

Adaptive Filter for Audio Noise Reduction Basing on Gradient Descent

Xiaohong Nie

Jiangsu Wireless Power Corporation, Nanjing Jiangsu
Email: 251573056@qq.com

Received: Jul. 14th, 2020; accepted: Jul. 29th, 2020; published: Aug. 5th, 2020

Abstract

Filter skill deals with noise in background to compensate channel's noise and reduce it. Commonly, we call the processor filter. The paper computes filter's coefficient for minimum variance. The paper advises that one filter can reduce channel's noise, base on gradient descent. For minimum variance between standard signal and noising signal, the filter's coefficients will iterate to the last result. The simulation shows the filter can reduce noise to enhance SNR.

Keywords

Filter Coefficient, LMS, Gradient, Adapt, Audio, Noise

基于梯度下降的自适应音频去噪声的研究

聂晓鸿

江苏无线电厂, 江苏 南京
Email: 251573056@qq.com

收稿日期: 2020年7月14日; 录用日期: 2020年7月29日; 发布日期: 2020年8月5日

摘 要

去噪声技术是对音频信号进行适当处理, 以补偿噪声畸变的技术, 通常把采用自适应去噪声的处理器称为自适应滤波器。本文方案以最小均方误差为准则来计算得到滤波器系数。本文设计了一种采用梯度下降的自适应算法的自适应调整滤波器, 根据标准参考信号和混有噪声信号进行自适应滤波运算, 迭代递归过程以混有噪声信号与标准参考信号最小均方误差为目标, 仿真结果显示这个滤波器可以很好地降低噪声提高信噪比。

关键词

滤波器系数, 最小均方误差, 梯度, 自适应, 语音, 噪声



1. 引言

在语音处理及复杂背景环境下音频数据交互系统中，由于各种电子设备功放喇叭噪声、外界干扰、语音回声、背景噪声的客观必然性，导致语音信号被数字采集接收时将不可避免地产生语音背景噪声或干扰，语言采集是系统前端采集部分，如果语音前端采集存在噪声，那么在后面处理阶段比如识别语音时会产生较大的误差。语音采集过程一般是语言经过麦克后振动产生电信号，然后由 ADC 进行数字采集得到数字语音信号。目前语音采集系统已经非常普遍地采用了高采样率及位宽的 ADC 模数转换器，ADC 模数转换器的数字采集频率可以达到 48 kHz 以上，采集位宽可达到 16 bit 以上。在某些对语音要求比较高的场合里，比如对语音有多样性处理要求的声卡，多样性语音处理声卡系统的语音采集频率可以达到 96 kHz，采集位宽达到 24 bit。理论上 ADC 采样的位宽越高、采样率越高，数字语音信号描述模拟语音信号的量化误差就会越小。

目前市场上语音识别技术应用前景十分广泛，但是这块领域还没有达到图像识别领域的成熟程度，其中一个主要限制因素就是在各种各样的场合环境下，语音信号的状态有很大差异，有些是语音发音差异的原因，有些是因为背景噪声、回声过大，比如船上、汽车上的背景噪声特别大。在智能音箱的语音激活检测中，在近场条件下信号比较清晰而且噪声小，简单算法就可以做语音激活检测。但在远场条件下，用户不能很近地接触智能音箱的语音采集系统，这时候人说话的语音小，其它噪声大比如水流声、电视语音、开水沸腾声、洗衣机噪声、开关门声、楼梯脚步声等等，做语音激活检测功能就必须采用复杂的处理方式。所以工业界也一直在努力解决这个问题，在这块领域有很多方向，比如谱减法、小波滤波、维纳滤波等。不管是谱减法、维纳滤波还是小波滤波法，在时域上的数学表达都等价于滤波器响应表示，因此通过各种算法处理方式去减少语音噪声，滤波器是其中一种重要的消除噪声的方法，得到降噪高性能的滤波器参数是一个重要的研究方向。普通数字低通滤波器是通过频域整块处理掉带外噪声，但带内噪声很难被处理，为了特殊地改变滤波器参数从而提高消除噪声干扰的性能，作者利用通用的自适应调整滤波器系数算法，使用仿真工具统计对比无处理的有噪声语音信号和处理滤波后的语音信号，结果显示自适应动态调整滤波系数的滤波器可以很好地消除干扰噪声。

