



为开发者赋能 为行业加速

2018年5月19日-20日 · 上海光大会展中心

主办方：LiveVideoStack
—— 音视频技术社区 ——



触宝电话如何改善通话质量

吴威麒

触宝科技 音频技术专家





技术开启新“视”界

Technology Bring New Vision

2018.10.19-20 北京丽亭华苑酒店

LiveVideoStackCon 2018正在招募出品人、讲师

自荐或推荐：speaker@livevideostack.com

大会购票通道



LiveVideoStack
— 音视频技术社区 —

CSDN

内容摘要

- 01** 声音问题
- 02** 回声消除
- 03** 语音降噪
- 04** 音量增益控制
- 05** 啸叫抑制
- 06** 变速变调
- 07** 总结与展望



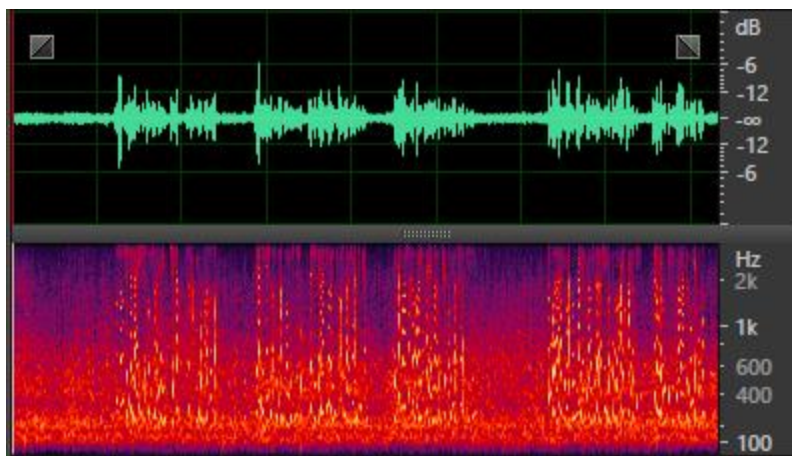
免费电话

第一款真正的免费电话

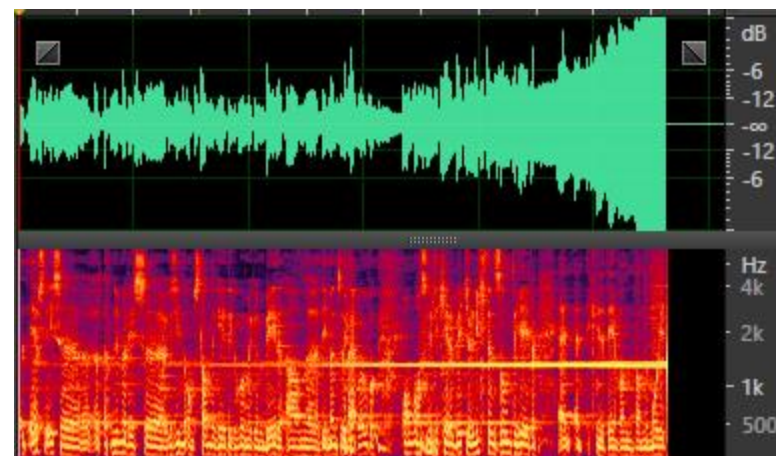
用触宝打电话，永久免费。无需注册，安装后即可在WIFI和3/4G网络下免费打电话。对方无需安装，无需联网，随时随地免费畅聊。

