自适应回声消除原理

# 摘要

实时通话对于人们来说越来越重要，尤其是在双工通话过程中，由于扬声器和麦克风会同时工作，麦克风总会采集到扬声器发出的声音，麦克风采集到的这各声音会传回远端，影响通话质量，这就是声学回声的过程，所以声学回声消除对于提高通话质量就显得尤为重要。本文通过分析几种基本的和改进版的自适应滤波器，分析不同的自适应滤波器在不考虑近端说话人时的回声消除中的原理实现，以及对回声的消除能力以及收敛速度和运算时间进行对比，结合实际数据集上的测试结果来分析各滤波器的优劣。通过测试结果得出哪种滤波器能更好的在实际工程中应用，以及只有自适应滤波器还不足以完全消除回声的原因。

# 引言

声学回声是指扬声器播出的声音在近端说话人听到的同时，也通过多种路径被麦克风拾取到传回远端。多路径反射的结果产生了不同延时的回声，包括直接回声和间接回声，直接回声是指由声器播出的声音未经任何反射直接进入麦克风，间接回声是指由扬声器播出的声音经过不同的路径(如房屋或房屋内的任何物体)的一次或多次反射后进入麦克风所产生的回声的集合。房屋内的任何物体的任何变动都会改变回声的通道，因此，这种回声的特点是多路径的、时变的。自适应回声消除的原理的基本思想是估计回声路径参数，模拟声音从扬声器播放在封闭环境内传播的路径得出一个混响声音，从麦克风采集的声音中减去这个混响声音。

下图为自适应滤波器的原理框图，在不考虑近端说话人即s(n)为0的情况下，x(n)代表来自远端的信号, d(n)是经过回声通道而产生的回声。D端是近端麦克风，麦克风只采集到的房间叠加的回声。对回声消除器来说，接收到的远端信号作为一个参考信号，回声消除器根据参考信号由自适应滤波器产生回声的估计值，将从近端带有回声的语音信号减去，就得到近端传送出去的信号 。在理想情况下，经过回声消除器处理后，残留的回声误差降为0，从而实现回声消除。

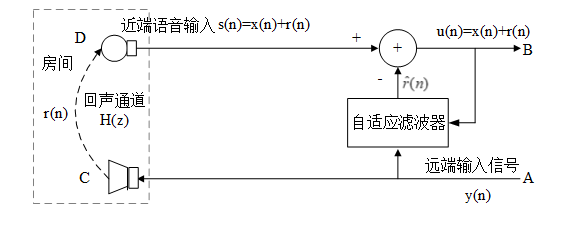


图1 自适应滤波器框图

由于通话过程中的环境非常复杂，所以对自适应滤波器的设计有一下性能指标：

1. 收敛速度：麦克风或者扬声器在房间中的位置总是不固定的，所以麦克风采集到的扬声器播放的混响声音总是在变化的，即混响路径总是在变化，所以滤波器的收敛速度一定要快。
2. 算法复杂度：通话要考虑实时性的问题，那么良好的自适应滤波器应该保持收敛速度的同时尽量降低算法复杂度。

# 自适应滤波器

## 1.1最小均方算法(Least Mean Square, LMS)

LMS算法是线性自适应滤波器也是随机梯度下降法中的重要一员，在1960年由Widrow和Hoff提出。其特点是算法简单，通常分为以下两个步骤：第一步，通过输入信号计算出线性滤波器的响应；第二步，通过计算出期望信号与输出信号的误差来调整滤波器参数。

设计一个N阶的自适应滤波器，每次输入的参考信号x(n)长度为M，则滤波器参数w(n)的大小为N\*M，则滤波器的输出为



则y(n)为在室内环境中估计的扬声器播放的远端语音，而在麦克风处采集的真实由扬声器播放的远端语音作为我们的期望信号d(n)，则误差信号e(n)可以通过d(n)与y(n)求差值得到：



我们的目标就是将误差e(n)最小化，并将e(n)传输至远端通话者，采用最小均方误差(Minimum Mean Square Error, MMSE)准则来最小化代价函数J(w)：



通过计算代价函数J(w)对w的导数，令其导数为0：



根据最速下降法原理则滤波器系数的更新公式为：



其中μ为步长因子，μ越大，算法收敛越快，但稳态误差也越大；μ越小，算法收敛越慢，但稳态误差也越小。为了保证算法能得到一个稳态解的w(n)，应该使得μ在以下范围内取值，通常来说μ值较小能有效滤除掉梯度噪声。