2. 自适应滤波数学理论

自适应调整参数滤波器的架构由滤波系数乘法器、延时单元和加法器组成。假设滤波器的滤波系数为 w_n ，当输入为 $w(t)$ 时，输出有：
$$y(t) = \sum_{n=-N}^N w_n x(t - nT_s)$$
。自适应滤波器处理结果输出 $y(t)$ 为输入 $w(t)$ 经过 $2N + 1$ 个不同时延单元输出结果的累加和。在 $t = kT_s$ 时进行离散采样取值，根据计算累加得到自适应滤波输出结果为：
$$y(kT_s) = \sum_{n=-N}^n w_n x(kT_s - nT_s)$$
，简写为 $y_k = \sum_{n=-N}^N w_n x_{k-n}$ 。

为降低系统采样时由外界产生的噪声，需要动态调整自适应滤波系数，使得以 k 为中心的前后 N 个单元在采样时时间 $t = kT_s$ 的噪声值趋向于零[1]。以得到极值为目标从而得到动态调整滤波器的系数，得到计算出极值的损失函数的方法有两种方式依据：以最小峰值畸变为准则得到极值和以最小均方误差为准则得到极值。其中，迫零算法为准则得到极值在处理的过程中忽略了加性噪声，会导致滤波器输出降噪性能下降。所以，本文选择得到最小均方误差为准则得到损失函数模型计算出最小均方误差。

3. 迭代计算实现理论

迭代计算的主要思想是先统计数据特性选取一个初始值，然后更加前面时间序列上的信号和数据特性来按一定学习因子步长比例的反复的累加减，每次累加减的结果都是以最小均方误差为准则去累加减，最终在迭代一定次数后，系统收敛到一个稳定状态。一般这个初始值会根据滤波器特性和随机性选择一个合适的初始值，同样的每次迭代的学习因子步长也会根据实际情况选择一个合适的初始数值，这样可以避免迭代计算进入不收敛状态或者学习停止状态。

递归迭代算法的原理

为了简单使用下面给出的自适应滤波器算法并保证算法的收敛和稳定性，有必要对信号的本质从概率的角度上做一些基本假设，输入信号可以看作是随机变量的一个矩阵[2]。首先，信号都具有普遍性，处理输入信号的基本统计方法有很多方式，如统计均值：

$$\eta = E\{x\} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \quad (1)$$

评估滤波器系统输出结果误差需要使用函数来表示。用离散的差的累加和取绝对值去估算误差：

$$e[n] = d[n] - y[n] \quad (2)$$

其中 $d[n]$ 是参考的标准信号， $y[n]$ 是通过自适应滤波器计算的输出值结果。一般常用的成本损失函数是基于最小二乘法的均方误差函数：

$$j = E\{e[n]^2\} = \overline{(d[n] - y[n])^2} \quad (3)$$

使用它可以降低自适应算法的计算负担[1]。

自适应滤波器的输出是通过卷积[3]计算得到的。其中滤波器的卷积系数 f_k 要以定义的成本函数 J 为最小值的方向去求极值。公式 $y[n] = x^T[n] f = f^T x[n]$ 用矩阵相乘方便的改写了累加和表达。其中 $f = [f_0, f_1, \dots, f_{L-1}]^T$ ，是 $(L \times 1)$ 阶矩阵，符号 T 表示矩阵的转置。代入矩阵相乘得到：

$$e[n] = d[n] - y[n] = d[n] - f^T x[n] \quad (4)$$

根据推导，均方误差函数可以表示为：

$$J = E\{e^2[n]\} = E\{d[n] - y[n]\}^2 = E\{d[n] - f^T x[n]\}^2 \\ E\{d[n] - 2d[n]f^T x[n] + f^T x[n]x[n]\} \quad (5)$$

