

一种基于视频会议系统的回声消除算法

作者：吴维农 李果 温永怡
(所在单位：重庆电力调度通信中心)

摘要：回声是影响视频会议系统性能的主要因素之一。本文分析了在视频会议系统中声学回声的形成原理，并介绍了几种主流的回声消除算法。文章分析了各算法的不足，并提出了一种新的回声消除算法，实验证明本文提出的算法是可行的，并且能够达到较好的回声消除效果。

关键词：视频会议；回声消除；AEC；NLMS；LPC

0 引言

视频会议系统采用的是分组交换技术，利用因特网来传输语音数据。这样虽然提高了资源的利用率，同时也存在一些弊端，比如语音质量比较差。导致这一问题的因素很多，一是网络延时和丢包，这主要是由于网络的性能和数据传输特点造成的，这一问题可通过改善网络得到缓解；二是回声，在正常通话过程中，通常会有回声现象，这一现象将直接导致与会方听音不清晰，严重影响会议质量，这一问题通常采用回声消除算法来解决。

1 回声消除算法

当远端语音由近端的扬声器放出来时，在室内的声波可直接传到麦克风或是经由一连串反射再传到麦克风，这些都称为声学回声(acoustic echo)，前者称为短回声(Short echo)，后者称为长回声(Long echo)，当声学回声由麦克风传回远端时，远端就会听到不久之前自己发出的声音，这样会明显降低语音质量，造成不舒适感，这种声学回声在视频会议系统中非常常见。

针对声学回声的形成原理，本文设计了一个回声消除算法对其进行消除处理，算法框图如图 1 所示。

从图 1 可以看出，回声消除算法可分为三个部分：自适应滤波器、语音对讲检测器和残余

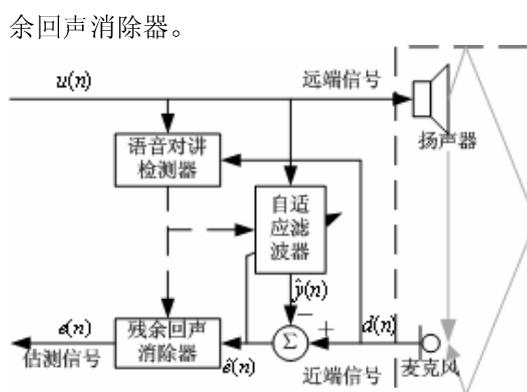


图 1 回声消除算法框图

Fig.1 Block diagram of AEC algorithm

1.1 自适应滤波器

这是回声消除算法的核心部分，常用的自适应算法中，可分为二大系列：最小均方(Least Mean-Square, LMS)算法与递归最小平方(Recursive Least-Square, RLS)算法。LMS 算法使用误差梯度来调整更新自适应滤波器参数，该算法结构简单、速度快、计算量小、稳定性较好；与 LMS 算法相比，RLS 算法执行时收敛速度较快，但是运算量大则是其主要的缺点[1]，这严重的影响了该算法的实用性。因此，实际使用的回声消除算法通常采用 LMS 系列。

本文采用的是一种改进的 LMS 算法——归一化最小均方算法(Normalized LMS, NLMS)。与 LMS 算法相比，NLMS 具有如下优点[2]：①NLMS 算法减轻了 LMS 算法中梯度噪声过大的问题；②NLMS 算法的收敛速度更快[3]。

1.2 语音对讲检测器

语音对讲检测器 (Double Talk Detection, DTD) 用于检测通话过程中远端和近端是否处于讲话的状态。自回声消除器开始发展以来, 许多不同的语音对讲检测器方法开始被提出。针对声学回声消除的应用也有很多学者提出不同的解决方法, 互相关系数即为一种相当不错的判断机制, 它将远端信号与估测信号之间的互相关系数用来判定双向谈话状态。另外有学者使用远端信号与麦克风接收的近端信号构成的互相关系数来判定。这两种算法的思路比较类似, 考虑到估测信号不能完全代表去除回声的近端信号[4], 因此, 本算法采用后者作为语音对讲检测器。

1.3 残余回声消除器

经过 NLMS 算法处理后, 仍然残留少量回声信号, 当近端没有说话时, 通常远端仍然能够听到这些回声, 这些被称为残余回声。残余回声的成分很复杂, 很难像 NLMS 算法那样构造较为简单的模型来模拟, 因而, 在实际应用中, 一直是一个难以解决的问题。本文对残余回声的成分进行了分析, 并提出了一种较为切合实际应用的残余回声信号消除算法。