全球2亿用户共同的选择

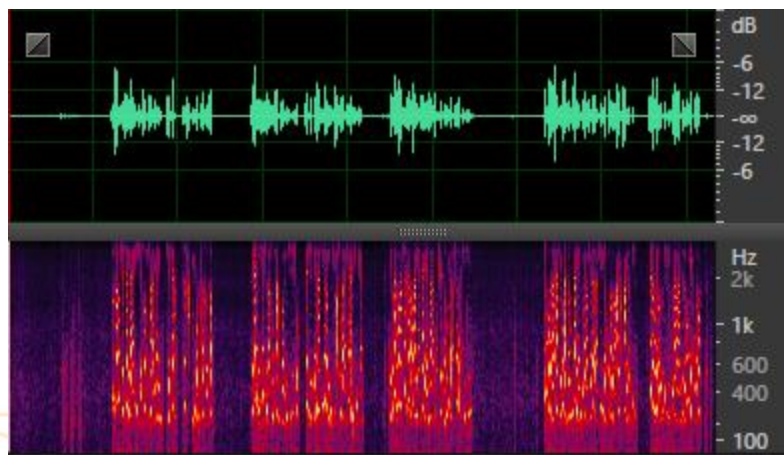
声音问题



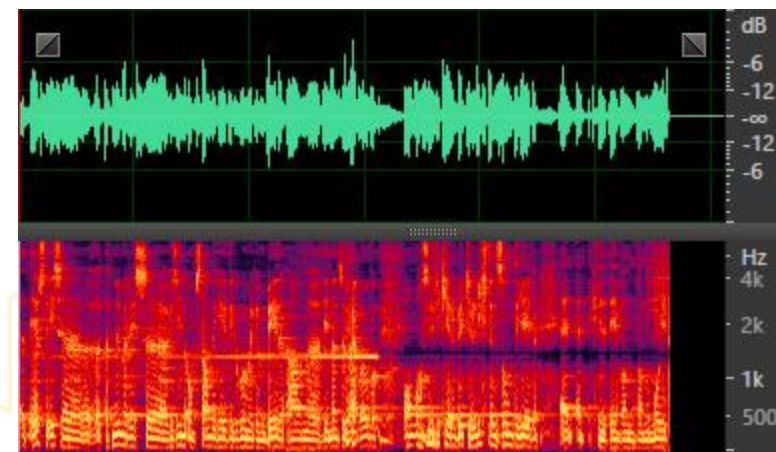
含噪语音



单频啸叫



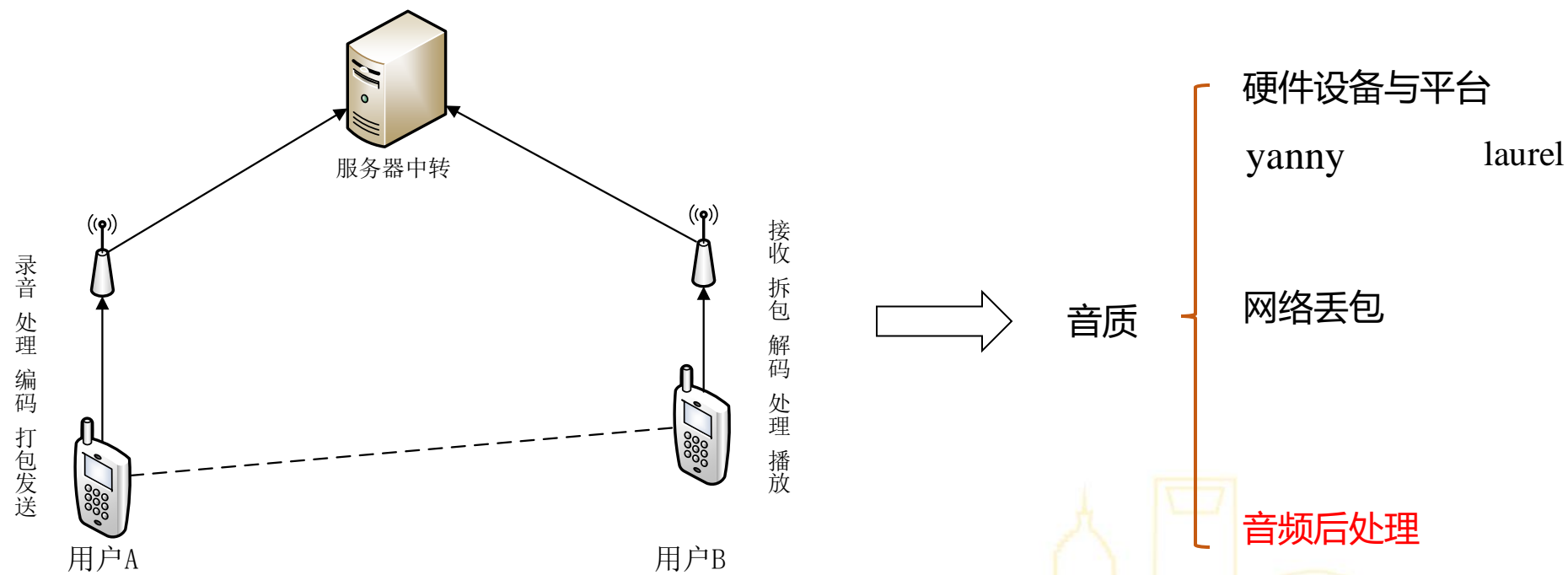
去噪语音



啸叫抑制



音质分析



1 ----• 回声消除

2 ----• 语音去噪

3 ----• 音量控制

4 ----• 啸叫抑制

5 ----• 编解码

6 ----• 端点检测

7 ----• 混音

8 ----• 去混响

9 ----• 音效

回声消除

回声的产生

基本框架

线性回声及建模

非线性回声及处理

近端干扰

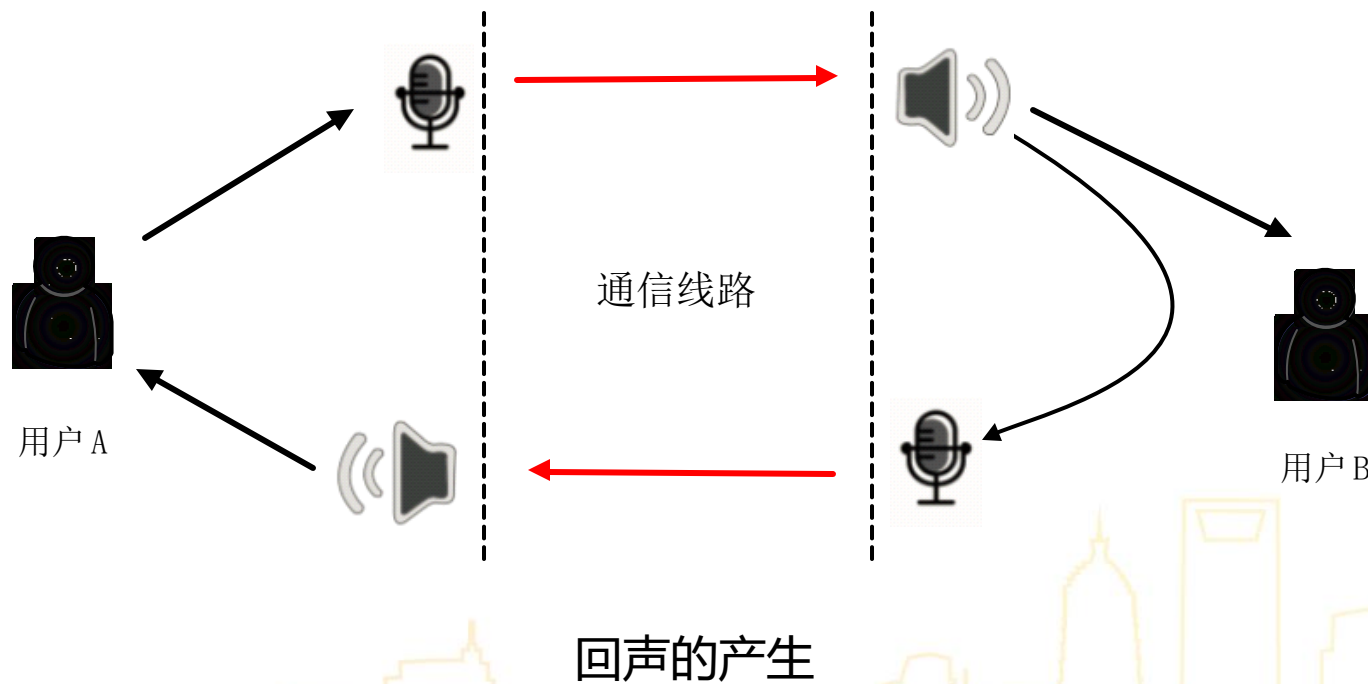
信号对齐

计算量

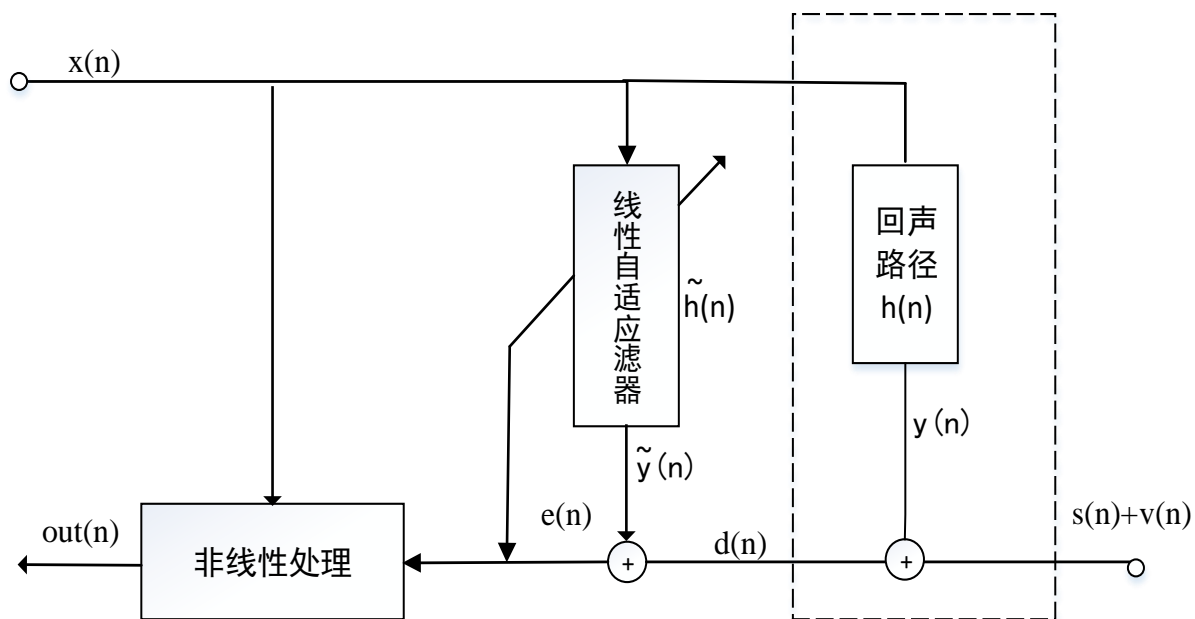
PBFDFAF

效果对比

回声的产生



基本框架



$x(n)$ 远端参考语音
 $h(n)$ 回声路径