将上式对 f 求导数求极值，原则是取函数的最小值，令微分后得到的梯度等于 0。结果：

$$\nabla = \frac{\partial J}{\partial f^T} = E\{-2d[n]x[n] + 2x[n]x[n]f_{opt}\} = 0 \quad (6)$$

可以得到：

$$f_{opt} = E\{x[n]x^T[n]\}E\{d[n]x[n]\} = R_{xx}^{-1}R_{dx} \quad (7)$$

按得到最小极值的方式计算得到自适应滤波器系数 f_{opt} 。

让均方误差取最小极值用最小二乘法计算，矩阵计算量大，所以本文采用梯度下降方法来计算实现。

LMS 算法也是最速下降法的一种实现[2]，根据梯度下降的方式推导得：

$$f[n+1] = f[n] - \frac{\mu}{2} \nabla[n] \quad (8)$$

根据 $\nabla[n] = -2e[n]x[n]$ 推导结果为:

$$f[n+1] = f[n] + \mu e[n]x[n] \tag{9}$$

4. 迭代计算自适应去噪滤波器系数的数学原理建模仿真实现

为了保证高性能同时利于工程实现, 自适应滤波器的阶数选择为 32, 采样点数为 27,960 个点, 音频信号是一段中文语音信号, 经过噪声干扰后, 在处理端以最小均方误差为目标迭代计算出滤波器系数, 然后通过自适应滤波器得到输出信号, 并统计对比有滤波器处理后的信号和原始带噪声信号的信噪比, 仿真流程图见图 1。