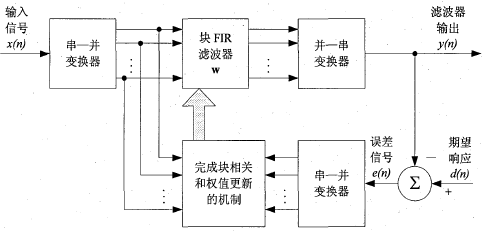


LMS算法是自适应滤波器算法中最基础的算法，其原理简单，也是众多其它LMS变种算法的基础，但是在实际回声消除中却不会采用，原因有以下：LMS算法滤波器的系数是逐点更新的，因为实际通话过程中是拿不到未来语音信息的，所以每一点采样求解梯度总会存在误差；由于系数是逐点更新的，所以跟踪性能较差，当通话环境发生变化时，滤波器性能会严重下降.

## 1.2分块最小均方算法(Block LMS)

为了进一步提高LMS算法的跟踪性能，在LMS算法的基础上提出了分块处理的LMS算法。LMS算法时每进来一个采样点就更新一次滤波器参数，而Block LMS算法时每K个采样点才更新一次滤波器参数。

输入的序列x(n)经过一个串—并变换器之后被分成若干个长度为L的数据块，然后将这些数据块一次一块的输入阶数为M的块滤波器中，在每收集一块数据后，进行滤波器系数更新，而不像LMS算法那样是逐点更新的。



在累计了L个采样点之后，就可以得到Block LMS滤波器权重更新的公式：



其中，μ为步长因子，Block LMS更新一次参数相当于LMS更新了L次，其优点是运算量比LMS的小得多，收敛速度也能和LMS算法保持一致。

## 1.3归一化最小均方算法(Normalized Least Mean Squares, NLMS)

NLMS算法同样是LMS算法的改进版，其解决的问题是滤波器参数调整会与输入值成正比。如下列公式所示，w(n+1)的值会受输入信号x(n)的影响，如果x(n)的变化非常大，那么就会放大梯度噪声。



所以我们要对输入信号x(n)进行变换：



从上式可以看出，步长μ(n)与x(n)成反比，步长从固定值μ变为自适应值μ(n)。而在NLMS的基础上为了增强此算法的鲁棒性，引入了两个超参数α和β，μ(n)变为下列式子：



其中β一般设为一个较小的整数，防止输入信号x(n)的内积过小使得μ(n)过大而引起滤波器稳定性下降，其中α为新的步长因子。

NLMS算法通过归一化的操作改善了LMS算法收敛速度慢的缺点，但是x(n)之间的相关性比较大时，会影响μ(n)使得收敛速度明显下降。

## 1.3变步长LMS(Variable Step-Size LMS, VSS LMS)

VSS LMS算法的主要思想是在滤波器工作的初始阶段采用较大的μ值，以加快收敛速度，而在后面阶段采用减小的步长值μ以减小稳态误差。其中最关键的是怎么调整整个过程中μ值的变化。





其中两个超参数α和γ的作用分别是平滑和调整误差信号，在实际处理过程中还会加上截幅操作：



传统的LMS算法通常采用固定步长μ，且有时候为了减小自适应算法的结果失调程度，通常的μ值都较小，其，代价就需要更长的时间来达到稳态解，所以可变步长的μ就显得非常重要。

## 1.4递归最小二乘法(Recursive Least Square, RLS)

MMSE准则适合处理平稳序列，因为MMSE是一个均匀加权的最优化问题，也就是说每一个时刻的误差信号对代价函数的贡献是相同的。RLS重新定义了代价函数：



其中λ为遗忘因子，取值在0<λ≤1，结合上述公式，当时刻i离n越近对误差e(n)的加权就越大遗忘的越少，而距离n时刻越远的误差加权就越小，遗忘的越多。

标准RLS算法的执行流程：

1.初始化权重值为w(0)=0，初始化输入信号的自相关矩阵的逆矩阵p(n)，p(0)= ，δ是一很小的正数，I是单位矩阵

2.计算先验误差：

3.计算增益向量：

4.逆矩阵更新：

5.滤波器更新：

RLS算法的基于的是先验误差修正参数w(n)，而LMS算法是基于后验误差，所以RLS算法的收敛速度和LMS算法相比，其跟踪性能和收敛速度都远优于LMS算法，但是代价是计算量非常大，且输入信号的自相关矩阵接近奇异时RLS收敛速度和跟踪性能会严重恶化。