 $y(n)$ 回声信号
 $y(n)=x(n)*h(n)$

$d(n)$ 近端麦克风录制信号
 $d(n)=y(n)+s(n)+v(n)$
 $s(n)$ 近端语音
 $v(n)$ 近端噪声



LMS线性滤波器



$$\mathbf{x}^* \tilde{\mathbf{h}} \rightarrow d(n) = y(n)$$

均方误差意义最小:

$$e(n) = y(n) - \mathbf{x}^* \tilde{\mathbf{h}}$$

$$E[e(n)^2] = 0$$

$$E[y(n)^2] - 2\tilde{\mathbf{h}}^T E(\mathbf{x} * y(n)) + \tilde{\mathbf{h}}^T E(\mathbf{x}\mathbf{x}^T) \tilde{\mathbf{h}} = 0$$

最优滤波器系数

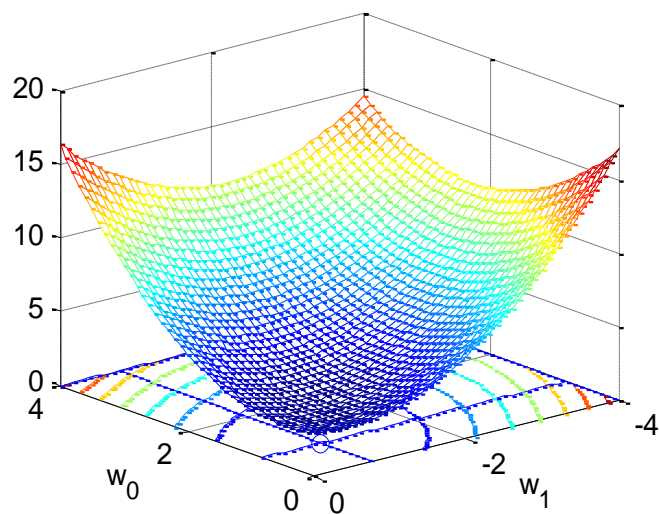
$$\mathbf{h}^* = R_{xx}^{-1} E(\mathbf{x} * y(n))$$

