的值。其中脉冲响应方差算法 (VIRE) 把抽头最大值位置的方差作为一个参数。回音消除通常以失配 MIS 作为评价标准:

$$MIS = 10 \lg \|h - w\|_2^2 / \|h\|_2^2 \quad (1)$$

h 是真实回声冲击响应, w 是预测的滤波器系数, 当 MIS 越小表示收敛越好, 根据这一度量标准, 本文提出的 DTD 算法主要利用滤波器抽头系数最活跃位置相邻的 n 个抽头的 2-范数作为一个检测参数, 以克服 VIRE 方法的对双向通话出现的反映滞后性。由于回声路径的脉冲响应处于活跃的部分一般较短, 所以为了减少运算复杂度, 本文只计算抽头系数最大值后相邻的 50 个系数的 2-范数。

图2所示的结构中, 远端信号 $x(k)$ 产生回声 $d(k)$, 回声加上可能近端信号一起形成麦克风信号 $y(k)$ 。主滤波器 $h_1(k) = [h_1(k), \dots, h_1(k-N+1)]^T$ 产生的预测回声 $d_1(k)$, 抑制回声后的信号 $e_1(k)$ 由麦克风信号 $y(k)$ 减去预测回声 $d_1(k)$ 。

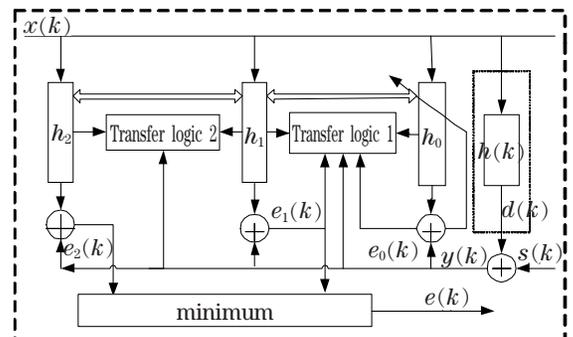


图2 双辅助滤波器结构的回声消除器
Fig.2 Acoustic echo canceller with dual auxiliary filter structure

$$e_1(k) = y(k) - d_1(k) = y(k) - h_1(k) \otimes x(k) \quad (2)$$

其中 $x(k) = [x(k), \dots, x(k-N+1)]$, “ \otimes ”表示卷积运算, 类似的方法得到。

$$e_2(k) = y(k) - d_2(k) = y(k) - h_2(k) \otimes x(k) \quad (3)$$

$$e_0(k) = y(k) - d(k) = y(k) - h_0(k) \otimes x(k) \quad (4)$$

试探滤波器系数一直更新, 更新算法采用 NLMS 算法^[1]:

$$h_0(k+1) = h_0(k) + \frac{\mu e_0(k) x(k)}{\|x(k)\|^2 + \varepsilon} \quad (5)$$

其中 $\|x(k)\|$ 是向量 $x(k)$ 的 2-范数, μ 是步长控制参数, 为了防止式(4)除数为 0, 一个远小于 1 的 ε 值被引入。当系数 $h_0(k)$ 逼近某一回声路径时, $h_0(k)$ 的幅度最大值的位置应该处在某值附近不变, 于是当试探性滤波器系数 $h_0(k)$ 与主滤波器 $h_1(k)$ 的最活跃的位置的后 50 个系数欧式距离发生大的变化

时会认为可能有近端语音出现,再结合剩余回声信号 e_0 和 e_1 的能量判断是回声路径改变还是近端信号的出现,也就是:

$$[\max \text{ value}, \text{Index0}] = \max(\text{abs}(\mathbf{h}_0(k))) \quad (6)$$

$$[\max \text{ value}, \text{Index1}] = \max(\text{abs}(\mathbf{h}_1(k))) \quad (7)$$

$$\xi_1 = \left\| \sum_{i=\text{Index0}}^{\text{Index0}+50} h_0(i) - \sum_{i=\text{Index0}}^{\text{Index0}+50} h_1(i) \right\|_{l_2} \quad (8)$$

$$\text{能量计算公式为: } p(\lambda, n) = \frac{1}{n} \sum_{k=1}^n \lambda(k) \times \lambda(k) \quad (9)$$

$$\text{erle}(\lambda) = \frac{\sum_{i=k}^{k-n+1} y(i) \times y(i)}{\sum_{i=k}^{k-n+1} \lambda(i) \times \lambda(i)} \quad (10)$$

