音频信息隐藏与水印算法

钮心忻、杨榆、雷敏
北京邮电大学信息安全中心
yangyu@bupt.edu.cn

音频信息隐藏技术

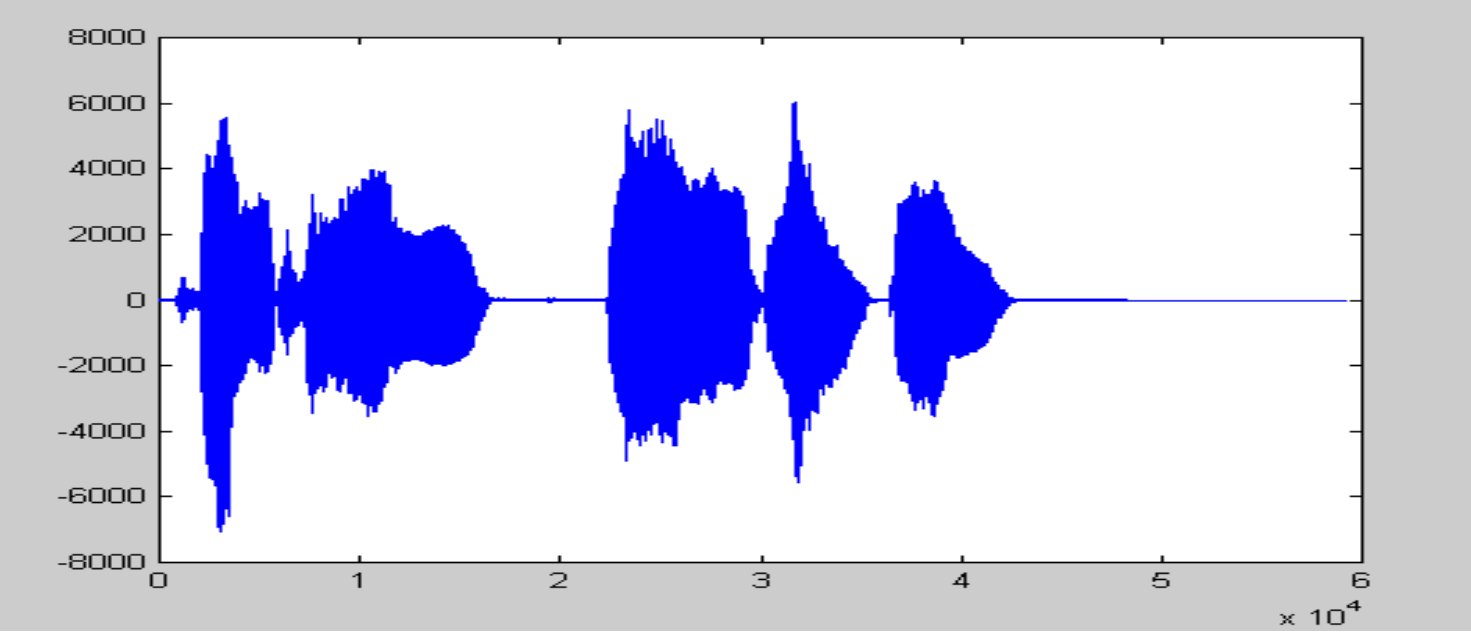
- 音频信号的特点
  - 一维信号
  - 人耳听觉系统（HAS）比人眼视觉系统（HVS）灵敏得多
- 对音频信息隐藏技术的要求
  - 透明性
  - 鲁棒性（强鲁棒，抗模数转换）
  - 同步要求
  - 盲检测