2 NLMS 算法

图2所示是 NLMS 自适应滤波器结构图, 从图中可以看出, NLMS 算法是由横向滤波器和自适应控制算法组成的, 其中 $u(n)$ 为远端信号, $d(n)$ 为近端信号, $\hat{y}(n)$ 为 NLMS 算法估算出的回声信号, $\hat{e}(n)$ 是估测信号。

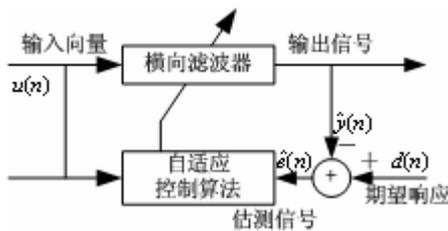


图2 NLMS 算法原理

Fig.2 NLMS algorithm

NLMS 算法的思想是: 假设 $u(n)$ 与 $\hat{y}(n)$ 是多延迟且线性相关的, 通过横向滤波器和自

适应控制算法来逼近回声通道参数 $\hat{w}(n)$, 从 $d(n)$ 中提取出与 $u(n)$ 相关的信号 $\hat{y}(n)$, 并将其扣除, 从而得到去除回声信号的 $e(n)$ 。

其数学表达式[5]如下:

$$\hat{y}(n) = \hat{w}^H(n)u(n) \quad (1)$$

$$\hat{e}(n) = d(n) - \hat{y}(n) \quad (2)$$

$$\hat{e}(n) = d(n) - \hat{w}^H(n)u(n) \quad (3)$$

由(3)式可见, 回声信号能否被滤除干净, 取决于 $\hat{w}(n)$ 是否能够完全的模拟回声通道。在迭代计算中, $\hat{w}(n)$ 的逼近公式如下:

$$\hat{w}(n+1) = \hat{w}(n) + \frac{\tilde{\mu}}{\delta + \|u(n)\|^2} u(n)\hat{e}^*(n) \quad (4)$$

其中:

1) $\tilde{\mu}$ 为步长因子, 是一个正的实数, $\tilde{\mu}$ 的选取至关重要, 为了确保收敛, 必须满足 $0 < \tilde{\mu} < 2$ 。 $\tilde{\mu}$ 取值过大, 虽然可以加快自适应滤波器的收敛速度, 但是 $e(n)$ 中残余的回声信号会更多, $\tilde{\mu}$ 取值过小, 则使收敛速度变慢。

2) δ 为补充因子, 也是一个正的实数, 它是为了防止 $\|u(n)\|^2$ 过小而人为添加的一个参数。

由以上推导可见, 在 $\hat{w}(n)$ 几乎不可能完全模拟出回声通道, 因此, $\hat{y}(n)$ 中也不可能包含 $d(n)$ 中所有的回声信号。实验证明, 在经过 NLMS 算法后, $\hat{e}(n)$ 中仍然残留部分回声信号。

3 残余回声信号处理

经过 NLMS 算法处理后, 接下来需要进一步处理残余的回声信号[6], 经过对残余回声信号的分析, 其中可能包含如下成分:

1) 假设 NLMS 算法中的假设前提成立 (即 $u(n)$ 与 $\hat{y}(n)$ 是多延迟且线性相关的), 则由于 NLMS 算法固有缺陷, 导致有部分线性回声信号残余。

2) $u(n)$ 是待播放的数字信号, 而在扬声器播放和麦克风采样之间, 包括数模/模数转换过程以及语音信号在空间的多径传播, 因此, 很难保证 $u(n)$ 与 $y(n)$ 之间是完全线性

的,也就是说,在残余的回声信号中还包含部分非线性多延迟的回声信号。

由于残余回声信号各成分性质不同,采取的处理方式也就不同,下面将详细介绍残余回声信号的处理方法。

3.1 线性残余回声信号处理

针对这一类信号,我们做如下假设:

1) 在 $d(n)$ 中所有与 $u(n)$ 线性相关的成分都在 $\hat{y}(n)$ 中有所体现。

2) 在 $\hat{e}(n)$ 中的线性残余回声信号与 $\hat{y}(n)$ 是线性相关的。

那么,这类信号可以通过与 $\hat{y}(n)$ 自相关的方式去除。

3.2 非线性残余回声信号处理

这一类信号随机性比较大,因此,这里直接将其视为噪声。然而,它与环境噪声不同,它是由远端声音存在与否决定的,针对这一类信号,我们采用 DTD 判定对讲状态,然后再针对具体情况处理。