## 1.5 比例归一化最小均方算法(Proportionate Normalized Least Mean Square, PNLMS)

通过对回声路径模型的分析，发现回声能量中较活跃系数均在时域聚集，且比重很小，其数值只有很少不为零的有效值，大多数都是零值左右，这就是回声路径具有的稀疏特性。而由于回声路径的稀疏特性，如果滤波器具有很长的阶数，实际过程只有一小部分在发挥实际作用。因此如果回声路径比较长的时候，基本的LMS以及NLMS算法的收敛速度是达不到要求的，因此根据回声稀疏特性的原理，提出了PNLMS算法，其算法的核心思想便是按比例分配滤波器的权值向量大小，该算法对回声消除的发展具有非常重要的意义。

PNLMS 算法拥有初始收敛速度的优势，尤其在回声路径具有稀疏性条件下，其收敛速度则更加明显。PNLMS算法采用与滤波器抽头稀疏成正比的可变步长参数来调整算法收敛速度，利用其抽头稀疏的比例值来判断当前权重稀疏所属的活跃状态，根据状态的不同，所分配的步长大小也有所差异，活跃抽头系数分配较大的步长参数，这样可以加速其收敛，而不活跃的抽头系数则相反，通过分配其较小的步长参数来提高算法的稳态误差。每个滤波器抽头被分别赋予了不同估计值，算法的稳态收敛性得到了明显改善。滤波器权值更新公式可表示为：



式中G(n)为N维的对角增益矩阵，对滤波器各权值系数分配步长权重，表示为



其中



其中



δp 为在滤波器初始阶段调整更新参数的因子，ρ的取值将影响算法的整体收敛速度，较大的ρ值会使得收敛速度减缓，但也不宜取值过小，一般取5/M。若将ρ取值为 1，则PNLMS 将变为 NLMS 算法。

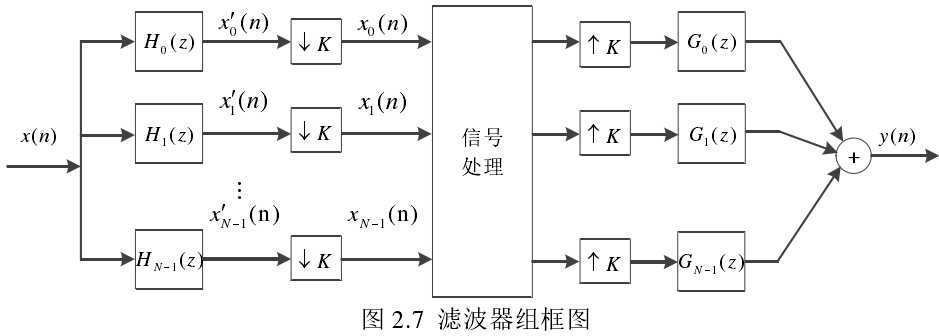
PNLMS算法对于稀疏的回声路径，在初始阶段拥有快速的收敛速率，与此同时降低了稳态误差，所以如果对于非稀疏的情况则收敛速度会严重下降。

## 1.6子带自适应滤波器(Subband Adaptive Filter, SAF)

在声学回声消除的应用中，远端输入语音信号的相关性较高，然而传统的全带LMS算法和NLMS算法等都是基于信号无关性的前提的，所以很难满足系统对收敛速度的要求。远端信号的相关性在时域表现为语音信号的相关矩阵特征值的扩散度，在频域上的表现为语音的频谱动态范围。

一般来说，语音信号都具有很大的相关性，因为相比于白色信号，语音信号的频谱动态范围明显更大。因此为了提高LMS算法的收敛速度，其中一种方法就是降低输入信号的相关性，而另一种方法就是子带自适应滤波算法，子带自适应滤波器算法是基于频域对信号进行处理。

