

Speex 算法中回声抵消算法的分析和 dsp 实现

刘德青¹, 黄孝建¹, 李敬²

1. 北京邮电大学信息与通信工程学院, 北京 (100876)

2. 清华大学电子工程系, 北京 (100084)

E-mail: freeliuniu@yahoo.com.cn

摘要: 在 VoIP 电话中, 由于扬声器播放出来的声音被麦克风拾取后发回远端, 加上网络和数据处理等各种延迟的影响, 使得远端通话者能听到自己的回声, 严重影响了通话的质量。因此, 控制和消除回声是 VoIP 电话的技术关键之一。本文分析了 Speex 中的回声抵消算法, 深入分析了回声抵消算法中的 LMS 算法和自适应滤波器系数更新的原理, 介绍了算法的定点化的实现方式。最后通过仿真, 表明此回声抵消算法能够有效的改善通话的语音质量。

关键词: 回声抵消; 自适应滤波器; 定点化

中国分类号: TB535+.2

1 引言

在电话会议和网络语音通信中, 经常遇到的一个问题就是声学回声的干扰具体的表现形式是在本地, 远端呼叫者的语音被扬声器以声波的形式播放出来, 声波通过在房间内的传播和反射后被本地的麦克风接收, 通过这个回路, 声音被传送回呼叫端, 呼叫者就听到了自己的回声。回声抵消方法是利用自适应滤波器来估计声音的传输函数, 并使用这个估计值在反馈中抵消掉回声信号。

2 回声抵消的原理

回声抵消器在网络中的工作位置如图 1 所示, 回声抵消器是一个典型的横向自适应滤波器, 由远端语音信号 $x(n)$ 作为参考输入, 另一个输入为近端输入 $d(n)$ 。 $d(n)$ 由两个成分构成, 参考信号的回声 $w(n)$ 和由通话设备传来的近端语音信号 $v(n)$ 。滤波器的输出是估计的回声信号。由近端信号中减去估计回声信号得到误差信号 $e(n)$, 误差信号再返回到回声抵消器中用来调整滤波器的参数, 使误差的均方值达到最小。滤波器的参数 $W(n)$ 随着误差的变化而变化, 每隔一个码元间隔调整一次, 调整的幅度与选取的步长参数有关。在选择步长时, 须要在稳态误差和收敛速度之间折衷考虑。研究表明, 步长较小时, 失调就比较小, 但需要较长的收敛时间; 而步长较大时, 收敛时间变短, 但是稳态误差变大, 失调较大。

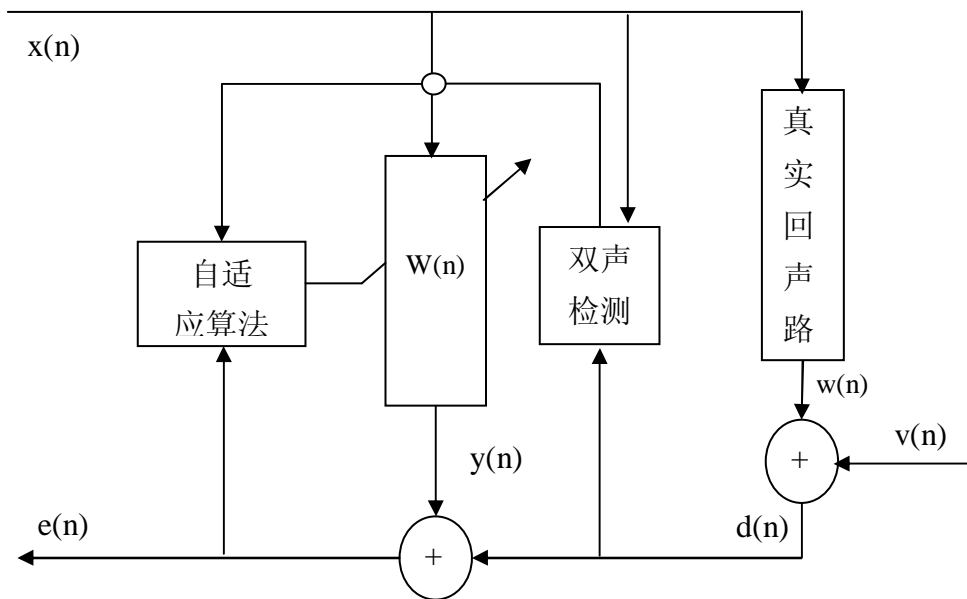


图 1 回声抵消结构图

2.1 最小均方 (LMS) 算法

在回波抵消中经常采用归一化的最小均方误差 LMS 算法。Speex 算法中的回声抵消算法叫做 MDF (MultiDelay block Frequency-domain) 算法, 这个算法就是采用了频域的 LMS 算法, 用频域的 LMS 是为了减少滤波器阶数很大带来的大的运算量, 而且如果直接使用 LMS 算法, 每进入一个输入数据就要更新一次权值, 那样就使算法的复杂度很大, 所以使用批处理的 LMS 算法, 这种算法是滤波器权值是在对收集到的一批数据后才进行一次更新的。