时间域音频算法

- 最低有效位方法
  - LSB
  - Lease Significant Bit
- 回声隐藏法

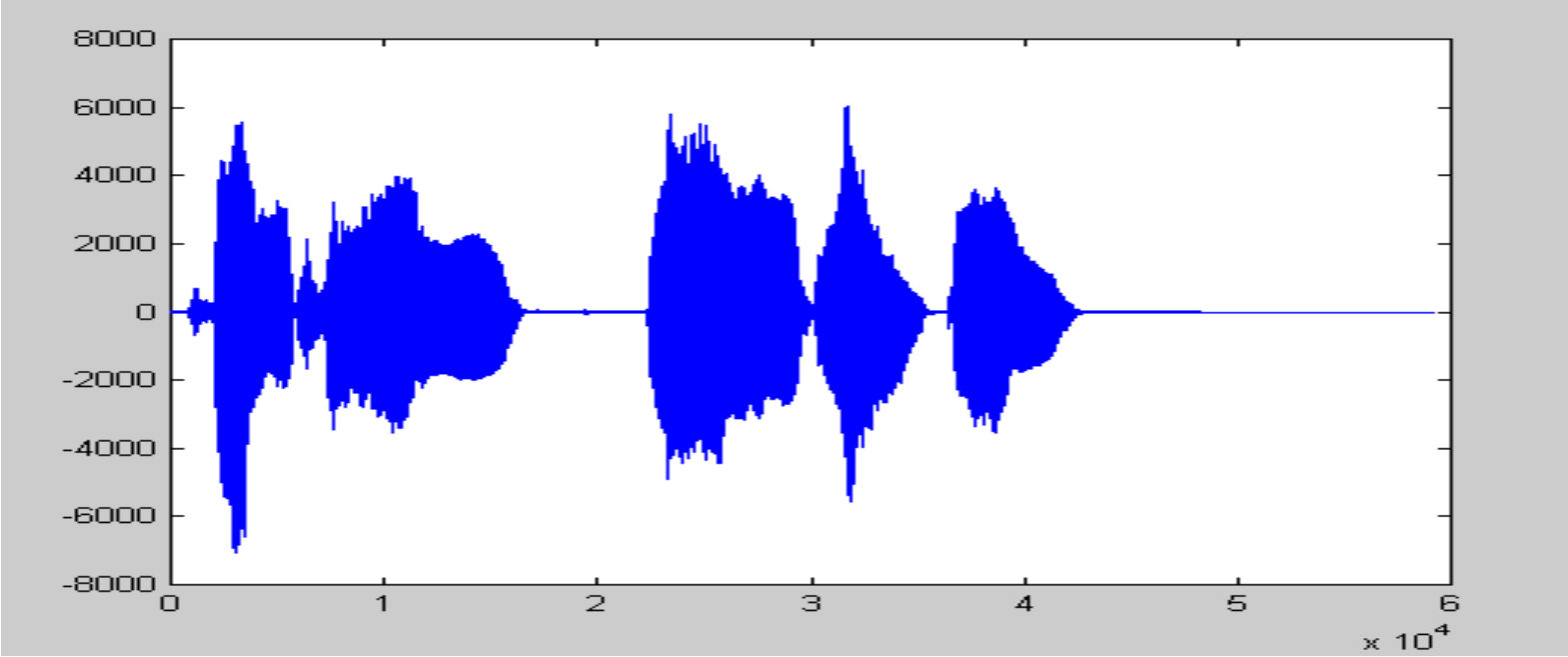
LSB原理

- 原始语音信号（“床前明月光”）



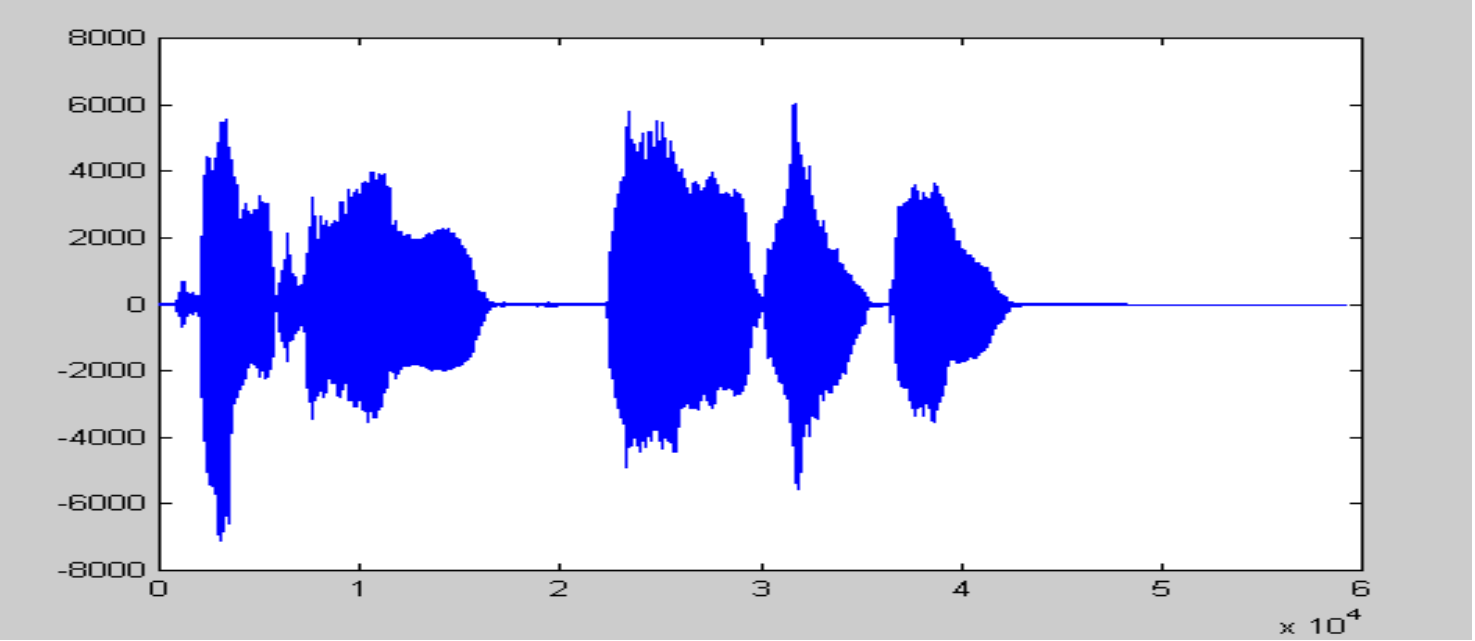
LSB原理（1）

- 去掉低2比特位的语音信号（声音信号听不出差别）



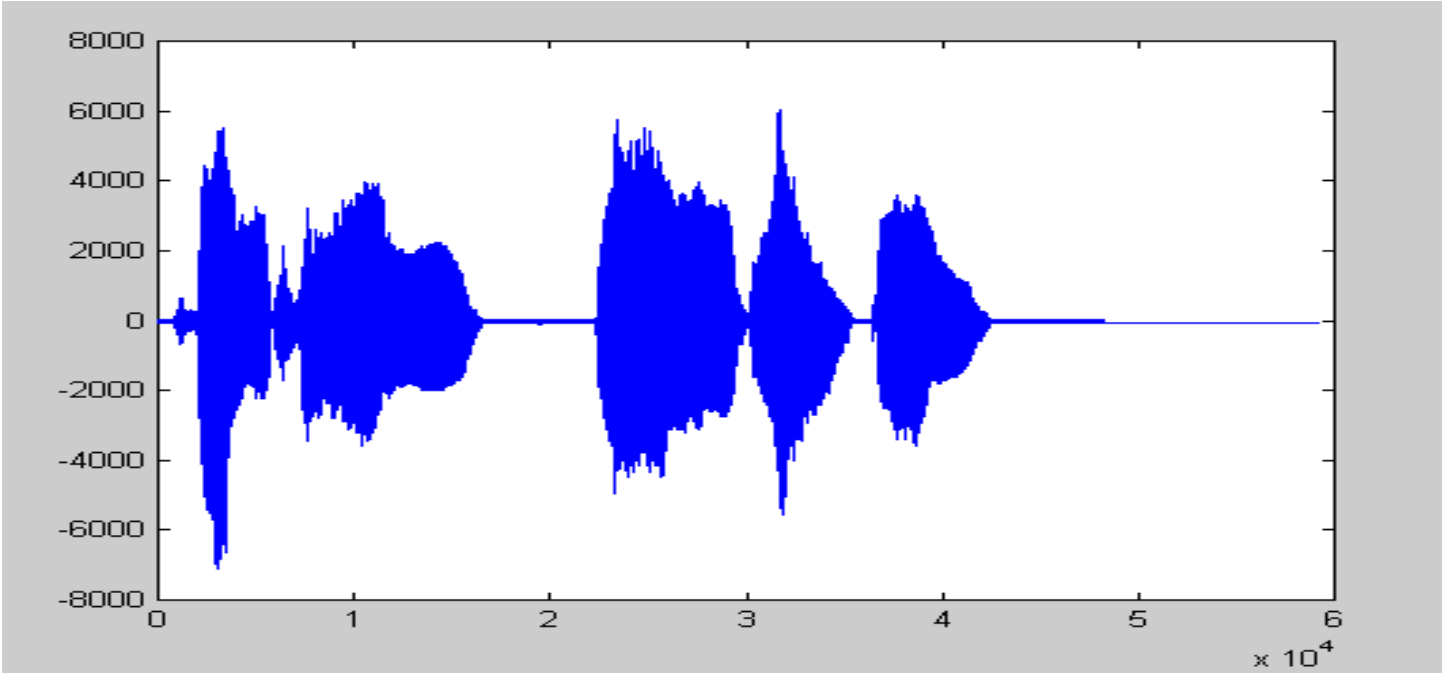
LSB原理（2）

- 去掉低4比特位的语音信号（声音信号听不出差别）



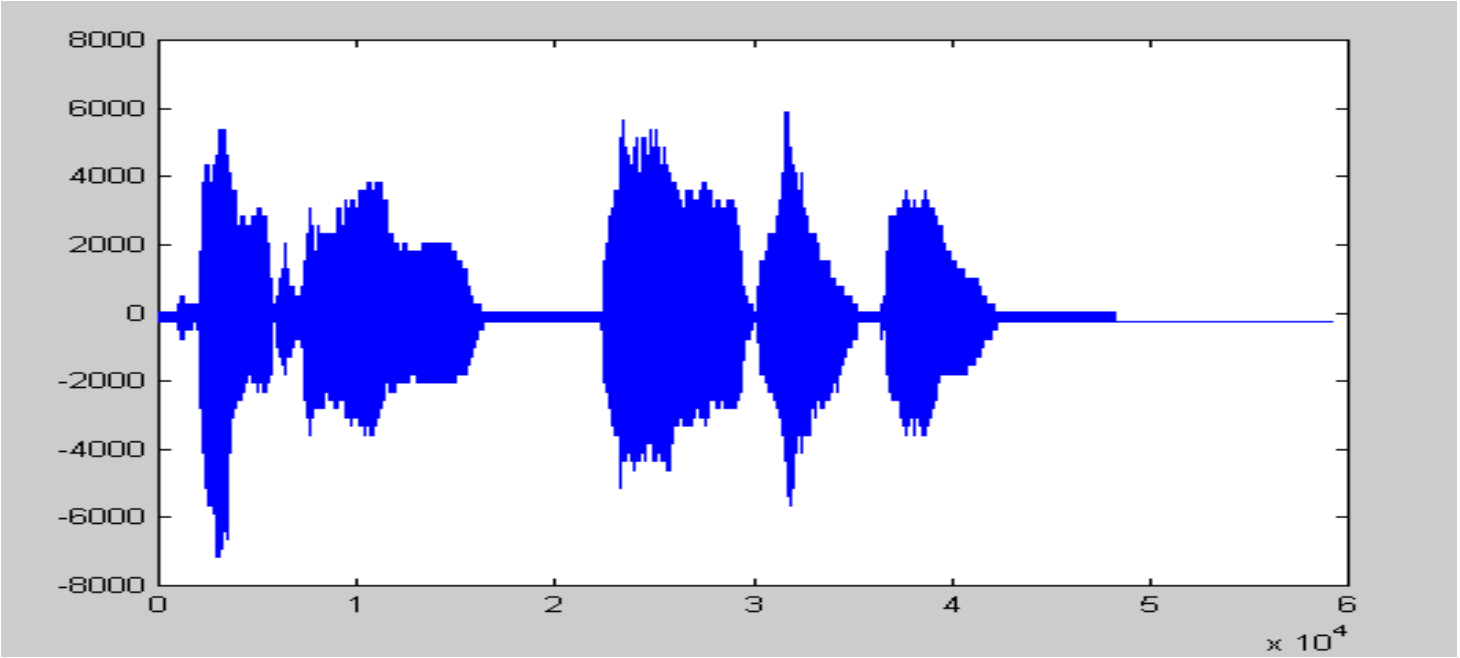
LSB原理（3）

- 去掉低6比特位的语音信号（声音中有极少的背景噪音，不易被察觉）



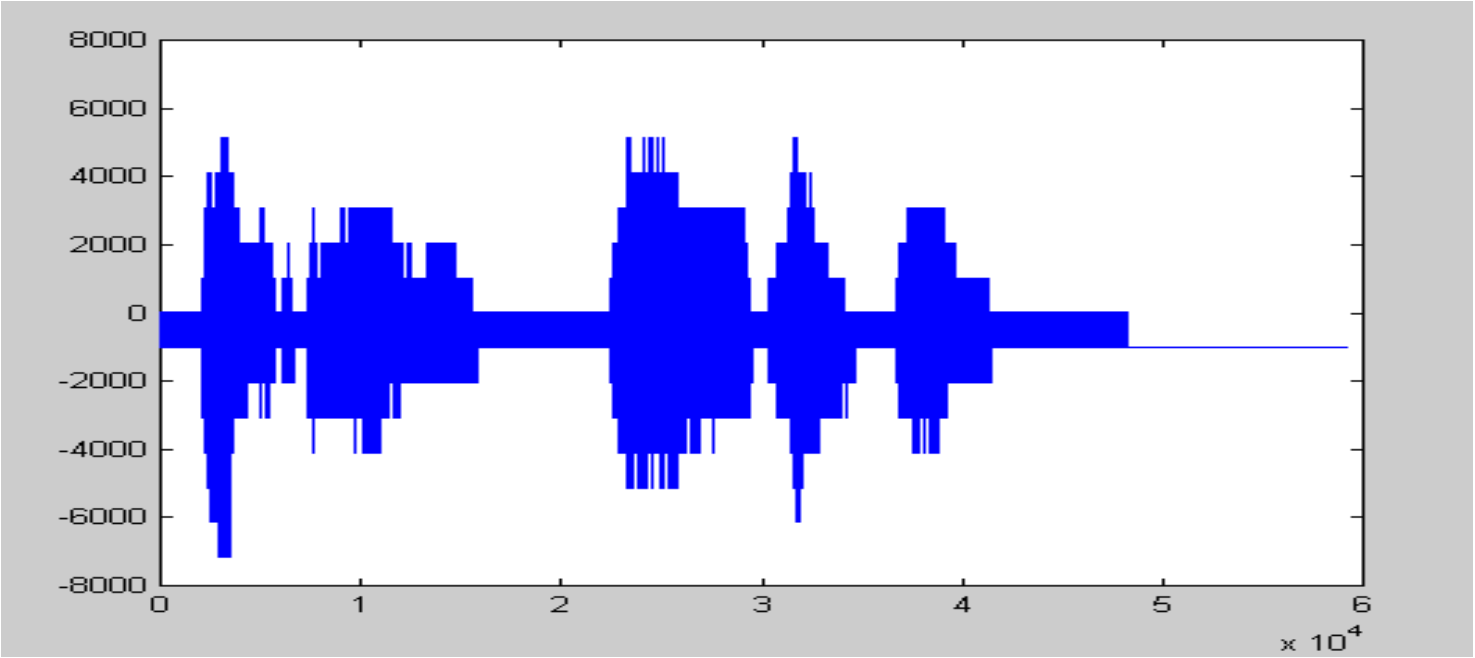
LSB原理（4）

- 去掉低8比特位的语音信号（声音中有较明显的背景噪音）



LSB原理（5）

- 去掉低10比特位的语音信号（声音中有很强的噪音，但话音仍较清晰）



LSB原理（5）

- 结论
  - 数字化音频中，低有效比特对音质贡献弱。
  - 改变低有效比特不会显著影响音质。

LSB算法

- LSB算法实现
  - 嵌入：用水印替换最低（或次低等）有效比特

0110 0011 0101 0111 0111 0110

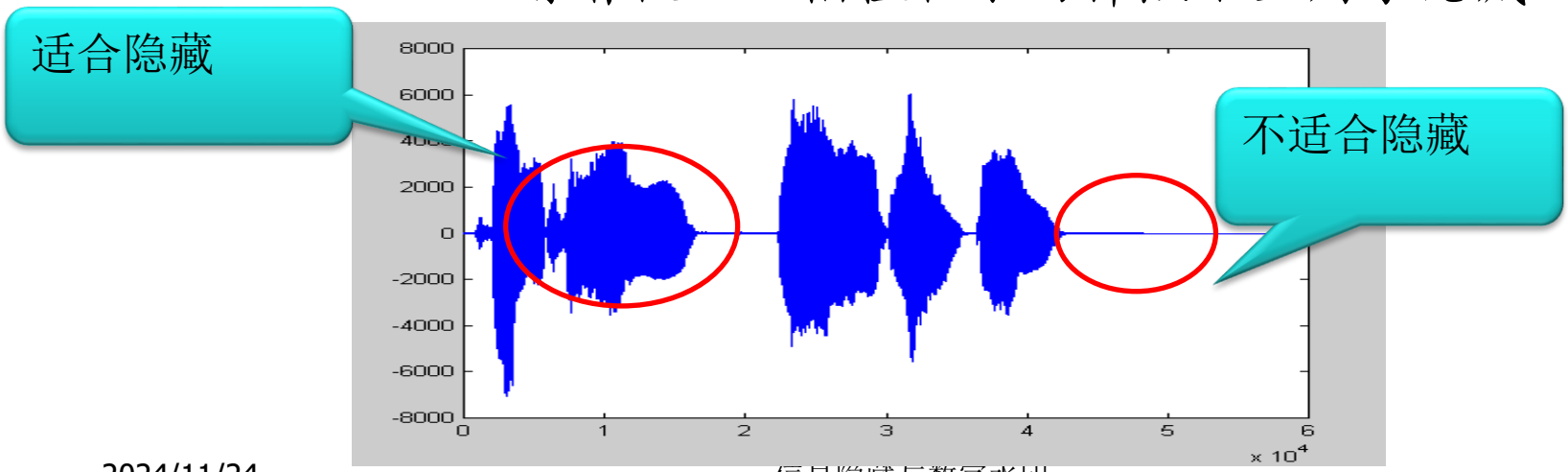
0 0 1

0110 0010 0101 0110 0111 0111

  - 提取：提取最低（或次低等）有效比特组合为水印。

LSB算法

- LSB算法设计
  - 选择样点
    - 样点幅值大小与其掩蔽能力有关
    - 静音段——幅值较小的样点不宜用于隐藏。



LSB算法

- LSB算法设计
  - 选择比特位
    - 低比特位对音质影响小，但容易受到干扰。
    - 例如：幅值为6(110B)的样点，哪怕幅度仅变化1，其多个比特位也会发生变化，
    - 若幅值减小1，变为5（101B），则最低、次低有效比特位变化；

LSB算法

- LSB算法设计
  - 选择比特位
    - 若幅值增大1，变为7（111B），则最低有效比特位发生变化。
    - 若在次低或第3比特位隐藏水印，则不容易受噪声干扰，但嵌入前后样点幅值的变化幅度由1上升到2或4。

LSB小结

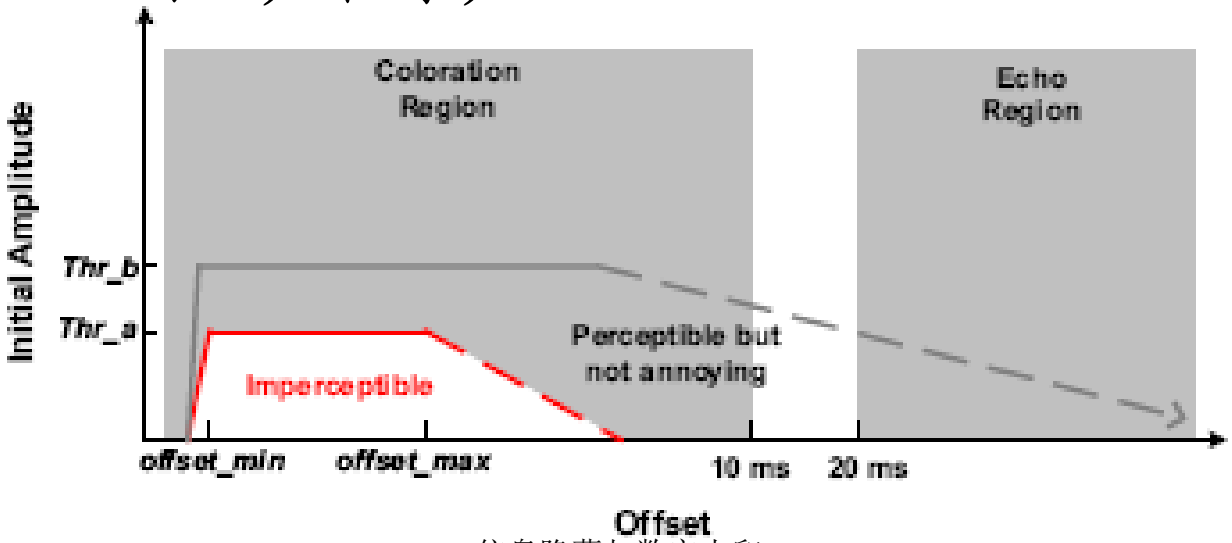