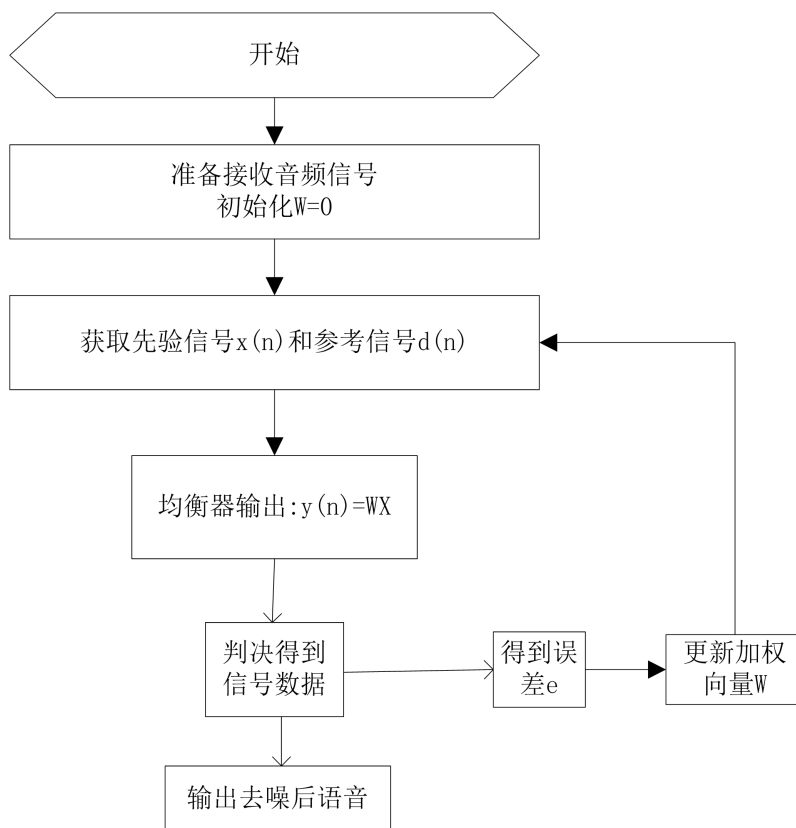


Figure 1. Theory of adaptive filter

图 1. 自适应原理实现图

首先, 在数学建模仿真时, 有个很重要的参数需要处理: 需要选取一个合适的学习步长因子参数 μ 。当学习步长因子参数 μ 选取过小时, 在计算自适应滤波器系数时, 在梯度方向上前进太小导致系统迭代进入学习停止状态, 迭代次数逼近无穷大, 这种过长学习停止状态显然违背了本文要易于工程化应用实现的初衷, 所以学习步长因子参数 μ 不应该过小。当学习步长因子参数 μ 选取过大时, 在计算自适应滤波器系数时, 在梯度方向上前进太大导致系统迭代进入超出正常数值范围即所谓梯度爆炸, 计算出的自适应滤波器系数前后差值非常大超过 0.5 比值, 系统计算系数进入了不收敛不稳定状态, 这种状态计算出来的系数显然是不具备正常降噪功能的, 所以学习步长因子参数 μ 不应该过大。所以作者对学习步长因子参数 μ 做了指数化动态缩减处理, 不完全和 ADAM 的自适应调整学习因子系数一样, 但目标方向思想是一致的: 即判断梯度方向最大量的大小, 如果梯度方向最大量很大那么就动态降低学习步长因子

参数 μ ，这样可以避免梯度爆炸。反之，梯度方向最大量很小那么就动态加大学习步长因子参数 μ ，这样可以避免系统过早进入学习停止状态。

另外，在数学建模仿真时，自适应滤波器的初始值选取也很重要：需要选取一个合适的滤波器初始值。当滤波器初始值很大或者过小从而偏离最终收敛状态数值时，系统需要花很长迭代次数和时间去让滤波器系数学习收敛到最终状态。所以自适应滤波器的初始值选取一个比较接近最终收敛状态数值是非常有意义和必要的。本文对滤波器系数初始值的选取方法：随机生成一组随机数，然后做归一化处理。这样可以提高系统工作计算的效率，减少不必要的一些迭代计算。