其中 k 是当前样本, n 是窗口的长度, λ 作为一个信号变量(可以是 $y(k), x(k), e(k)$), 当试探滤波器的系数幅度的最大值连续保持在同一位置的次数大于某个阈值 cnt , 滤波器处于收敛过程中, 麦克风信号能量与剩余回声信号能量的比率($erle$)越大表明收敛得越好, 在整个收敛过程中拷贝 $erle$ 值最大时的主滤波器系数到备份滤波器, 这样能保证备份滤波器一定是收敛的, 出现双向通话时利用备份滤波器来预测回声, 为了让主滤波器从发散状态快速恢复, 当双向通话状态结束时, 拷贝主滤波器的系数到试探滤波器, 由于收敛过程具有跟随效应, 当某个样本被检测为需要更新主滤波器系数时, 下 N_{update} 也被认为需要更新。由于回声路径的改变也会造成抽头活跃位置的变化, 为了区分是双向通话还是回声路径的改变造成的这种变化, 本文还利用剩余回声能量的大小作为另一个参数来区分回声路径的改变^[3,5], 整个算法描述为:

(1) 根据公式(2)、(3)、(4)分别计算 $e_1(k), e_2(k), e_0(k)$, 保留 $\min(e_1(k), e_2(k))$ 的值。

(2) 根据公式(5)更新试探滤波器系数, 同时由公式(9)计算 $e_0(k), e_1(k)$ 能量值。

(3) 是否是双向通话结束, 是则拷贝主滤波器到试探滤波器。

(4) 公式(6)得到 Index0 与上个样本得到的值的差小于 3 吗? 不是则清 counter, 跳到 6。

(5) 变量 counter 加 1, counter 是否大于阈值 cnt , 是则根据公式(10)计算主滤波器产生的 $erle(e_1)$, 如该值比上次大则拷贝主滤波器到备份滤波器, 并记录此时的剩余回声能量 backupPower。

(6) 计算 Index1, 并与 Index0 比较, 如差值的范围小于 3 则认为还处于收敛过程中, 否则进入步

骤 9(可能发散的状态处理)。

(7) 公式(10)计算的试探滤波器的 $erle(e_0)$ 值大于 1.5, 且 counter 大于 cnt 则认为是收敛, 下 N_{update} 个样本需要更新主滤波器系数, 然后重复步骤(1)。

(8) 如连续二帧 $power(e_0, n) < power(e_1, n)$ 则认为是收敛, 可以更新主滤波器系数, 然后重复步骤 1, 否则, 需要比较试探滤波器与主滤波器的差异性, 根据公式(8)计算 ξ_1 , 若 $\xi_1 > \text{thresh1}$ 且 $power(e_1, n) > power(e_0, n)$, 则有双向通话发生, 拷贝备份滤波器到主滤波器, 重复步骤(1)。

(9) 由于主滤波器与试探滤波器的最活跃位置发生了大的变化(大于 3), 所以这个过程基本判断为双向通话发生或是回声路径发生改变。当连续两帧 $power(e_0, n) < power(e_1, n)$, 同时 $erle(e_0) > 1.8$, 则回声路径发生了改变, 主滤波器需要更新系数, 否则, 计算试探滤波器与备份滤波器的差异性, 根据公式(8)计算 ξ_1 , 若 $\xi_1 > \text{thresh2}$ 且 $erle(e_1) > erle(e_0) \times 0.91$, 则判断为双向通话发生, 拷贝备份滤波器到主滤波器, 重复步骤(1)。

4 算法的实现仿真

录制一段 20s 的语音信号, 如图 3 所示, 根据 G.168 标准^[6]的模型生成的回声信号叠加了一段近端语音信号合成信号 $y(k)$, 如图 4, 前半部是回声模型 3, 后半部采用回声模型 1 生成的回声, 如图 7, 算法实现时, 为了节省计算的开销, 每隔 M 个有效样本判断一下回声路径的状态, 远端信号的噪声电平能量利用动态检测方法, 即采用了每帧长度 160, 交叠 50% 的窗计算 $x(k)$ 的能量, 比较相邻二帧的能量差是否在 6.6×10^4 内, 如连续这样的条件成立达 8 次, 则判断为噪声而不需要更新任何滤波器的系数。剩余回声能量采用每帧长度 240, 交叠 37.5% 的矩形窗, $N_{\text{update}}=128, N_{\text{hold}}=256, cnt=\frac{N}{M}$, 滤波器长度 N 取 1024, M 取 8, thresh1 由于处在准收敛区, 所以初始值取为 0.03, 每次进入双向通话时按公式:

$$\text{thresh1} = \text{thresh1} \times 0.97 + (1 - 0.97) \times \xi_1 \quad (11)$$

进行更新。而 thresh2 处在趋发散区, 初始值为 0.112, 但是当进入该区时, 主滤波器还处在更新状态则该初值变为 0.088, 以后每次也按类似公式(11)的方法更新 thresh2。运行上述算法后得到的剩余回声信号如图 6, 检测到的双向通话位置如图 5 所示, 图 7 中既有近端语音同时在 8×10^4 的位置后回声路径发生了变化, 检测到的双向通话位置和回声取消后的结