这时就要将输入语音序列分成批, 常用的分批方法有两种: 重叠保留法和重叠相加法。在文献[1]中, 得到对于频域 LMS 算法, 重叠保留法的运算量比较小, 而且 50%重叠时效果最好。但是在回声抵消程序里直接用这种算法, 有比较高的延时, 还需要大阶数的 FFT 等缺点, 因此在 MDF 算法中我们用小的 FFT, 但是频繁的更新系数。

在 MDF 自适应滤波器中, 在文献[2]中, 假设滤波器阶数 L 是处理帧长 N 的整数倍, 即 $L=KN$, K 为整数。FLMS 就是当 $K=1$ 时 MDF 的特殊情况。采用重叠保留法的输入信号的频域表达式:

$$X_F(K, m) = \text{diag}\{FFT[X^T(m-1), X^T(m)]\}, \text{ m表示第几帧。}$$

其他块的频域表达式是由 $X_F(k, m) = X_F(k+1, m-1)$ 得到。则滤波器输出和误差信号为:

$$Y(m) = IFFT[\sum_{k=0}^{K-1} X_F(k, m)W_F(k, m)] \text{ 的后半部分。}$$

$$E(m) = FFT\{0, \dots, 0, [D(m) - Y(m)]^T\}^T$$

其中 $W_k = [W_k^T, 0, \dots, 0]^T$, $D = [0, \dots, 0, D^T]^T$, $X(m) = [X^T(m-1), X^T(m)]^T$

权值更新公式 $W_k(m+1) = W_k(m) + \mu\phi(m)$, 其中 $\phi(m) = IFFT\{X(m)E(m)\}$ 的后半部分。

2.2 最佳步长的计算

在文献[3]中提到在 NLMS 算法中当假设近端不含回声的信号 $v(n)$ 和远端信号 $x(n)$ 为白噪声信号且他们不相关时，可近似得到使得滤波器失调最小的条件是：

$$\mu_{opt}(n) \approx \frac{\sigma_r^2(n)}{\sigma_e^2(n)} \quad (\text{当滤波器阶数趋于无穷大时这种近似为真})。$$

从而，可以近似得到：

$$\mu_{opt}(k,l) \approx \frac{\sigma_r^2(k,l)}{\sigma_e^2(k,l)}$$

其中 k 表示频率， l 表示第几帧。在这里 $\sigma_e^2(k,l)$ 是可以估计出来的，但是 $\sigma_r^2(k,l)$ 难以估计出来。

现在我们假设一个跟频率无关的量失调系数 $\eta(l)$ 表示这个自适应滤波器的失调。则：

$$\hat{\sigma}_r^2(k,l) = \hat{\eta}(l) \hat{\sigma}_y^2(k,l)$$

其中 $\hat{\eta}(l)$ 是估计的失调系数。这样对 $\sigma_r^2(k,l)$ 的估计就变成了对一个变化缓慢但是估计比较困难的 $\hat{\eta}(l)$ 的估计和一个变化快速但是比较容易估计的 $\hat{\sigma}_y^2(k,l)$ 的估计。

在 Speex 算法中，对于失调系数是这样估计的。首先设：

$$P_Y(k,l) = (1-\gamma)P_Y(k,l-1) + \gamma(|\hat{Y}(k,l)|^2 - |\hat{Y}(k,l-1)|^2)$$

$$P_E(k,l) = (1-\gamma)P_E(k,l-1) + \gamma(|E(k,l)|^2 - |E(k,l-1)|^2)$$

那么 $\hat{\eta}(l)$ 就等于估计出来的回声的能量 $P_Y(k,l)$ 和输出的信号能量 $P_E(k,l)$ 的线性衰减。

$$\text{即：} \hat{\eta}(l) = \frac{\sum_k R_{EY}(k,l)}{\sum_k R_{YY}(k,l)}$$

$$\text{其中：} R_{EY}(k,l) = (1-\beta(l))R_{EY}(k,l) + \beta(l)P_Y(k)P_E(k)$$

$$R_{YY}(k,l) = (1-\beta(l))R_{YY}(k,l) + \beta(l)P_Y(k)P_Y(k)$$

3 算法的定点 DSP 实现

定点 DSP 同浮点 DSP 相比,具有体积小、功耗低、运算速度快且价格低廉等优点，但是由于定点运算会带来溢出、截尾等降低精度的不良影响，而提高算法的精度往往是以降低处理速度为代价的。所以在浮点算法的定点 DSP 实现中，尤其要注意处理好速度和精度的关系。