子带自适应滤波器：将相关信号通过滤波器组分割成近似无关的各子带独立信号，然后对子带信号进行多速率抽取来获得采样信号，再进行信号的自适应处理。用于子带自适应滤波器的多速率抽取系统有下采样和上采样两种，主要通过抽取和插值两种方法来使系统获得不同的采样率。设采样因子是K，通过保留信号的K倍采样点可将采样速率从f减小为f/K，这样就降低了自适应算法的计算量。反之，上采样通过在信号相邻点间插入K-1个0使原采样率从f增大为Kf。



上图是一个N个子带的滤波器组，其中和分别是分析滤波器组和综合滤波器组的传递函数，分析滤波器组作用是将输入信号分割为子带信号，综合滤波器的作用是将子带信号重建为输出信号，在分析滤波器之后经过下采样在进行分块的自适应滤波，最后通过上采样和综合滤波器来重建信号。

SAF中的关键步骤就是分子带，且有个假设就是子带在频域上不会重叠，这在实际上是不可能实现的，所以在重建信号时总会引入误差。

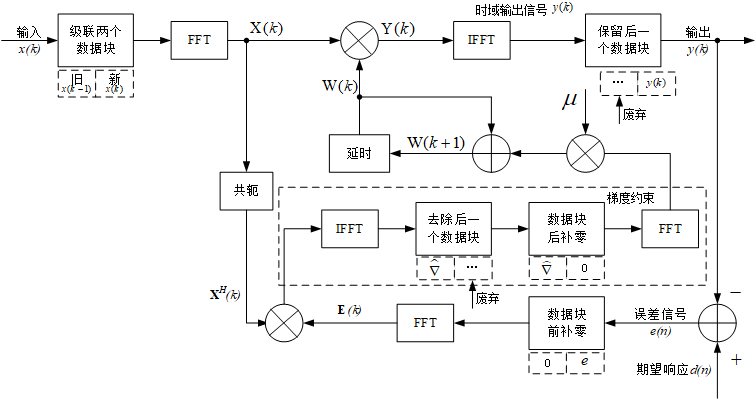
## 1.7频域分块自适应滤波器(Partioned Block Frequency Domain Adaptive Filter, PFADF)

频域分块自适应滤波器是现实回声消除工程中用的最多的自适应滤波器算法，由于回声客观的时延存在，且由于通话工程需要收发同步，那么实际中对滤波器的阶数要求非常高，起码是数以千计的，如果在时域中进行处理，其计算量会非常大，而如果在频域中计算可以节省运算量。

时域中自适应滤波算法求解线性卷积的运算量非常大，尤其是当回声路径长而复杂的时候，所以需要通过PFDAF算法将自适应滤波器的阶数分成FFT长度的整数倍子块，对输入信号的每个子块做FFT，然后在频域内进行自适应滤波。

在实际过程中通常使用重叠存储法或者重叠相加法计算线性卷积，为了保证有N个点的线性卷积和圆周卷积的结果一致，通常输入信号的长度为滤波器阶数的两倍即每次参加运算的数据为2N个点。

基于重叠保留法的流程图如下所示：



# 实验

## 2.1实验参数设置

在回声消除的常用数据集里挑选100条麦克风语音和对应数量的远端语音来进行测试，对比本文中提到的自适应滤波器算法。由于使用的数据集都是在实际通话过程中录制的，所以为了能发挥自适应滤波器的最佳性能，多次调参选出表现最好的那组参数进行对比，因为在阶数一定的前提下，阶数越高性能越好，但相应的延迟也会越高，而步长因子μ则通过多次实验选出表现最好的那次步长值。

自适应滤波器的输出结果通过计算SER来进行对比，其公式如下所示，其中mic是麦克风语音，err是残留的回声噪声，其比值越大那么算法的性能越好。



实验的条件是算法的时延不超过两秒，调整滤波器参数和步长值找出滤波器的最优结果。

## 2.2各滤波器最优结果对比

从表1看出，性能结果最好的就是NLMS算法，而基于先验误差的两种算法：RLS算法和AP算法程序上还存在bug，在随后的课题研究上再做补充。而基于频率域的自适应滤波器算法，在最后的性能表现上SER非常低，不符合理论表现，后面会给出我自己的分析。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | LMS | Block\_LMS | NLMS | VSS\_LMS | RLS | AP | SAF | PFADF |
| SER | 2.99dB | 1.79dB | 5.10dB | 1.27dB | X | X | 1.02dB | 0.06dB |
| Times | 1.48s | 0.45s | 1.56s | 1.60s | X | X | 2.20s | 0.95s |