梯度下降法

$$\mathbf{h}(n+1) = \alpha * \mathbf{h}(n) + 2\mu * \mathbf{x} * e(n)$$

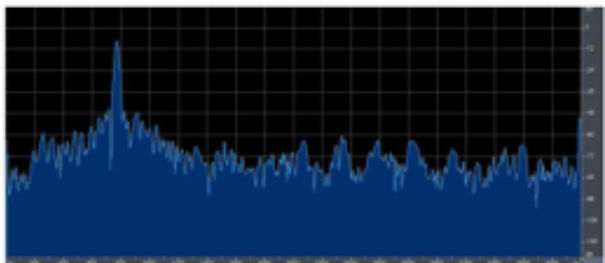


学习率

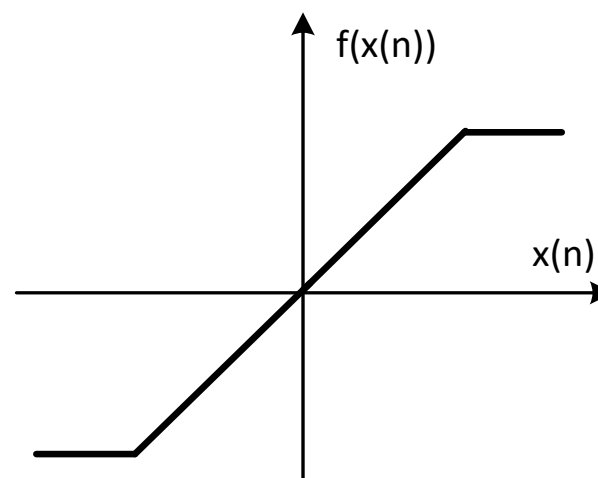


何时步长调大？
何时步长调小？

非线性回声



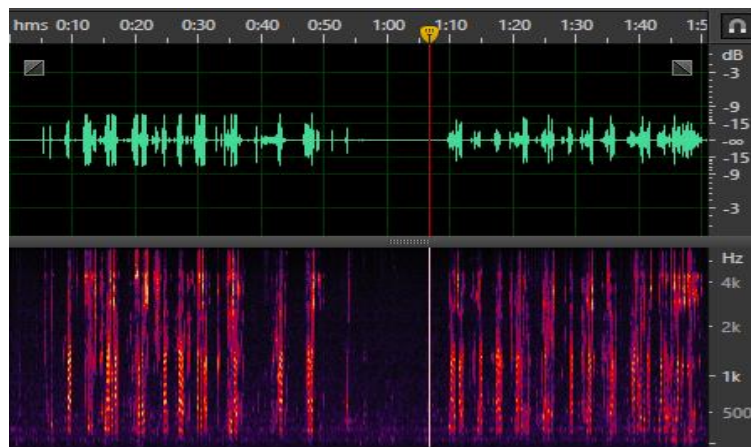
谐波失真



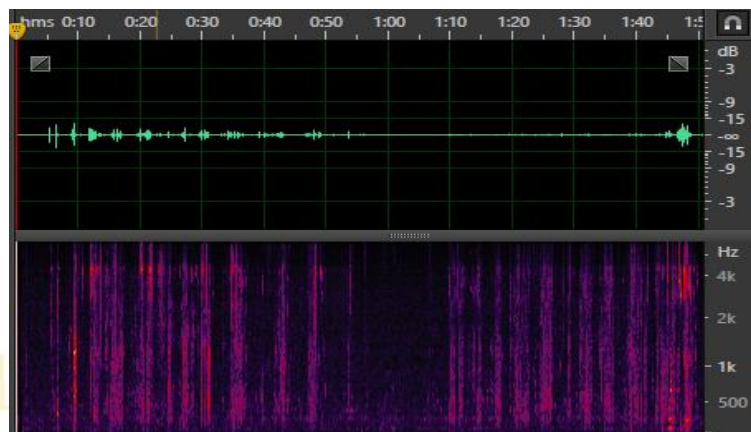