- LSB算法参数包括:
  - 样点和比特位置的选取
- LSB算法性能:
  - 透明度高
  - 容量大
  - 鲁棒性差

LSB算法

- 思考题
  - 某音频采用8比特无符号数量化，已知该音频使用了LSB算法嵌入了秘密信息，且部份样点值为127，125，110，则可从中提取的秘密信息是?
  - 我们学习生活中传送音频的方式有哪些？LSB适用于哪些应用?

回声隐藏

- 原理
  - 掩蔽效应：强信号的存在会使其附近的弱信号难以被感知。
  - 当回声与原声的间隔充分接近时，人耳难以区别回声和原声。



回声隐藏

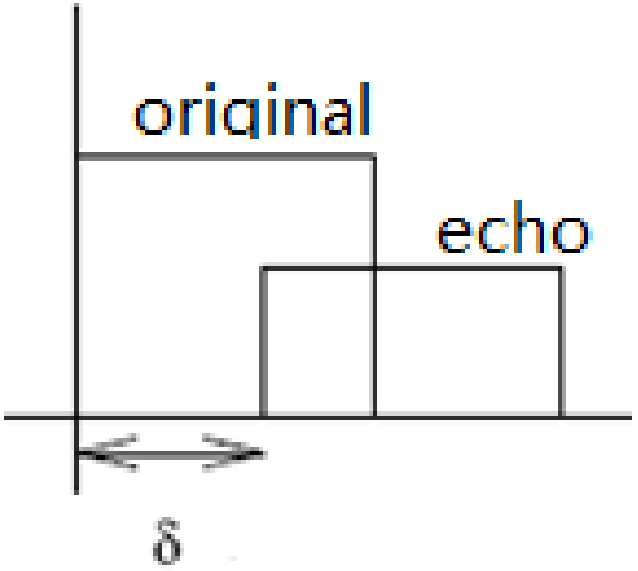
- 如何应用掩蔽效应隐藏秘密信息?
  - 回声和原声间的延迟在一定范围内人耳都难以察觉，
  - 亦即可以人为添加不同延迟的回声。

回声隐藏

- 如何应用掩蔽效应隐藏秘密信息?
  - 要隐藏二进制信息，只需用两种不同延迟的回声分别代表0、1比特。
  - 例如，回声延迟为1毫秒代表比特“1”，回声延迟为2毫秒代表比特“0”，这样，要隐藏0，那么我们在原声上添加延迟为2毫秒的回声。

回声隐藏

- 如何“生成”回声信号?
  - 回声信号，可简单模拟为，原始信号经过时延和幅度衰减后产生的信号。
  - 设原信号为x(t)，时延为δ，衰减为α，
  - 则叠加回声的信号为:
  - $y(t)= x(t) + \alpha x(t- \delta)$



回声隐藏

- 案例：“回声”的制造
  - 已知音频片段采样值为以下序列：
  - $x[i]$:10, 12, 14, 8, 6, 8
  - 请产生衰减系数为0.5，延迟为2个采样间隔的回声。
  - 请将上述回声叠加到原声，生成混合序列。

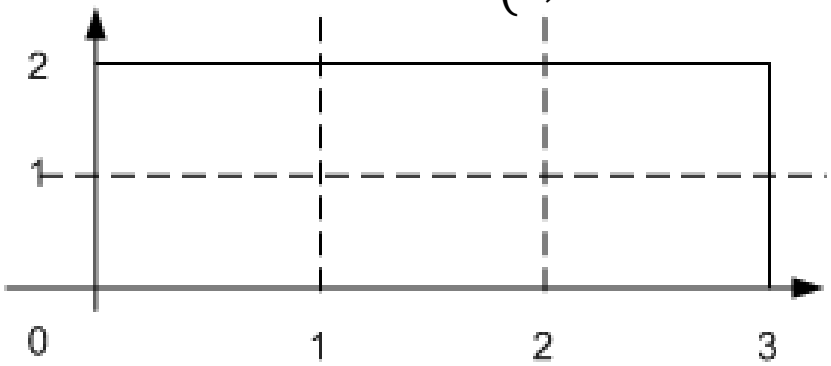
回声隐藏

- 案例：“回声”的制造
  - 解：衰减系数为0.5，则序列变为：
  - $0.5x[i]$:5, 6, 7 4, 3, 4
  - 延迟为2个采样间隔，因此回声为：
  - $0.5x[i-2]$:0, 0, 5, 6, 7 4, 3, 4
  - 混合序列为：
  - $y[i]=x[i]+0.5x[i-2]$:10, 12, 19, 14, 13 12, 3, 4

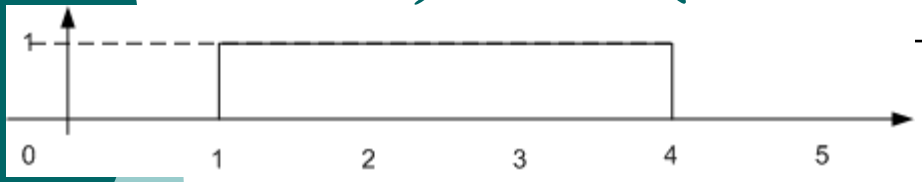
回声隐藏

- 例：若回声延迟为1毫秒代表比特“1”，回声延迟为2毫秒代表比特“0”，回声幅度衰减系数为0.5，请给出下面信号对应的0、1回声信号，以及在这个信号上嵌入比特“1”以后所得信号