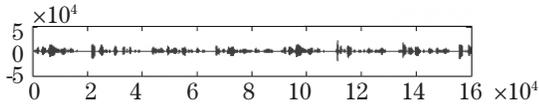


图3 远端信号
Fig.3 Farend signal

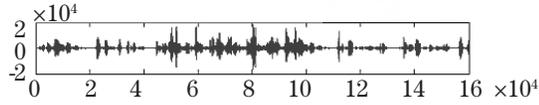


图4 Nearend signal containing one kind of echo path

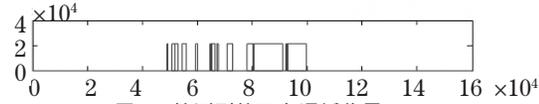


图5 检测到的双向通话位置
Fig.5 Double talk position

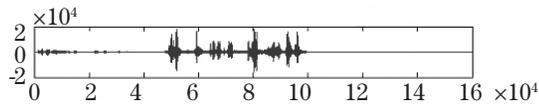


图6 剩余回声信号
Fig.6 Residual echo signal

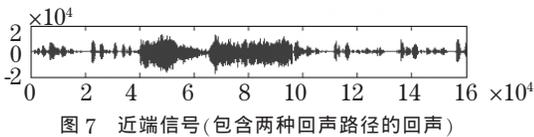


图7 Nearend signal containing two kinds of echo path

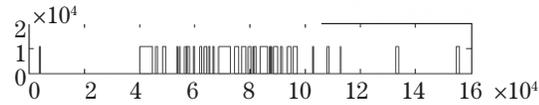


图8 检测到的双向通话位置
Fig.8 Double talk position

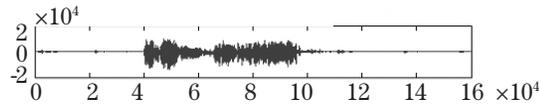


图9 剩余回声信号
Fig.9 Residual echo signal

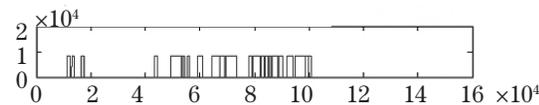


图10 GEIGLE 算法检测的 DT 位置
Fig.10 Double talk position by geigle

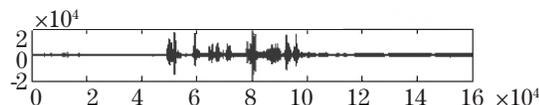


图11 GEIGLE 算法得到的剩余回声信号
Fig.11 Residual echo signal by geigle

果分别如图 8、9, 可见回声路径的改变本算法也能快速响应, 一些双向通话的误检也不影响回声取消的效果。使用 GEIGLE 算法得到的双向通话位置如图 10, 得到的剩余回声信号如图 11, 可见检测的误差较大, 回声抵消后的信号幅度大于图 6 的指示的信号^[7]。

5 小结

本文在双辅助滤波器回声抵消器结构的基础上, 提出基于滤波器活跃抽头位置判断滤波器的收敛发散状况, 比较试探滤波器与主滤波器最活跃抽头位置后连续的 n 个抽头的欧式距离的大小来判断是否有双向通话的存在, 同时这个区分滤波器间差异性的参数的阈值不断更新的, 由于双辅助滤波器这种特殊的结构, 克服了过去回声抵消器由于误判造成的滤波器发散的矛盾。试验证明了这种算法能快速响应双向通话的出现与消失。

参 考 文 献

- [1] Per Ahgren. On system identification and acoustic echo cancellation[J]. Uppsala University, 2004.
- [2] Ye H, Wu B X. A new double-talk detection algorithm based on the orthogonality theorem[J]. IEEE Trans. on Communications, 1991, 39(11): 1542-1545.
- [3] Akihiro Hirano, Kenji Nakayama. Double-talk resistant acoustic echo canceller with double filters[A]. International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2003)[C]. Awaji Island, Japan, December 7-10, 2003: 367-370.
- [4] OU Xiongbing, CHEN Zhe. An echo canceller based on the structure of dual-auxiliary filters[A]. Kyoto, Japan: International Workshop on Acoustic Echo and Noise Control[C]. 2003, 09: 35-38.
- [5] 蒙淑艳. 自适应回声抵消和噪声消除算法研究[D]. 长春: 吉林大学, 2004.
MENG Shuyan. Study on adaptive echo cancellation and noise reduction[D]. Changchun: Jilin University, 2004.
- [6] ITU-T Recommendation G.168(08/2004), Digital network echo cancellers[S].
- [7] 唐邦友, 夏秀渝, 何培宇, 等. 可区分回波通道变化和双端语音的检测算法研究[J]. 四川大学学报(自然科学版), 2007, 44(1): 73-76.
TANG Bangyou, XIA Xiuyu, HE Peiyu, et al. Study on detection algorithms able to distinguish echo path changed from double talk[J]. Journal of Sichuan University (Natural science edition), 2007, 44(1): 73-76.