幅度失真

非线性处理

中心削波



相关性抑制



能量法

互相关

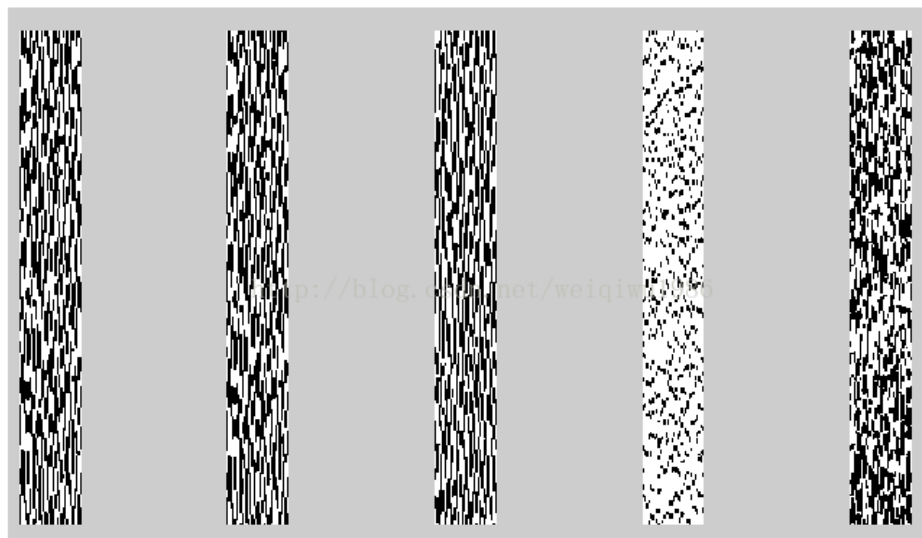
双滤波器

分类识别

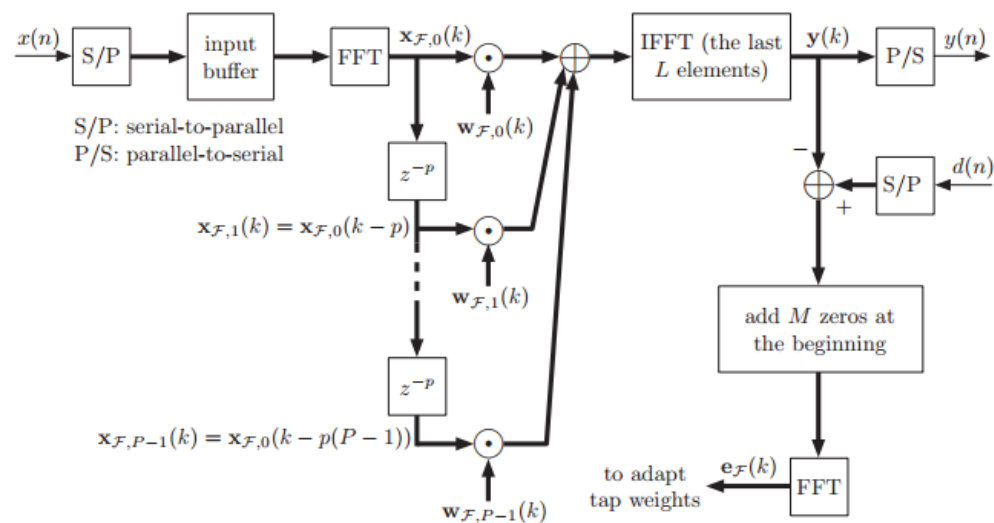
互相关

共振峰匹配

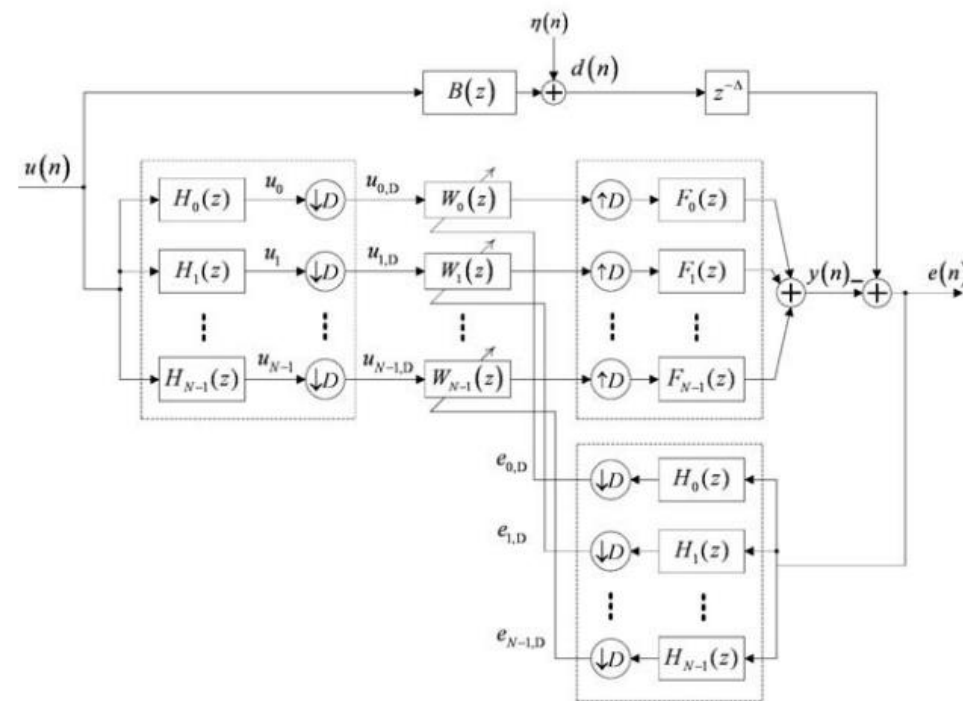
二值谱匹配



指纹匹配

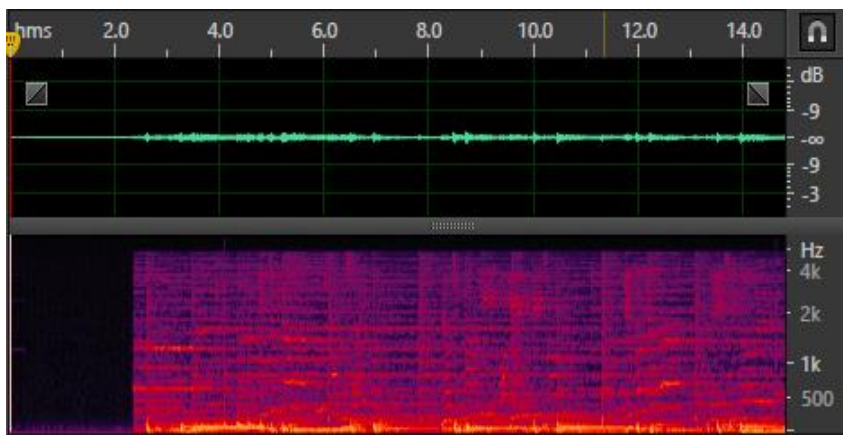


PBFDAF结构图

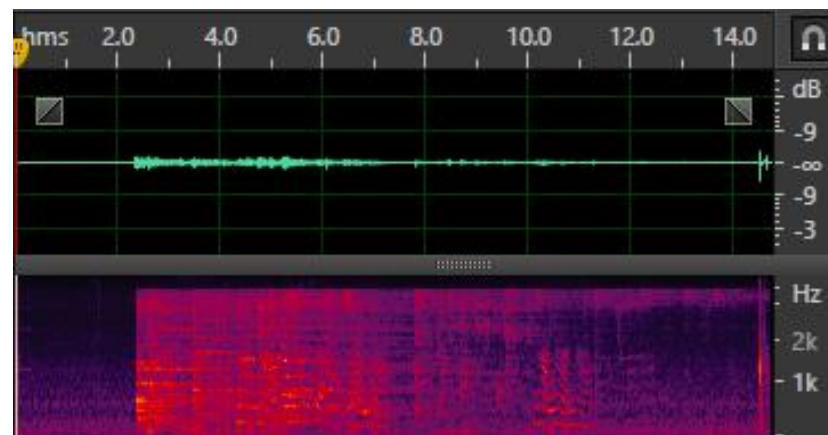


