

移动声学回声介绍

生成日期: 2025-10-30

WebRtcAec_Process接口如上, 参数reported_delay_ms为当前设备需要调整延时的目标值。如某Android设备固定延时为400ms左右, 400ms已经超出滤波器覆盖的延时范围, 至少需要调整300ms延时, 才能满足回声消除没有回声的要求。固定延时调整在WebRTCAEC算法开始之初作用一次, 为什么target_delay是这么计算? $\text{int target_delay} = \text{startup_size_ms} * \text{self} \rightarrow \text{rate_factor} * 8; \text{startup_size_ms}$ 其实就是设置下去的reported_delay_ms, 这一步将计算时间毫秒转化为样本点数, 16000hz采样中, 10ms表示160个样本点, 因此target_delay实际就是需要调整的目标样本点数

$\text{aecpc} \rightarrow \text{rate_factor} = \text{aecpc} \rightarrow \text{splitSampFreq} / 8000 = 2$ 我们用330ms延时的数据测试: 如果设置默认延时为240ms, overhead_elements次被调整了-60个block, 负值表示向前查找, 正好为 $60 * 4 = 240\text{ms}$, 之后线性滤波器固定index=24, 表示 $24 * 4 = 96\text{ms}$ 延时, 二者之和约等于330ms, ②大延时检测是基于远近端数据相似性在远端大缓存中查找相似的帧的过程, 其算法原理有点类似音频指纹中特征匹配的思想。大延时调整的能力是对固定延时调整与线型滤波器能力的补充, 使用它的时候需要比较慎重。需要控制调整的频率, 以及控制造成非因果的风险。

双耦合声学回声消除算法的主要贡献体现在两个方面。移动声学回声介绍

他的是线性回声传递函数。基于这样的数学假设, 我们收到的信号y就可以表示成发射的信号x分别跟这样两个传递函数进行卷积之后的结果。接下来我们对这个模型进行了适当的简化, 简化主要是基于数学分解, 我们假设非线性的传递函数, 可以分解成线性跟非线性这样两个系统函数的组合形式, 就会得到中间的方程。接下来对中间的方程进行变量替换, 就得到这个表达式, 这个表达式它的物理意义很清晰, 我们从可以看到, 整个回声路径是可以表示成线性回声路径跟非线性回声路径二者之和的形式, 这是它的物理意义。

2. 双耦合自适应滤波器, 基于这样一个数学模型, 接下来我们就构建了一种新的滤波器结构, 称之为双耦合自适应滤波器。这个滤波器跟传统线性的自适应滤波器相比, 主要有两个方面的不同, 个不同是传统的线性滤波器只有一个学习单元, 而我们的这个滤波器有两个学习单元, 分别是这里的线性回声路径滤波器, 我们用Wl来表示。还有非线性的回声路径滤波器, 我们用Wn来表示。第二个不同就是, 我们在这两个滤波器之间还加入了一个耦合因子, 这个耦合因子目的就是为了协同二者更好的工作, 让二者能够发挥出比较大的效能, 甚至能够起到 $1+1>2$ 的效果。

移动声学回声介绍介绍非线性声学回声消除的公开文献也少之又少。

喇叭发声单元跟麦克接收单元之间, 通常是需要做隔振处理的, 如果没有隔振处理的话, 那么在喇叭发声的过程中, 他所产生的振动会通过物理方式传递到麦克接收端。对麦克接收到的声学信号进行调制, 而这

种振动本质上是一种随机的、非线性的振动，所以它必然会带来非线性失真。手机声学特性调研我们之前针对市面上主要的手机机型做过一次调研，主要调查声学特性。结果我们很惊讶地发现，市面上超过半数的手机机型，声学特性不够理想，对应这里面的“较差”和“极差”这两档。我们平时用手机开外音玩游戏，或者语音通话时，经常会出现漏回声问题和双讲剪切问题，就与手机声学特性不佳有直接联系。当然这组数据只是针对手机这种电子产品，市面上类似于手机这样的电子产品还有很多，它们应该也有类似的问题。这组数据告诉我们，非线性失真问题在我们生活中的电子产品里是一个普遍存在的问题，我相信对这个问题的研究将会是一个很有价值也很有意义的方向。