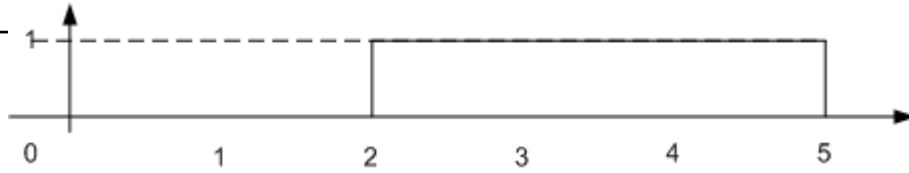
$$f(x)=\begin{cases}2, & 0\le x\le 3\\ 0, & x<0\ or\ x>3\end{cases}$$



回声隐藏



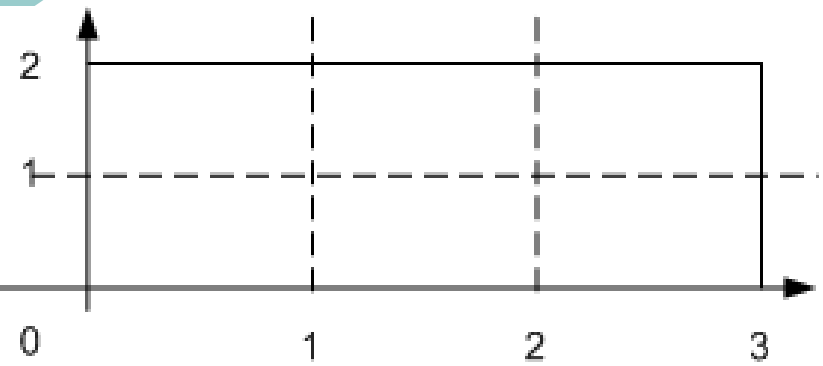
标识比特“1”的回声信号



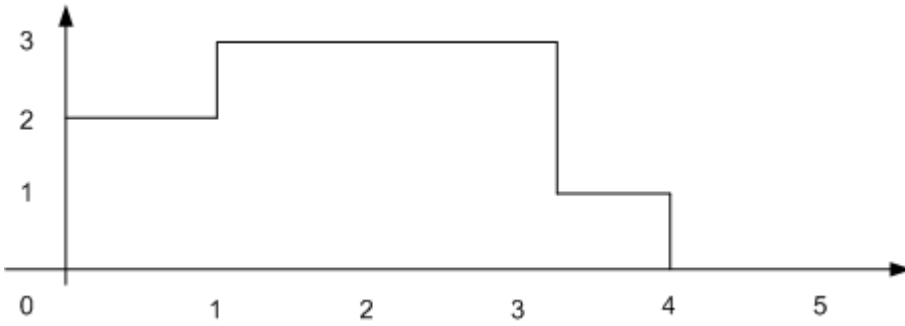
标识比特“0”的回声信号

$$g_1(x) = 0.5f(x-1) = \begin{cases} 1, & 1 \le x \le 4 \\ 0, & x < 1 \ or \ x > 4 \end{cases}$$

$$g_0(x) = 0.5f(x-2) = \begin{cases} 1, & 2 \le x \le 5 \\ 0, & x < 2 \ or \ x > 5 \end{cases}$$



原始信号



叠加回声“1”后的合成信号

$$f(x) = \begin{cases} 2, & 0 \le x \le 3 \\ 0, & x < 0 \ or \ x > 3 \end{cases}$$

$$y(x) = f(x) + g_1(x) = \begin{cases} 2, & 0 \le x \le 1 \\ 3, & 1 < x \le 3 \\ 1, & 3 < x \le 4 \\ 0, & x < 0 \ or \ x \ge 4 \end{cases}$$

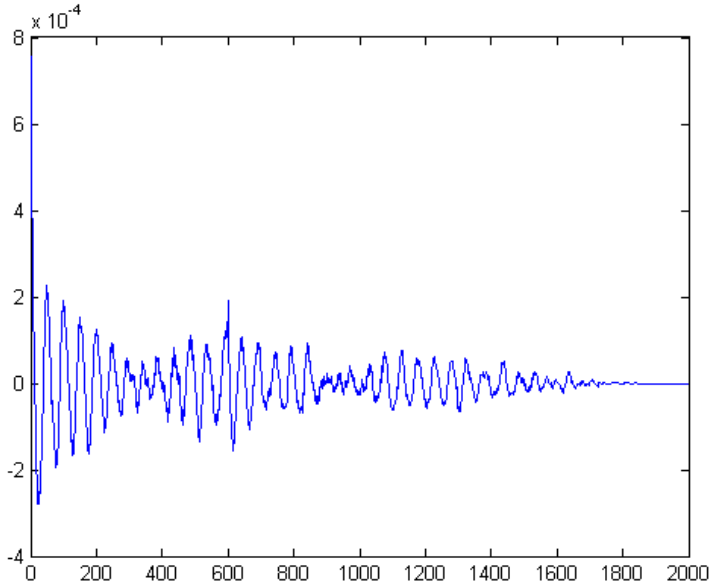
回声隐藏

- 如何提取水印？
  - 水印信息，即0、1比特，与回声延迟相关联，
  - 提取信息最自然的思路就是检测回声延迟。
  - 回声是原始信号平移、线性缩放后生成的信号，与原信号相关性较强，
  - 可否使用自相关系数检测回声延迟?
  - 自相关（Autocorrelation）是信号与其自身平移信号的互相关： $R(t) = \int_{-\infty}^{+\infty} f(\tau) f(t-\tau) d\tau$
  - 常用语分析信号自身的重复模式（如周期信号，谐波等）。

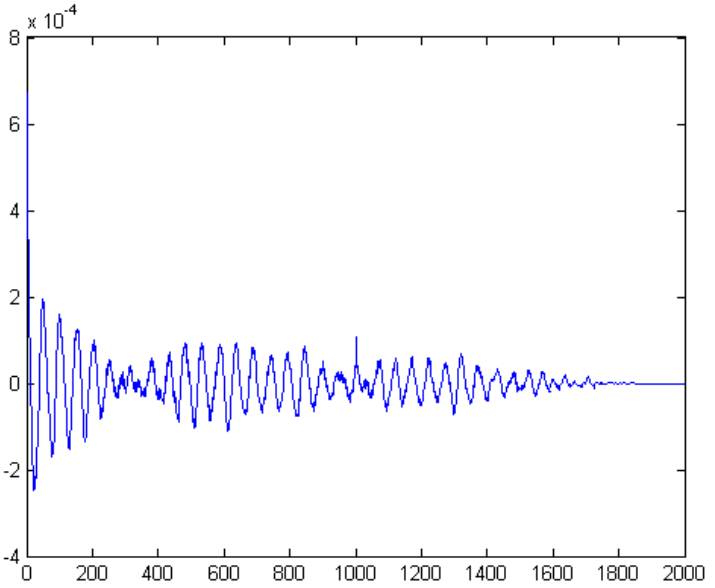
回声隐藏

图示语音信号采样率为8000Hz，即每(1/8000)秒，或每(1/8)毫秒产生一个采样点。回声信号比原声信号延迟N个样点，即延迟N*(1/8)毫秒。因此，若N等于600，则延迟为75毫秒。若N等于1000，则延迟为，125毫秒。

自相关（回声延迟为600个样点）



自相关（回声延迟为1000个样点）



语音信号自身具有相关性，当回声延迟足够大时，能通过自相关系数检测延迟，但这样条件下的透明性和容量性能不佳。

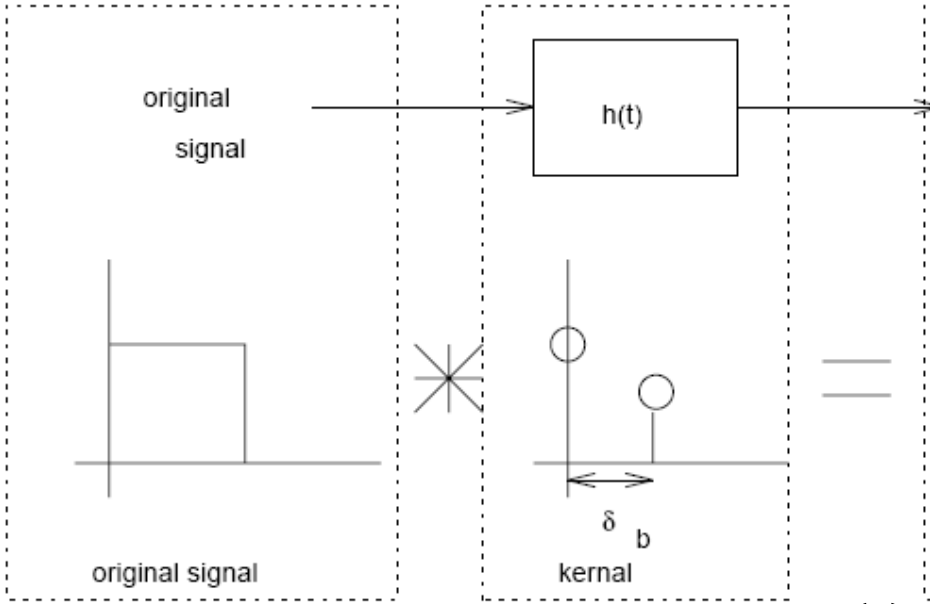
略过深入讨论

回声隐藏（深

性质名称	函数的卷积积分
交换律	$x(t) * h(t) = h(t) * x(t)$
分配律	$x(t) * [g(t) + h(t)] = x(t) * g(t) + x(t) * h(t)$

性质名称	函数的卷积积分
	$x(t) * \delta(t) = x(t)$
延时特性	$x(t) * \delta(t - t_0) = x(t - t_0)$

- 如何提取水印？
  - 研究水印嵌入过程，可建模为线性系统



$x(t) * \delta(t) = \int x(t-\tau)\delta(\tau)d\tau = x(t)$

original

echo

$x(t) * \alpha\delta(t - \Delta)$

$= \alpha x(t - \Delta)$

$\left[\delta(t) = \left\{ \begin{matrix} \infty & (t=0) \\ 0 & (t \neq 0) \end{matrix} \right\} \left(\int_{-\infty}^{+\infty} \delta(t)\, dt = 1 \right)\right.$

$$y(t) = x(t) * \left(\delta(t) + \alpha\delta(t - \Delta)\right)$$
$$= x(t) * \delta(t) + x(t) * \alpha\delta(t - \Delta)$$
$$= x(t) + \alpha x(t - \Delta)$$