闭环型子带滤波器结构图

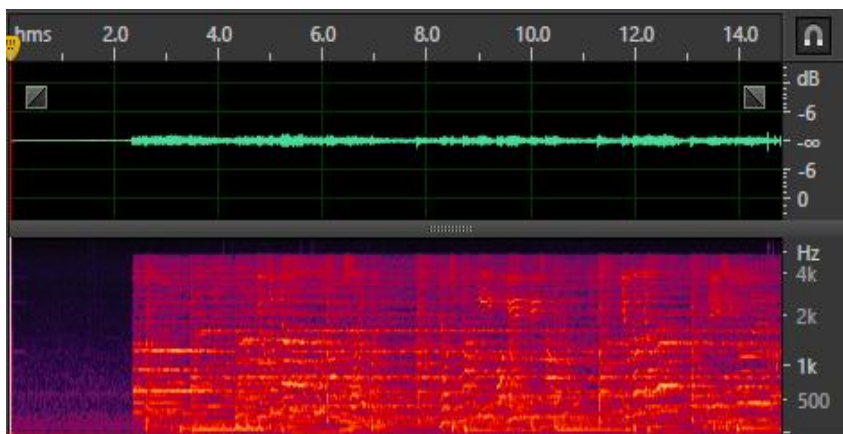
效果对比



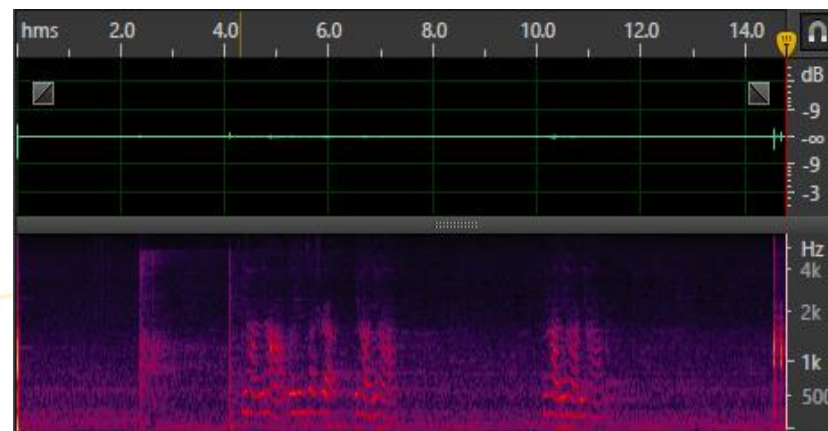
ref



webrtc



mic



ours

去噪

噪声类型与估计

维纳滤波

webrtc的实现方法

语音质量客观评估

噪声类型与估计

高斯白噪声

粉红噪声

风扇噪声

电流声

敲键盘声

写字沙沙声

空调噪声

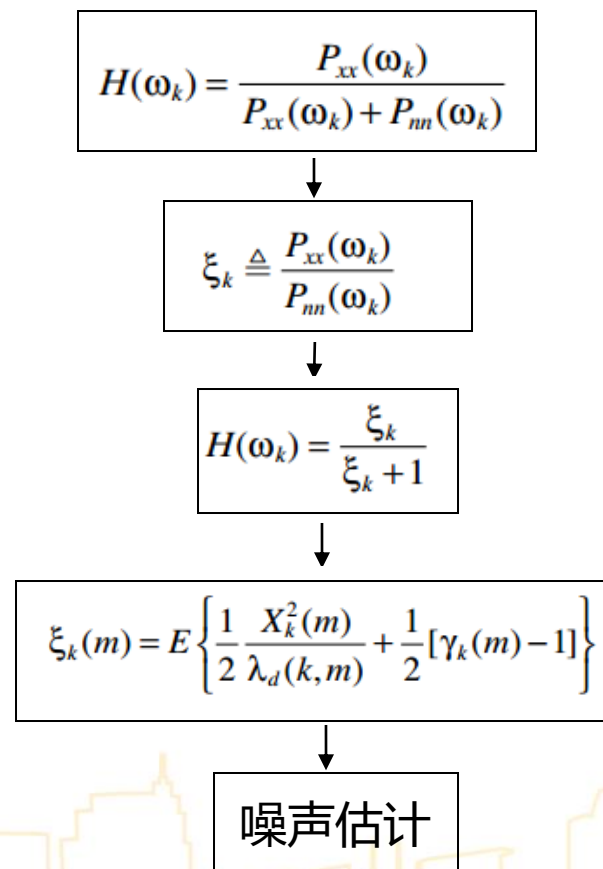
车内噪声