23. 避免厅堂音质缺陷的方法主要是从厅堂的体形设计和吸声材料布置两方面入手，消除产生音质缺陷的条件。例如，为了消除回声，应在可能引起回声的部位布置强吸声材料，使反射声减弱；另一种方法是调整反射面角度，将后墙与顶棚交接处作成比较大的倾角，将声音反射给后区观众，彻底消除回声，取得化害为利的效果。为了消除声聚集现象，应尽量控制厅堂界的曲面弧度，采用凸形结构，并在弧面上布置合适的吸音材料。为了消除音质缺陷，可根据厅堂内声源的位置。采用几何作图法，用声线的分布找出各种声缺陷的条件和部位，再采取必要的措施进行抑制。24. 回声指强度和时间差大到足以引起听觉将它与直达声区分开来的反射声。从单一声源产生的一连串可分辨的回声则叫多重回声，当室内两个界面之间距离大于一定数值，且吸声量不足时，在中间声源发出的声音就可能产生多重回声。回声会影响听音注意力，影响声音的清晰度，破坏立体声聆听的声像定位效果。25. 颤动回声当声源在平行界面或一平面与一凹面之间发生反射，界面距离大于一定数值时会出现颤动回声。发生颤动回声时，声音有连续的重叠声，并有颤抖的感觉。颤动回声会引起听力疲劳，使人感到厌烦。

声学回声的功能怎么样？

在线性的回声场景里，双耦合的非线性滤波器是处于休眠的状态，所以它的值是趋于0的，这个时候起主导作用的是线性滤波器。接下来我们再看一下右边的非线性声学回声场景。我们假设非线性的失要出现在 t_1 到 t_2 这个时间段内，大家可以看到黄色线在这个时间里，出现了一次突变，对于NLMS算法，当出现非线性失真之后，它的线性滤波器会去逼近非线性失真。但是由于学习的速度跟不上滤波器变化的速度，所以它跟真实的值之间总是存在一个比较大的gap。同时当非线性失真消失之后，它还需要一段时间恢复到正常状态，因此在整个时间段里，都会出现回声泄露的问题。接下来我们再看双耦合算法，在非线性失真出现之后，线性滤波器会进入到一种相对休眠的状态，就是前面所提到的耦合机制，会降低它的更新速度，所以在整个非线性出现的这段时间里，他的值是缓慢变化的。进入非线性失真状态之后，非线性滤波器开始工作，它会快速非线性特性的变化，而当非线性失真消失之后，非线性滤波器又进入休眠状态。将这两个滤波器结合起来，就可以实现对整个声学回声路径的变化进行有效。这里只是给出了一个示例，实际情况往往要复杂很多。接下来我们对这2个滤波器做了特性比较，主要是从4个不同的维度。

我们把声学回声消除这个技术变成一张实体的插件（设备插卡），在系统中，为实现次回声过滤。移动声学回声介绍

回声消除AEC[AcousticEchoCancellation]一般指的是声学回声消除，其主要用于抑制产品本身发出的

声音。移动声学回声介绍

噪声抑制和声源分离同属于语音增强的范畴，如果把噪声理解为广义的噪声三者之间的关系, 噪声抑制需要准确估计出噪声信号，其中平稳噪声可以通过语音检测判别有话端与无话端的状态来动态更新噪声信号，进而参与降噪，常用的手段是基于谱减法(即在原始信号的基础上减去估计出来的噪声所占的成分)的一系列改进方法，其效果依赖于对噪声信号估计的准确性。对于非平稳噪声，目前用的较多的就是基于递归神经网络的深度学习方法，很多Windows设备上都内置了基于多麦克风阵列的降噪的算法。效果上，为了保证音质，噪声抑制允许噪声残留，只要比原始信号信噪比高，噪且听觉上失真无感知即可。单声道的声源分离技术起源于传说中的鸡尾酒会效应，是指人的一种听力选择能力，在这种情况下，注意力集中在某一个人的谈话之中而忽略背景中其他的对话或噪音。该效应揭示了人类听觉系统中令人惊奇的能力，即我们可以在噪声中谈话。科学家们一直在致力于用技术手段从单声道录音中分离出各种成分，一直以来的难点，随着机器学习技术的应用，使得该技术慢慢变成了可能，但是较高的计算复杂度等原因，距离RTC这种低延时系统中的商用还是有一些距离。噪声抑制与声源分离都是单源输入。

移动声学回声介绍