回声隐藏（深入讨论）

- 鉴于自相关系数法对于回声定位的缺陷，必须采用其他方式。

1、考察回声嵌入过程：

$$\delta(t) = \begin{cases} \infty & (t=0) \\ 0 & (t\neq 0) \end{cases} \left(\int_{-\infty}^{+\infty} \delta(t)\,dt = 1 \right)$$

$y(t)$表示叠加回声后的信号，

$x(t)$表示原始信号，α表示衰减系数，

Δ表示延迟，$\delta(t)$表示单位冲击信号（单位冲击函数），

$*$表示卷积，则：

$$\begin{aligned} y(t) &= x(t) + \alpha x(t-\Delta) \\ &= x(t) * \big(\delta(t) + \alpha\delta(t-\Delta)\big) \end{aligned}$$

回声隐藏（深入讨论）

2、令$\mathrm{h}(t)=\delta(t)+\alpha\delta(t-\Delta)$，如果能分离$x(t)$和$\mathrm{h}(t)$，则定位延迟（确认$\Delta$）就容易得多。

3、由$y(t)=x(t)*h(t)$，及傅里叶变换特性可知：

$Y(j\varpi)$表示信号$y(t)$傅里叶变换信号，类似地，

$X(j\varpi)$和$H(j\varpi)$分别表示$x(t)$和$h(t)$傅里叶变换信号，则：

$$Y(j\varpi)=\mathcal{F}\big(y(t)\big)$$
$$=\mathcal{F}(x(t)*h(t))$$
$$=\mathcal{F}\big(x(t)\big)\cdot\mathcal{F}\big(h(t)\big)$$
$$=X(j\varpi)\cdot H(j\varpi)$$

回声隐藏（深入讨论）

4、为了进一步分离信号，对等式两边求对数：

$$\ln[\mathrm{Y}(j\varpi)] = \ln[\mathrm{X}(j\varpi)] + \ln[\mathrm{H}(j\varpi)]$$

5、对上式求傅里叶逆变换可得：

$$\hat{y}(t) = \mathcal{F}^{-1}\{\ln[\mathrm{Y}(j\varpi)]\}$$

$$= \mathcal{F}^{-1}\{\ln[\mathrm{X}(j\varpi)]\} + \mathcal{F}^{-1}\{\ln[\mathrm{H}(j\varpi)]\}$$

$$= \hat{x}(t) + \hat{h}(t)$$

相较于常规变换$x(t) = \mathcal{F}^{-1}\{\mathcal{F}(x(t))\}$，$\hat{x}(t) = \mathcal{F}^{-1}\{\ln[\mathcal{F}(x(t))]\}$在展开傅里叶逆变换之前增加了对数运算，称$\hat{x}(t)$为$x(t)$的倒谱信号。

式中，$\hat{y}(t)$、$\hat{x}(t)$和$\hat{h}(t)$分别是$y(t)$、$x(t)$和$h(t)$的倒谱信号。经过同态分析（傅里叶变换-〉求对数-〉傅里叶逆变换），信号$x(t)$和$h(t)$的卷积计算转变为倒谱信号$\hat{x}(t)$和$\hat{h}(t)$的求和计算。

最后，利用$h(t)$（$\hat{h}(t)$）的特性，回声延迟可以准确定位。

自然对数x=0处的泰勒级数(麦克劳林级数) $\ln(x+1)=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}x^n \quad \forall x\in(-1,1]$

回声隐藏（深入讨论）

6、进一步分析核函数。

已知：$h(t)=\delta(t)+\alpha\delta(t-\Delta)$

则：$H(j\varpi)=\mathcal{F}\big(h(t)\big)=\mathcal{F}\big(\delta(t)\big)+\alpha\mathcal{F}\big(\delta(t-\Delta)\big)=1+\alpha e^{-j\varpi\Delta}$

又因为当$|x|<1$时，$\ln(1+x)=x-x^2/2+x^3/3-\cdots$,

而应用通常选择$\alpha<1$，所以：

$$\ln(H(j\varpi))=\alpha e^{-j\varpi\Delta}-(\alpha^2/2)e^{-j2\varpi\Delta}+(\alpha^3/3)e^{-j3\varpi\Delta}-\cdots$$

因此：

$$\hat{h}(t)=\mathcal{F}^{-1}\{\ln[H(j\varpi)]\}$$
$$=\alpha\delta(t-\Delta)-(\alpha^2/2)\delta(t-2\Delta)+(\alpha^3/3)\delta(t-3\Delta)-\cdots$$

即：

$$\hat{y}(t)=\hat{x}(t)+\alpha\delta(t-\Delta)-(\alpha^2/2)\delta(t-2\Delta)+(\alpha^3/3)\delta(t-3\Delta)-\cdots$$

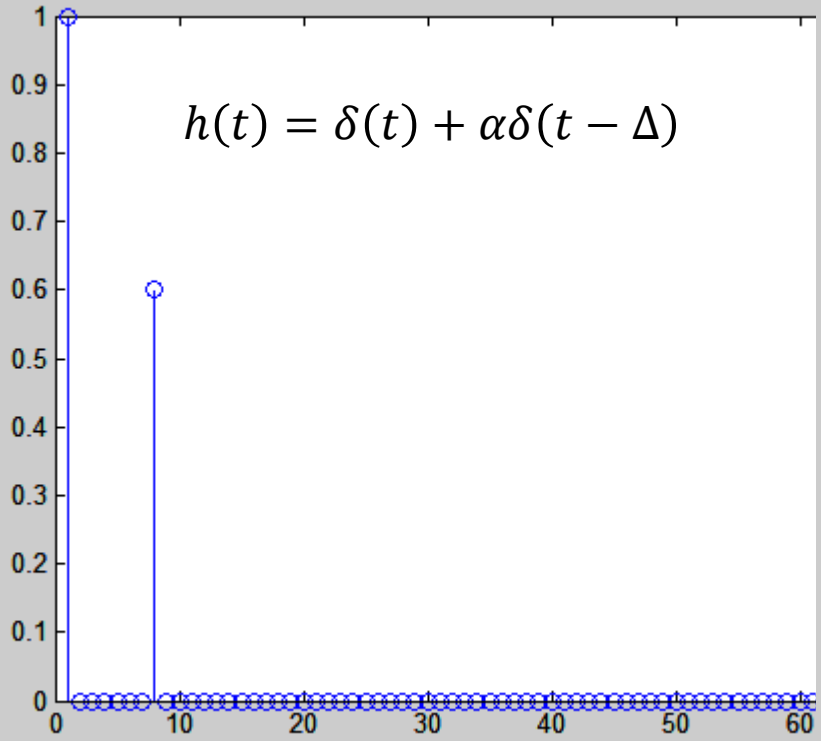
$\mathcal{F}[\delta(t)]=\int_{-\infty}^{+\infty}\delta(t)e^{-j\omega t}dt=e^{-j\omega t}\big|_{t=0}=1.$

即$\delta(t)$与1构成Fourier变换对 $\delta(t)\longleftrightarrow 1.$

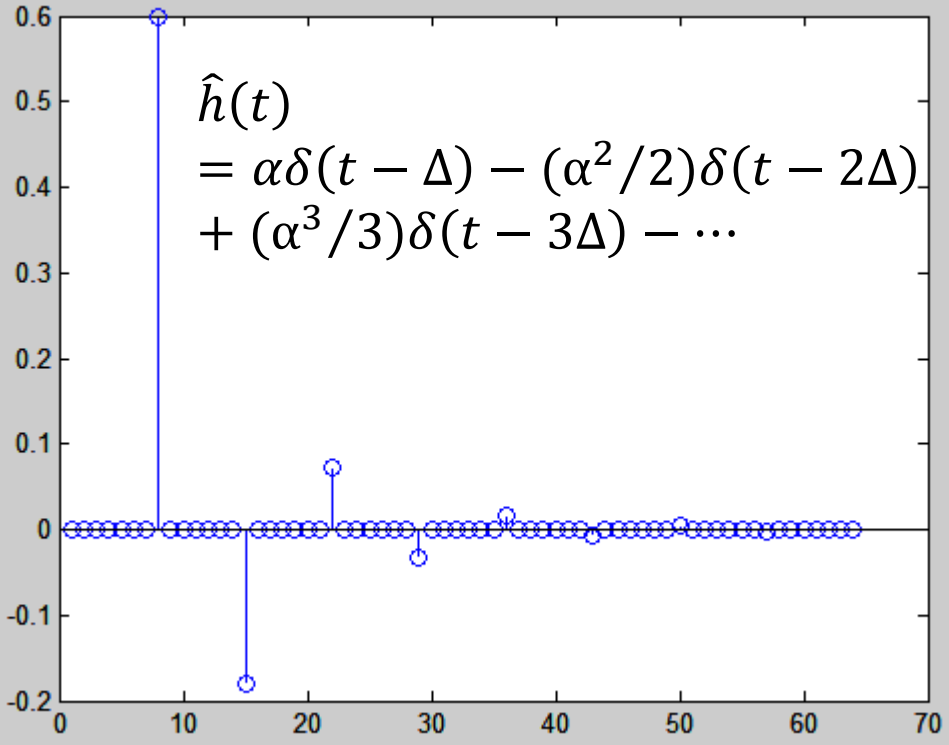
$\mathcal{F}\big(\delta(t-\Delta)\big)=\int_{-\infty}^{+\infty}\delta(t-\Delta)e^{-j\omega t}\,dt=e^{-j\omega t}\Big|_{t=\Delta}=e^{-j\omega\Delta}$