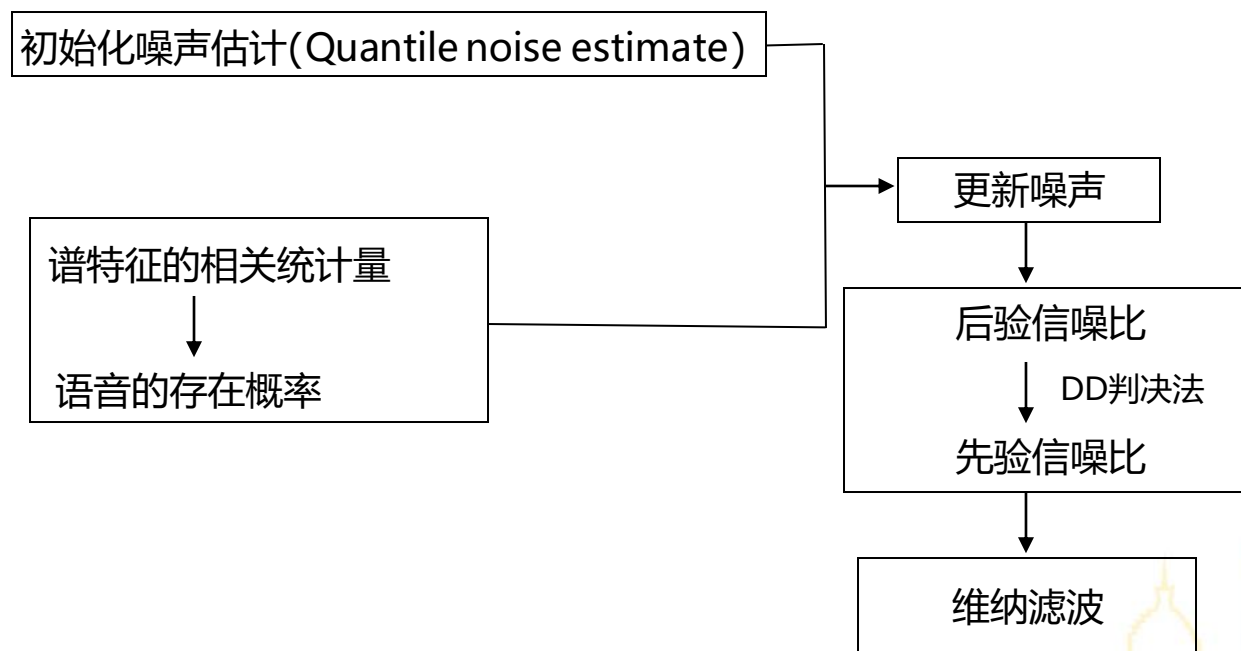
地铁噪声

机场噪声

餐厅内嘈杂噪声

最小值跟踪估计





语音质量主观评估

MOS	主观意见	听觉感受
4-5	优秀	很好
3-4	良好	清楚
2-3	一般	不太清
1-2	差	勉强
0-1	很差	极差

语音质量客观评估

- 1 SNR
- 2 PESQ
- 3 LLR
- 4 LSD

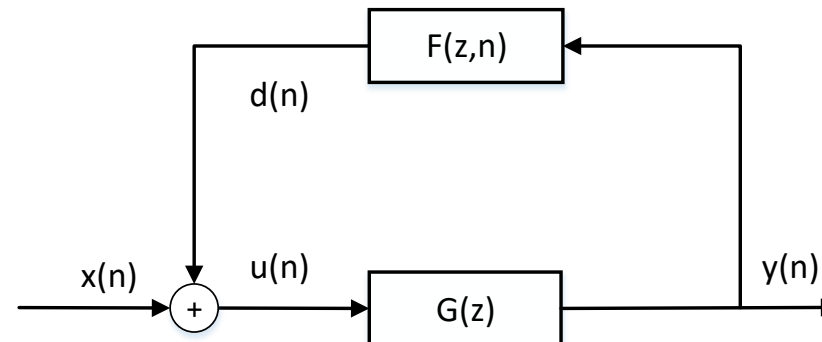
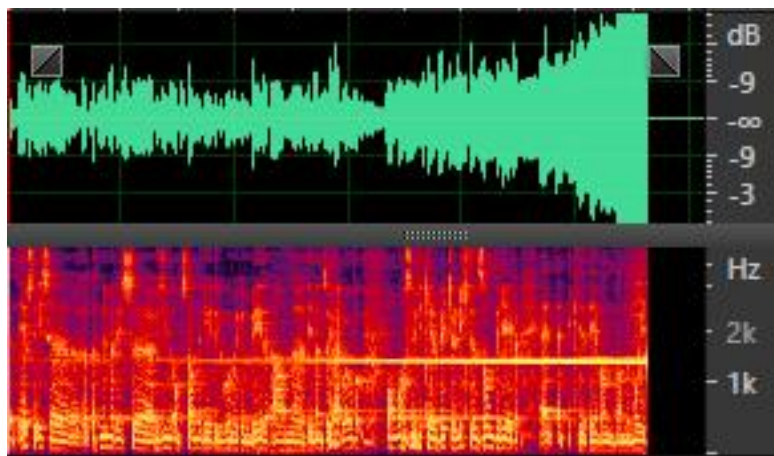
三种工作模式：