在定点运算系统中，数据都是用二进制的补码数来存储的，数据按照二进制补码的运算法则进行运算。对于小数，我们是对这个二进制的补码数加入一个定标值来表示的，然而，这个定标值并不实际存储在存储单元，而是完全根据程序员的理解来确定定标的位置，这个定标的位置可以用 Q 值来表示，例如 Q3 表示该数的最后三位用来表示小数。而 Q 值表示

法也有别的表达形式，例如在 ADI 公司的文档中，用的是 M.N 的格式来表示定点数，其中 M 表示了整数的位数，N 表示了小数的位数，QN 与 M.N 表达的定点数的格式是等价的。

有了 Q 值的概念，就可以方便地对该实数变量进行定标，从而找到合适的表达法，具体方法是估计出该变量的最大值，然后找到大于或等于这个值的最小的 2 的幂，就可决定 Q 值表达法中整数的位数，从而也决定了该数的表达法。

但是，有些变量的动态范围通过经验或者理论分析是可以确定的，但大部分参数的动态范围需要利用大量样本通过统计分析得到。所谓统计分析，就是用足够多的样本值来确定程序中参数的动态范围，这里的参数样本值一方面要有一定的数量，另一方面必须尽可能地涉及各种情况。当然，统计分析毕竟不可能涉及所有可能发生的情况，因此对统计得出的结果在程序设计时可采取一些保护措施，如适当牺牲一些精度，Q 值可以取比统计值稍大些，使用 DSP 芯片提供的溢出保护功能等。

此外，有些参数我们可以根据动态范围来选取合适的 Q 值，但是在实际情况下，有些参数的动态范围并不大，但是却对精度要求非常高，使得我们对于这样的参数就需要进行精度的扩展。

参数定标完成后就可以进行程序内部函数的改写。这部分工作，需要用到定点数学函数集。绝大多数的数学运算（如加、减、乘、除……）都可以用定点数学函数集里的函数替代，这些函数的主要作用是将浮点运算变为定点运算，另一个作用是在定点运算出现溢出时采用溢出饱和的措施。

在函数体内部用到的浮点型变量也要转换为定点类型。一般来说，随着输入参数定点化和定点数学函数集的使用，内部变量的精度级别比较容易确定，从而转换为对应定点类型。其中，作临时变量用的中间变量一般转换为 Word32 类型，以保证运算的中间结果的准确性。

4 仿真结果

4.1 软件仿真分析

在 Palmmicro 公司开发的 PalmAdsp 仿真器中运行 Speex 中定点化的回声抵消程序，可以得出运行该程序共需要 5764 个 PM（程序储存区）和 5719 个 DM（数据储存区），运算量需 33.3MIPS。

为了检测 MDF 算法的收敛速度，可以计算一个平方误差的 $e^2(k)$ 的短时平均，从图 2 中可以看到这种算法的收敛速度大约为 300ms 左右，这是由于程序中滤波器系数并不是每次都更新全部，而是只更新其中一部分，这一段时间就需要 100ms 左右。

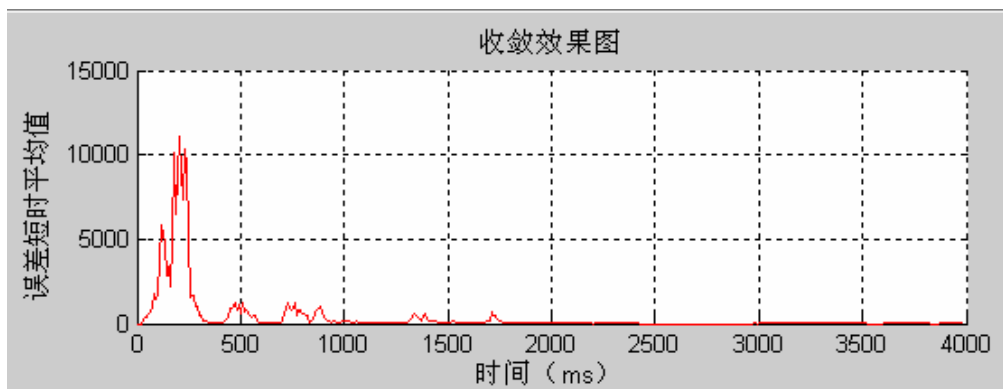


图 2 仿真结果

4.2 硬件仿真分析

在 Palmmicro 公司开发的 AR1688 芯片硬件平台上运行定点 DSP 的回声抵消算法代码。在网上抓包然后用 Palmmicro 公司开发的调试工具中解码,在 CoolEdit2000 语音工具软件中查看输出语音,同时可以直观的看到语音波形图,通过图形的对比,来比较说明回声抵消的作用。下面四幅图中,图 3 和图 4 两图是当电话不使用回声抵消模块时,在网上捕捉到的两边电话的声音,其中图 3 是原始语音,图 4 是回声;图 5 和图 6 两图是电话使用回声抵消模块时,在网上捕捉到的两边电话的声音,其中图 5 是原始语音,图 6 是经过回声抵消模块的输出。从波形开可以明显看出经过回声抵消模块的语音波形明显比未经回声抵消模块的语音波形小,从听觉上来听,未经回声抵消的电话有明显的回声,而经过回声抵消的回声基本上听不到回声。