回声隐藏（深入讨论）

简单回声核$h(t)$函数

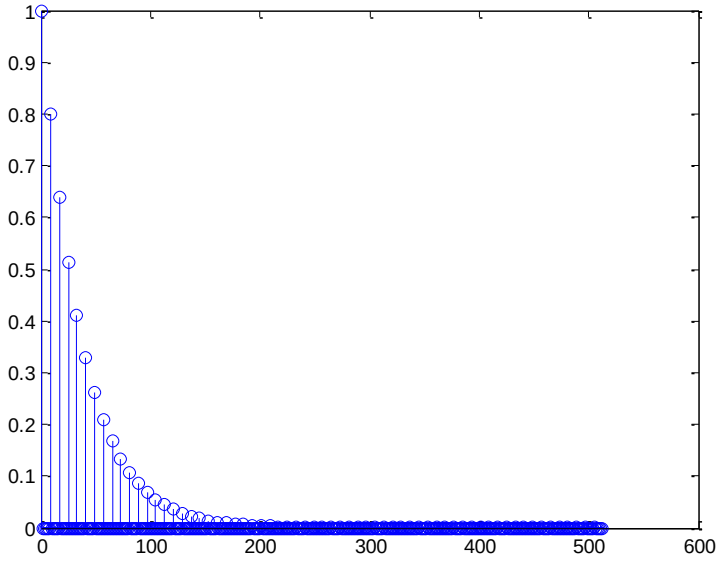


简单回声核倒谱（局部）

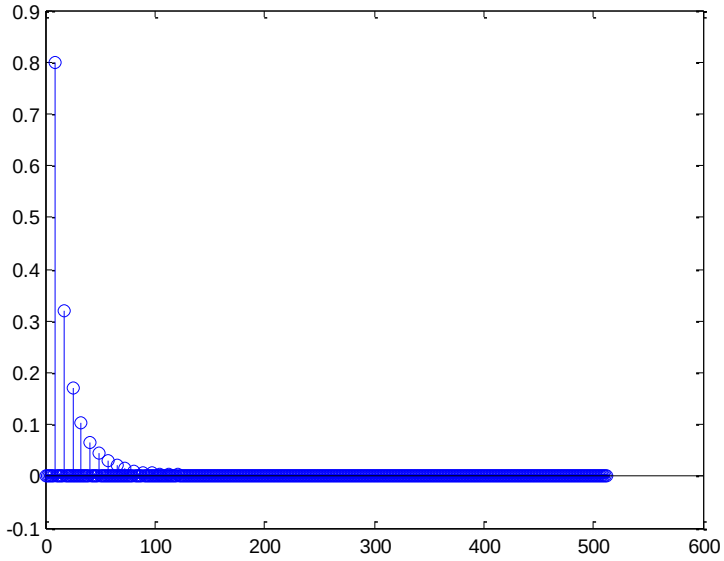


回声隐藏（深入讨论）

典型回声核

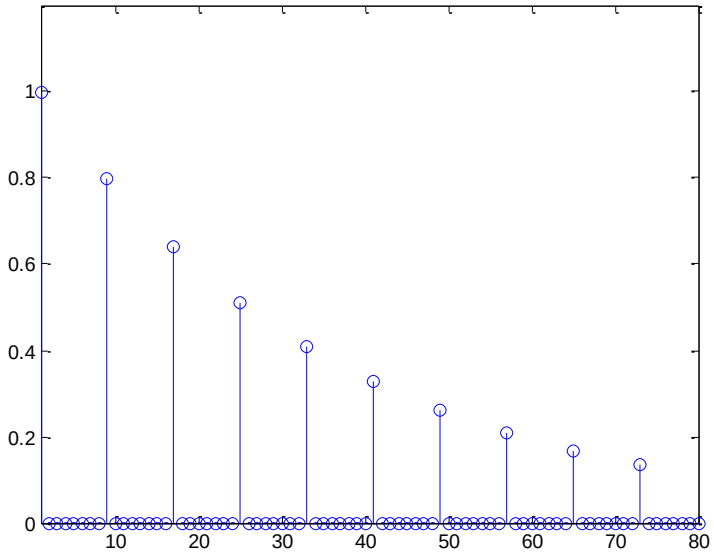


典型回声核的复倒谱

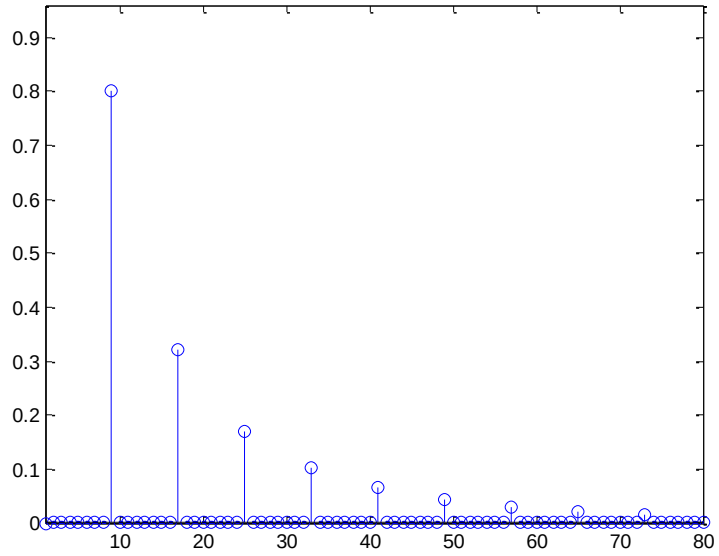


回声隐藏（深入讨论）

典型回声核（局部）

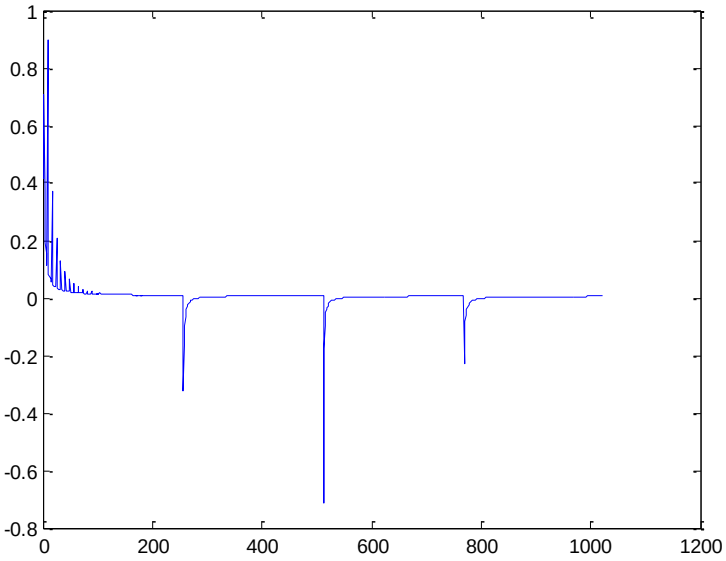


典型回声核的复倒谱（局部）

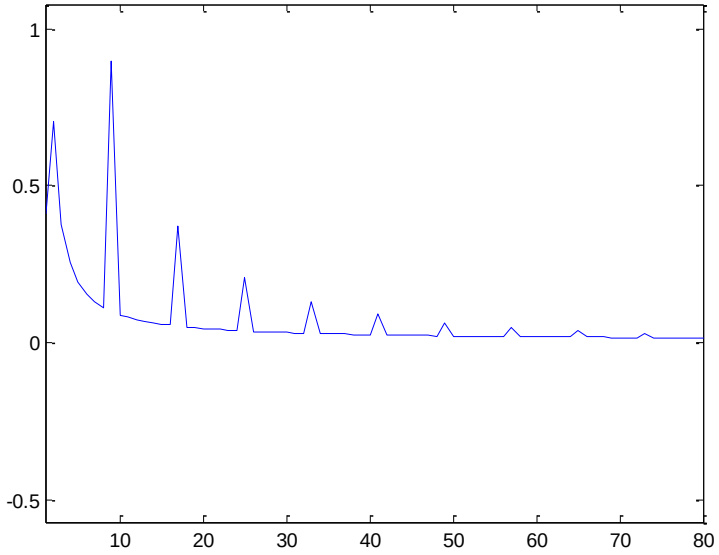


回声隐藏（深入讨论）

混合信号复倒谱

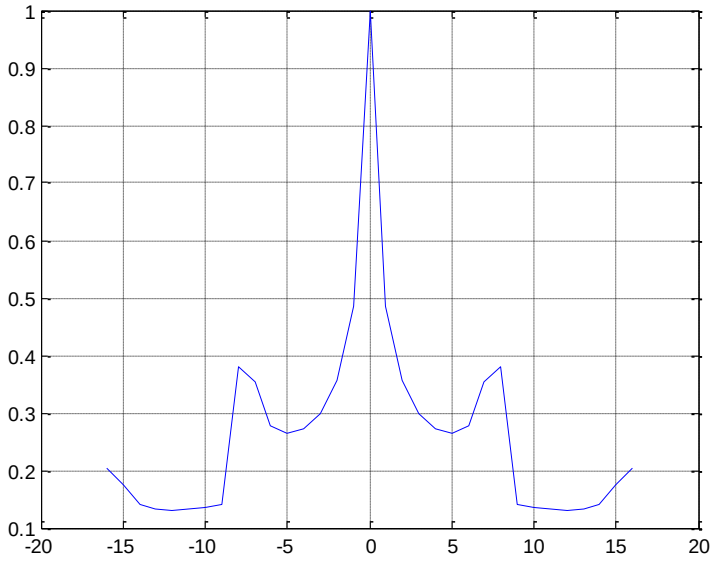


混合信号复倒谱（局部）

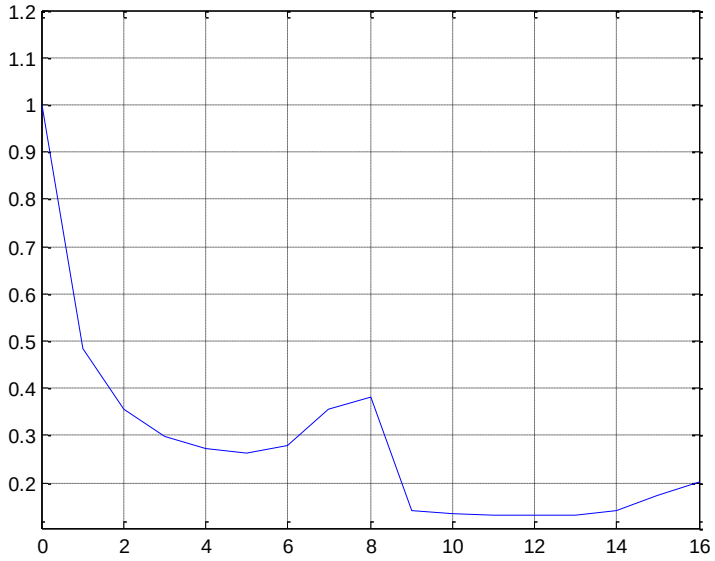


回声隐藏（深入讨论）

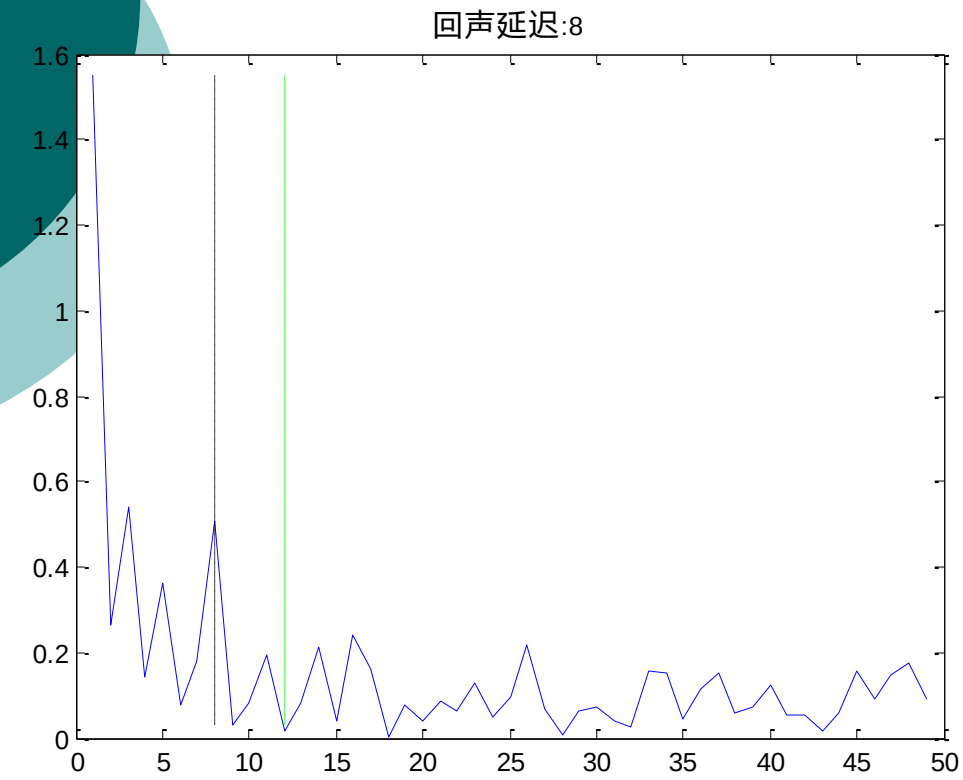
混合信号复倒谱自相关

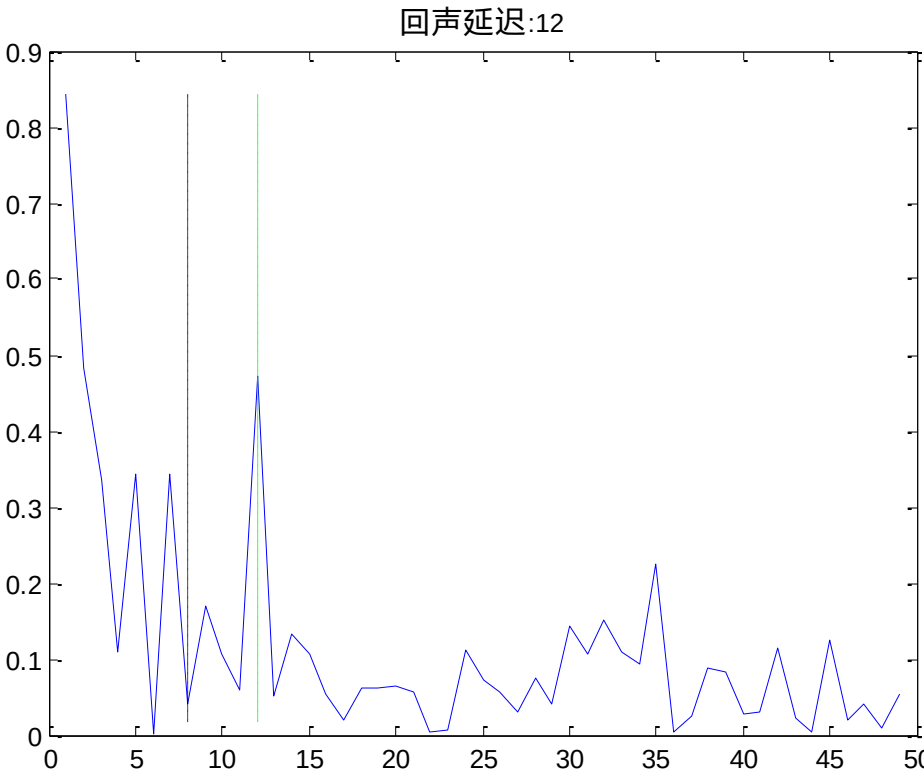


混合信号复倒谱自相关（局部）



回声隐藏（深入讨论）





回声延迟分别为8个和12个样点，左图为嵌入8个样点延迟回声后所得分段倒谱信号，右图倒谱信号对应延迟为12个样点。

回声隐藏

- 如何提取水印？
  - 使用倒谱自相关系数检测。
  - 倒谱自相关系数在回声延迟位置处有峰值

略过深入讨论

回声隐藏（深入讨论）

- 检测算法的进一步讨论
  - 已知计算倒谱包含对傅里叶变换信号求对数，如下所示：