kAdaptiveAnalog

kAdaptiveDigital

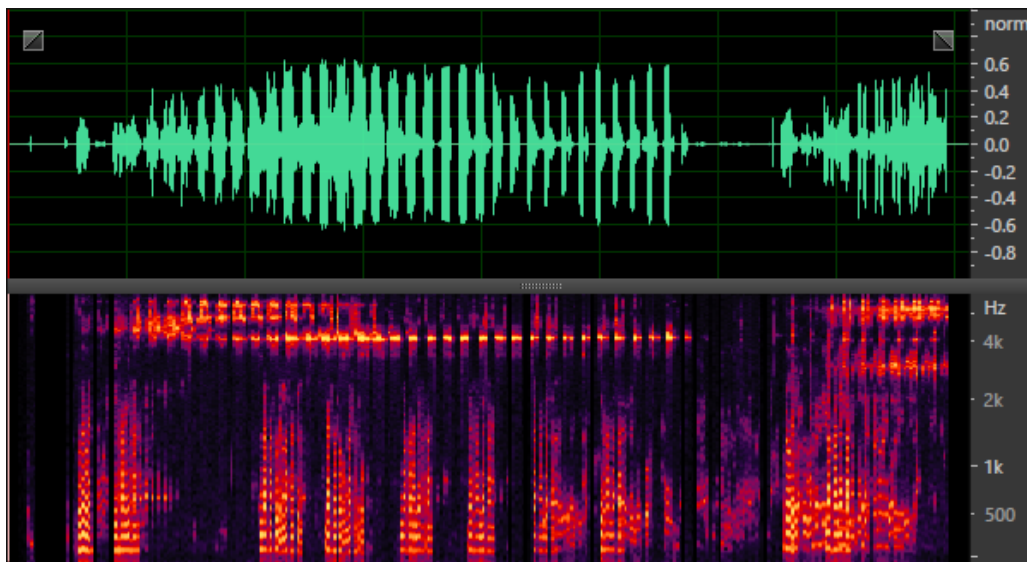
kFixedDigital

啸叫抑制

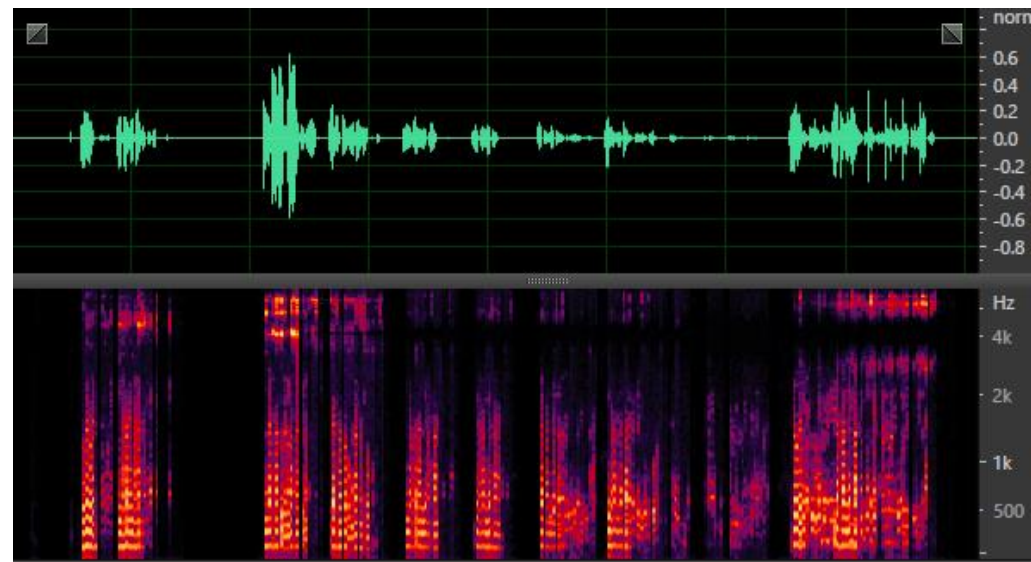


效果

频点不固定？
宽带频点？
多频点？
频点断续？



啸叫语音



抑制后的语音

时域法	参量法
剪贴法	相位声码器
SOLA、SOLA-FS、WSOLA	正弦模型

变速不变调

时域法	参量法
频率调制	相位声码器
变速不变调+重采样	正弦模型

变调不变速

- 应用场景

- 缓冲区包过多
- 抗丢包
- 预防啸叫
- Tomy猫
- 男声<->女声
- 老人<->小孩
- 老鼠吱吱声

AEC

- ◆ 高采样率
- ◆ 音乐信号
- ◆ 双声道立体声AEC
- ◆ 语音识别前端

ANS

- ◆ DNN降噪

AGC

- ◆ 人声检测

Thank You !



主办方：LiveVideoStack