自适应滤波器仿真结果对比语音数据如图 2 所示。

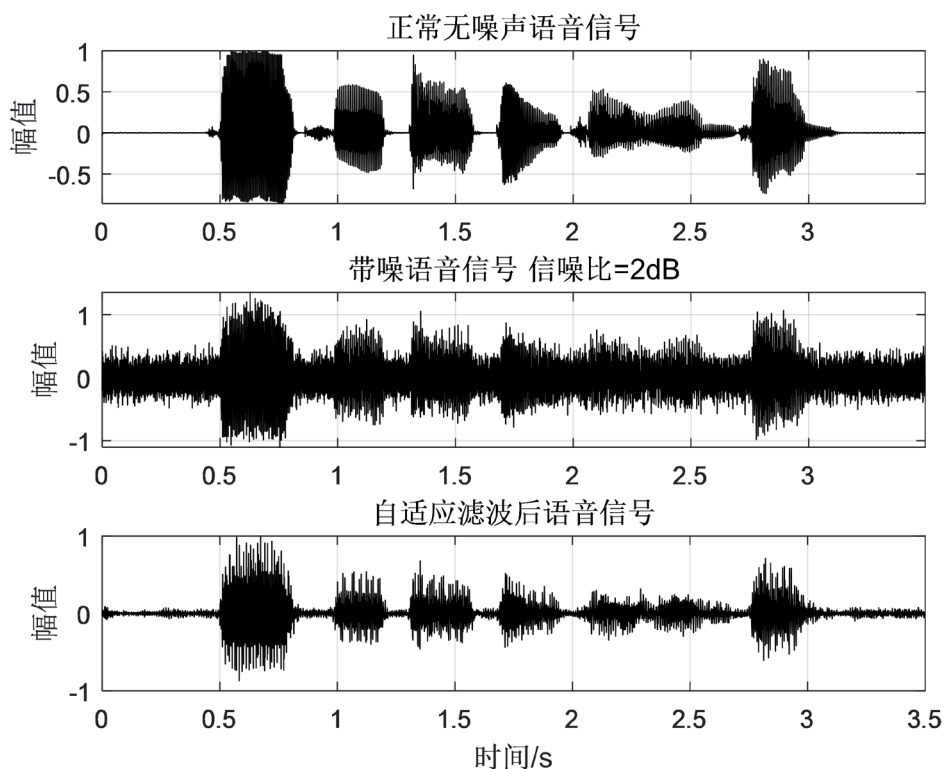


Figure 2. Statistics the signal of audios
图 2. 仿真统计对比的语音信号图

从图 2 中统计的语音信号对比中我们可看出，利用了自适应滤波器技术的语言型号噪声比较小，所以仿真结果证明了自适应滤波器能够降低接收端的噪声，得到更优化的语音信号数据。

5. 迭代计算自适应去噪滤波器系数的数学原理建模仿真实现

目前在语音增强去噪声这块领域使用深度学习神经网络去噪声是一个非常火热的方向，原理也是输入标准音频信号和混有噪声的信号预处理后送入全连接神经网络，根据标准参考信号和混有噪声信号进行误差运算训练神经网络参数，神经网络输入层、隐藏层、输出层参数迭代学习过程以混有噪声信号与标准参考信号最小均方误差为目标。基于深度神经网络的音频降噪方法包括：对数据进行预处理，得到带噪音频数据；训练 DNN 音频降噪模型，得到的神经网络音频降噪模型可以完成音色转换后的歌声的对数功率谱与纯净音频的对数功率谱之间的映射；结合训练好的神经网络音频降噪模型，输出的对数功率谱及相位信息，得到降噪后的音频数据[4]。

但是神经网络学习的过程比较复杂，计算量大，针对一种背景噪声训练一个成熟模型，如果换个背景噪声环境需要重新训练神经网络模型。

本文设计的自适应滤波器属于经典信号处理范围。相对于深度学习去噪声这个新领域，本文设计的方法计算比较简单易工程化实现，尤其是在一些低功耗的经济型的嵌入式的音频采集小系统里面易于工程化实现。

6. 结论

现实生活中的语音不可避免地要受到周围环境的影响，很强的背景噪声例如机械噪声、其他说话者的话音等均会严重地影响语音信号的质量；此外传输系统本身也会产生各种噪声，因此在接收端的信号为带噪语音信号[5]，实际语音处理系统处在不同条件下，系统背景会存在各种噪声、干扰等，甚至在这种条件下说话人自身状态也发生变化导致其说话发声也改变了清晰的状态，导致系统语音采集端采集的语音信号有噪声。在语音系统前端需要采用一些前端处理来降低噪声才可以保证语音的正确接收，自适应滤波技术就是通过提取噪声分布特征和信号特征差异来消除噪声的一种技术。作者做的主要工作便是让语音信号经过递归最小二乘法的自适应滤波器，输出噪声较小的结果，从仿真结果对比可以看出很好地消除了噪声，由于梯度下降计算复杂度已经降低很多，所以这种自适应滤波在工程上的应用前景很大。

参考文献

- [1] Douglas, S.C., Zhu, Q. and Smith, K.F. (1988) A Pipelined LMS Adaptive FIR Filter Architecture without Adaptation Delay. *IEEE Transactions on Signal Processing*, **46**, 175-179.
- [2] Simon Haykin 自适应滤波器原理[M]. 郑宝玉, 等, 译. 北京: 电子工业出版社, 2010.
- [3] 迈耶-贝斯. 数字信号处理的 FPGA 实现[M]. 陈青华, 等, 译. 北京: 清华大学出版社, 2017.
- [4] 卿朝进, 等. 一种基于绝对听觉阈值的深度学习语音增强方法[P]. CN109671446A 2019.10.12.
- [5] 赵力. 语音信号处理[M]. 北京: 机械工业出版社, 2019.