$$\ln\big(X(j\varpi)\big) = \ln(|X(j\varpi)| \cdot \angle(X(j\varpi)))$$
$$= \ln(|X(j\varpi)|) + ln(\angle(X(j\varpi)))$$

式中$\angle(X(j\varpi))$表示复数的相位，由于相位具有多值性，所以计算时实际使用了$\pm\pi$区间范围的主值。对这样的信号求傅里叶逆变换时，丢失了其原始相位信息，产生了误差。

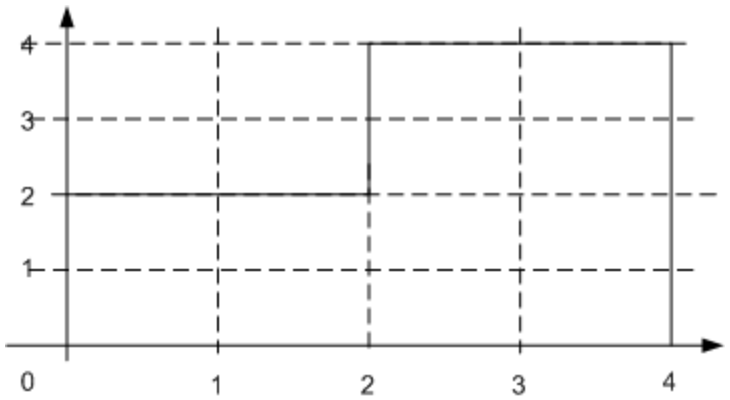
  - 至此，我们进一步细分，计算信号复对数产生的倒谱，准确称之为复倒谱。仅计算傅里叶变换信号模对数产生的倒谱，准确称之为实倒谱。

回声隐藏（深入讨论）

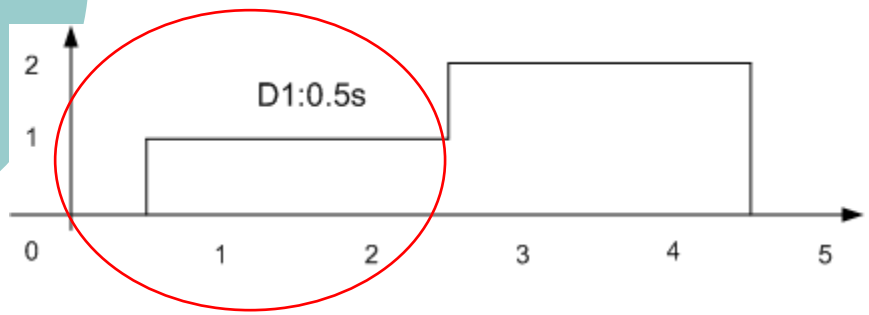
- 检测算法的进一步讨论
  - 若$x(t)$是最小相位信号，则其复倒谱$\hat{x}(t)$和其实倒谱$c_x(t)$具有以下关系：
  - $$\hat{x}(t)=\begin{cases}0, t<0\\ c_x(0), t=0\\ 2c_x(t), t>0\end{cases}$$
  - 实践中，语音信号多为（或近似为）最小相位信号，因此通常计算语音信号的实倒谱，完成回声信号的检测。

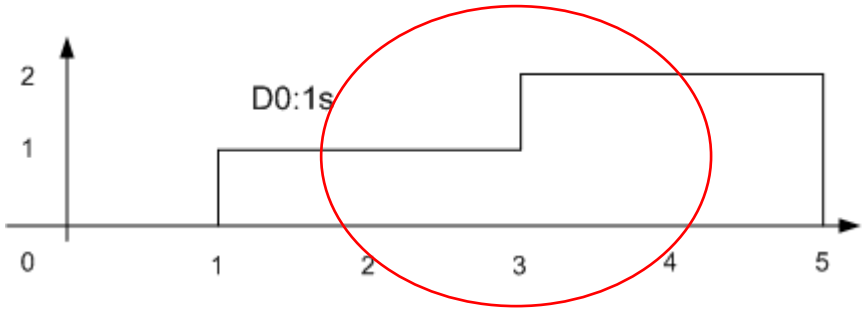
回声隐藏

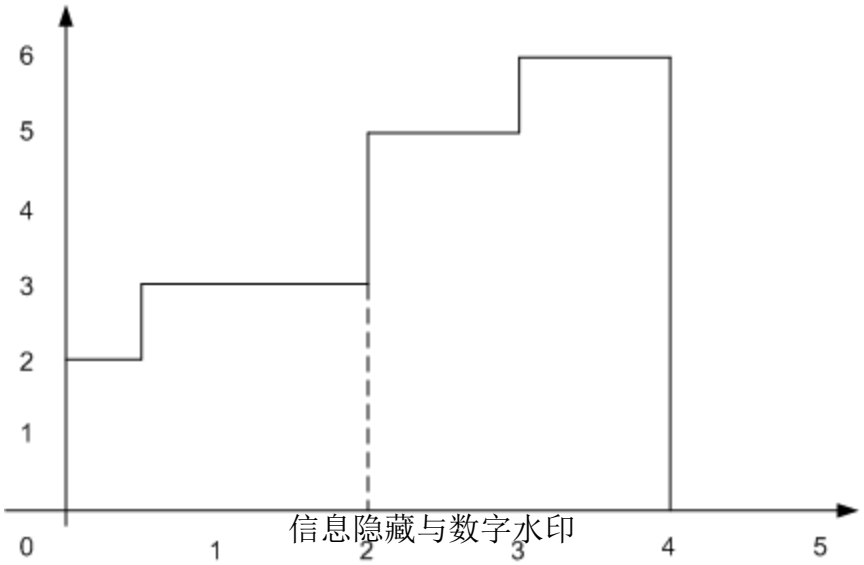
- 如何隐藏多个比特?
  - 语音信号分为多个片段，每个分段加入对应不同回声
- 例：若回声延迟为0.5毫秒代表比特“1”，回声延迟为1毫秒代表比特“0”，回声幅度衰减系数为0.5，分段长度为2毫秒，请给出下面信号嵌入比特“1,0”以后所得信号