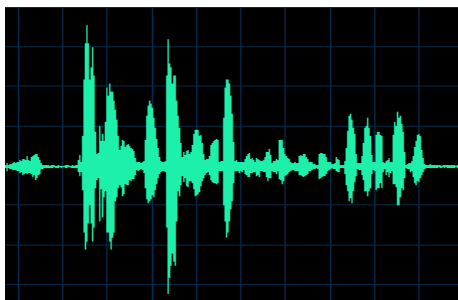


图 3 主叫话机的语音波形

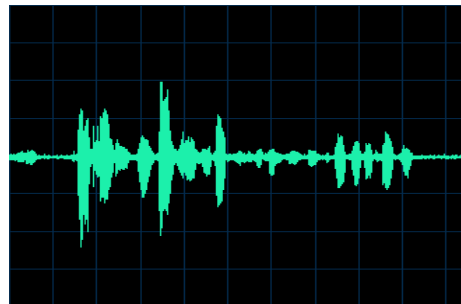


图 4 被叫话机未经回声抵消的语音波形

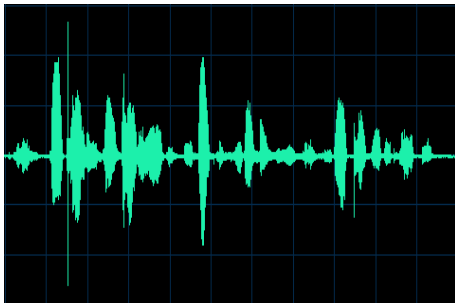


图 5 主叫话机的语音波形

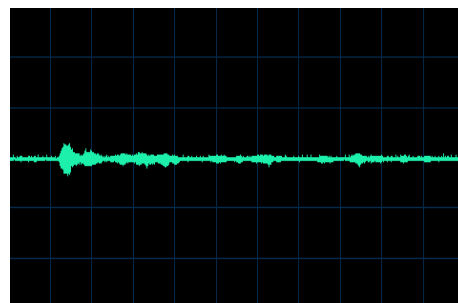


图 6 被叫话机经过回声抵消的语音波形

5 结束语

VoIP 正在成为一个热门的应用领域,语音质量的提高是其难点之一。语音编解码算法的性能将决定所传输的语音质量及占用的网络带宽。Speex 编码算法是一种随着网络质量变化而提供不同传输速率的变速率的编码算法,特别适合于 VoIP 的编解码。本文深入分析了 Speex 语音编码算法中的回声抵消的算法,仿真表明回声抵消使 Speex 算法在 VoIP 中应用时有更好的性能。

参考文献

- [1] 龚耀寰.《自适应滤波器》[M],北京:电子工业出版社,2003。
- [2] Jacob Benesty, Tomas Gansler. A Multidelay Double-Talk Detector Combined with the MDF Adaptive Filter [J]. Applied Signal Processing 2003:11, 1056-1063
- [3] Jean-Marc Valin. On Adjusting the Learning Rate in Frequency Domain Echo Cancellation With Double-Talk [J]. IEEE TRANSACTIONS ON AUDIO, SPEECH & LANGUAGE PROCESSING, VOL. XX, NO. YY.

Analysis And Simulation of the echo cancellation algorithm of Speex

Liu Deqing¹, Huang Xiaojian¹, Li Jing²

1.School of Information and Telecommunication Engineering, Beijing University of Posts and Telecommunications, Beijing (100876)

2.Department of Electronic Engineering, Tsinghua University, Beijing (100084)

Abstract

In VoIP phone, due to the phenomenon that sound broadcasted from loudspeaker could be picked up by microphone and sent back to the remote, together with the effect that network and data processing could bring various delay, remote callers can hear the echo of their own, which seriously affects the quality of communication. Therefore, echo controlling and cancellation is particular important in VoIP system. This paper analyzes echo cancellation algorithm of Speex, and investigates LMS algorithm and the updated principle of adaptive filter coefficients in depth. Besides, the specific fixed-point implementation is introduced. Simulation results demonstrate that the proposed algorithm can effectively improve the quality of voice call.

Keywords: echo cancellation;adaptive filter;fixed point