表1 声源不移动条件下的SER

## 2.2.1时域LMS自适应滤波器算法实验结果分析

仅从现有的实验结果来看，表现最好的是归一化的最小均方误差算法(NLMS)，相比于其他三种同类型的LMS算法：LMS、分块LMS以及变步长LMS，NLMS的优势在于其步长值是根据输入信号的大小自适应修正步长的大小，因此抗发散的能力更好，所以在调参的时候可以设置更长的滤波器阶数，有更多的调参空间，从而达到更好的性能。而对于其他三种LMS算法则受限于阶数不能太长，否则滤波器的输出就会严重发散。

LMS、NLMS、VSS\_LMS三种算法的计算时间相差不大，从实验结果也能看出进行分块操作的Block\_LMS算法运算速度是最快的，因为它不是按点更新权重而是按分块的，每一个数据块运算完再计算一次新的权重，但是由于是分块更新的准确度总是比不过逐点更新的其他三种自适应滤波器算法。

## 2.2.2时域基于先验误差的自适应滤波器算法

RLS算法和AP算法都是基于先验误差的修正滤波器的算法，所以在运算过程中涉及大量的矩阵求逆运算，尤其是遇到奇异矩阵运算速度就更慢，所谓的奇异矩阵指的就是语音的相关性比较高，所以在构成输入矩阵的时候不满秩，此时的求逆将会异常困难。而RLS算法运行了一晚上没有把200条语音的数据集运行完，所以不参与此次实验的对比，在随后的研究中会继续研究修改。

## 2.2.3基于频域的自适应滤波器算法

基于频率域的LMS算法在理论性能上应该比基于时域的LMS算法是更优秀的，但是在本次实验结果来看，却与理论性能相差甚远。频域自适应滤波器解决的就是当回声路径过长的问题。其思路是在频域分块上进行处理，不仅能够通过频域乘法实现滤波卷积，而且有效地减少了自适应滤波器的长度，从而降低了自适应算法的计算复杂度。此外，为了降低计算复杂度，FDAF还可以提供更高的收敛速度。这是由于滤波器更新时信号自相关矩阵的特征值扩散减小所致。

## 2.3实验结果与总结

从实验结果来看，只依赖于自适应滤波器去消除回声是不足够的，通常来说还要结合上非线性抑制滤波器来完成残余回声的消除，通常这一模块才是AEC中提升SER最重要的部分。而且自适应滤波器在通话过程中真正起作用的也是在单讲情况下，在这个情况下才能准确估计回声路径，而对于双讲情况是无法估计路径的。所以在真实应用中的回声消除一定要很好地结合前端的阵列和噪声抑制来完成保持近端语音的同时又很好地消除回声。

从上述实验结果来看，虽然基于频域的自适应滤波器的实际性能没那么理想，但是在实际过程中，回声消除不仅仅只考虑性能，还要保证时延的降低，所以基于频率域的自适应滤波器在实际应用还是更为普遍。

# 文献引用

[1] Widrow B, Hoff M E. Adaptive switching circuits[R]. Stanford Univ Ca Stanford Electronics Labs, 1960.

[2] Shin H C, Sayed A H, Song W J. Variable step-size NLMS and affine projection algorithms[J]. IEEE signal processing letters, 2004, 11(2): 132-135.

[3] Farhang-Boroujeny B, Chan K S. Analysis of the frequency-domain block LMS algorithm[J]. IEEE Transactions on Signal Processing, 2000, 48(8): 2332-2342.

[4] Akhtar M T, Abe M, Kawamata M. A new variable step size LMS algorithm-based method for improved online secondary path modeling in active noise control systems[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2006, 14(2): 720-726.

[5] Lee K A, Gan W S, Kuo S M. Subband adaptive filtering: theory and implementation[M]. John Wiley & Sons, 2009.

[6] Bershad N J, Feintuch P L. Analysis of the frequency domain adaptive filter[J]. Proceedings of the IEEE, 1979, 67(12): 1658-1659.

[7] Rogers S. Adaptive filter theory: by Simon HAYKIN. Prentice Hall; Upper Saddle River, NJ, USA; 1996; 989 pp.; $84; ISBN: 0-13-322760-X[J]. 1996.