1) 仅有近端在说话

此时不存在噪声,不需做任何处理。

2) 仅有远端在说话

在经过 NLMS 和互相关处理后,残余的信号可视为 $r_m(n)$ 。我们可采用 LPC 算法,如下所示:

$$e(k) = r_m(k) - \sum_{i=1}^P a_i(k)r_m(k-i) \quad (5)$$

其中 $a_i(k)(k=1,2,3,\dots,P)$ 是线性预测系数,这里采用 Levinson-Durbin 算法计算这些系数,并随着数据的更新而不断更新。

3) 远端和近端同时在说话

此时停止计算线性预测系数,而调用之前算出的线性预测系数预测当前语音信号中的非线性残余回声信号。

4 仿真实验与分析

接下来,对本文提出的算法进行仿真,实验数据如图 3 所示。从图中可以看出,扬声器播放的信号在麦克风采样的信号中有所体现,而经过回声消除后,麦克风采样信号中的回声

信号基本上被清除了。

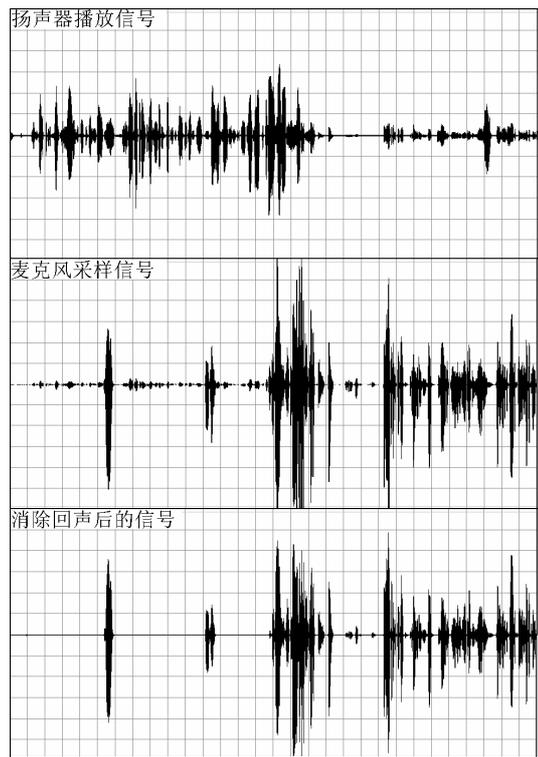


图 3 回声消除算法仿真

Fig.3 Simulation of AEC algorithm

5 结束语

回声消除一直是视频会议系统的难点之一,本文提出的回声消除算法,采用了当前主流的 NLMS 算法,并充分的考虑了视频会议系统的应用环境,有针对性的提出了残余回声信号处理算法。在视频会议系统的实际应用表明,该算法能够较好的消除回声信号,较大的提高了系统的通话质量。

参考文献:

- [1] J. Benesty, T. Gansler, D. R. Morgan, M. M. Sondhi, and S. L. Gay. *Advances in Network and Acoustic Echo Cancellation*, Berlin, Springer-Verlag, 2001
- [2] Simon Haykin. *Adaptive Filter Theory*. Prentice-Hall, 4th Edition 2002
- [3] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, Norwell Kluwer, 1997

- [4] S.M.Kuo and D.Vijayan, A secondary Path modeling technique for active noise control system, IEEE Trans. Speech Audio Processing, Vol. 5, P374-377, July. 1997.
- [5] S.L.Gay, and J.Benesty, Acoustical Signal Processing for Telecommu-
cation, Norwell, Kluwer, 2000.
- [6] V.Turbin, A.Gilloire, and P.Scalart, Comparison of three post-filter algorithms for residual acoustic echo reduction, IEEE Int. Conf. Acoustics, Speech, and Signal Processing, P307-310, 1997

Acoustic-Echo Cancellation Algorithm for Video Conference System

Abstract: The acoustic-echo is one of the main factors that influences the video conference system function. This article analyzed the element of the echo in video conference system, and introduce a few familiar acoustic-echo cancellation algorithm. The article analyzed the shortage of each algorithm, and put forward a kind of new acoustic-echo cancellation algorithm, the experiment proves this kind of algorithm can cancel echo effectively.

Keywords: Video Conference; AEC; NLMS; LPC