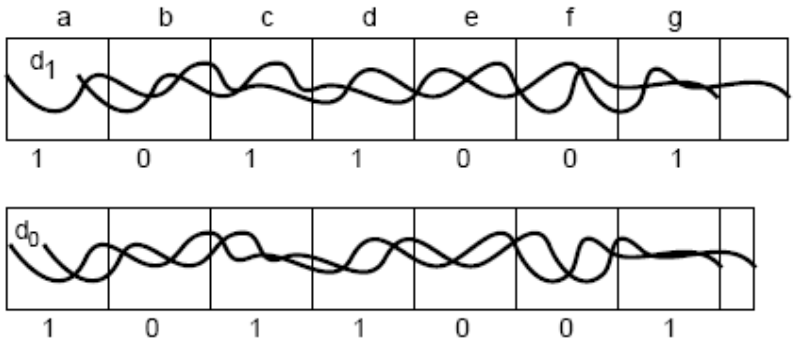


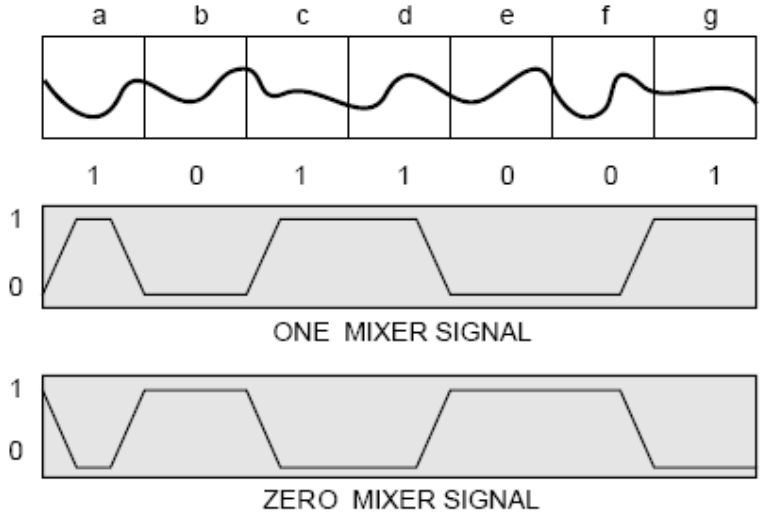
回声[illegible]



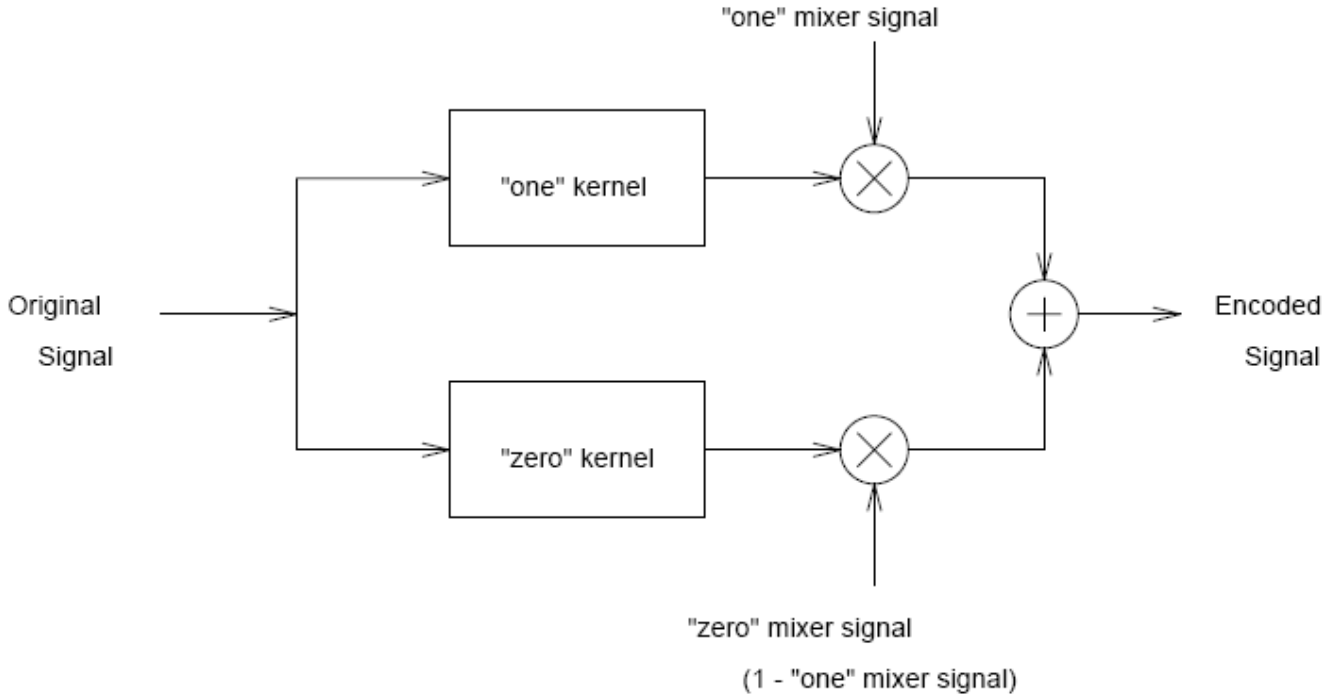





a b c d e f g
d_1
1 0 1 1 0 0 1
d_0
1 0 1 1 0 0 1


a b c d e f g
1 0 1 1 0 0 1
1
0
ONE MIXER SIGNAL
1
0
ZERO MIXER SIGNAL

- 回声隐藏系统图


"one" mixer signal
"one" kernel
Original
Signal
Encoded
Signal
"zero" kernel
"zero" mixer signal
(1 - "one" mixer signal)

回声隐藏

- 案例：回声隐藏
  - 已知音频片段采样值为以下序列：
  - $x[i]$:10, 12, 14, 8, 6, 8, 10,12,14,16,14,18, 16,18,20,22,20,22
  - 减系数为0.5，‘0’回声延迟为2个采样点，‘1’回声延迟为1个采样点，每6个样点为一个片段。
  - 已知秘密信息为(0,1,0)B，且隐藏后信号为：$y[i] = \delta x[i-d_0]\cdot m_0[i] + \delta x[i-d_1]\cdot m_1[i] + x[i]$
  - 请写出混合器信号$m_0[i], m_1[i]$（不用考虑过渡）。
  - 请写出隐写信号（即，携带秘密信息的信号$y[i]$）。

回声隐藏

- 案例：回声隐藏
  - 解：‘0’回声$\delta x[i-d_0]$为：
  - 0,0,5,6,7,4,3,4,5,6,7,8,7,9,8,9,10,11,10,11
  - ‘1’回声$\delta x[i-d_1]$为：
  - 0,5,6,7,4,3,4,5,6,7,8,7,9,8,9,10,11,10,11
  - 由$y[i]=\delta x[i-d_0]\cdot m_0[i]+\delta x[i-d_1]\cdot m_1[i]+x[i]$可知，每个时刻都有‘0’回声加权信号$\delta x[i-d_0]\cdot m_0[i]$和‘1’回声加权信号$\delta x[i-d_1]\cdot m_1[i]$，而同一时刻只能有‘0’回声信号或‘1’回声信号（否则难以检测）。所以，同一时刻，$m_0[i]$和$m_1[i]$，一个为0，另一个为1。

回声隐藏

案例：回声隐藏

解：因为6个样点为一个片段，而要隐藏的信息是（0,1,0)B，即第1、3分片‘0’回声有效，则：

i	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
$m_0[i]$	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	0	0
$\delta x[i-d_0]$	0	0	5	6	7	4	3	4	5	6	7	8	7	9	8	9	10	11	10	11
$x[i]$	10	12	14	8	6	8	10	12	14	16	14	18	16	18	20	22	20	22	0	0
$\delta x[i-d_1]$	0	5	6	7	4	3	4	5	6	7	8	7	9	8	9	10	11	10	11	0
$m_1[i]$	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0
$y[i]$	10	12	19	14	13	12	14	17	20	23	22	25	23	27	28	31	30	33	0	0

回声隐藏

- 稳健性分析
  - 演示一：
    - 利用回声隐藏原理嵌入和提取秘密信息。
    - 通过这个实验显示，回声隐藏算法能够在不显著影响音质的条件下，隐藏秘密信息。
    - 比对直接提取（未受攻击），和添加了25dB高斯噪声后提取的误码率。
    - 结果证明，回声隐藏算法抗高斯噪声攻击较强。

回声隐藏

- 稳健性分析
  - 演示二：
    - 对隐写语音分别进行A律、mu律压扩编解码。
    - 比对直接提取（未受攻击），和编解码后提取的误码率。
    - 结果证明，回声隐藏算法能够抵抗A律、mu律等压缩编码攻击。

回声隐藏

- 稳健性分析
  - 演示三：
    - 对隐写语音进行低通滤波，截至频率为3.4kHz
    - 比对直接提取（未受攻击），和低通滤波后提取的误码率。
    - 结果证明，回声隐藏算法能够抵抗低通滤波攻击

回声隐藏

- 稳健性分析
  - 演示四：
    - 对隐写语音进行播放、录制。
    - 比对直接提取（未受攻击），和模数/数模转换后提取的误码率。
    - 结果证明，回声隐藏算法能够抵抗模数/数模转换攻击。

回声隐藏

- 稳健性分析
  - 抗噪声性能
  - 抗A律、mu律压缩性能
  - 抗滤波性能
  - 抗D/A 、A/D转换性能
  - 适合使用电话信道的保密通信

回声隐藏

- 思考题
  - 回声隐藏参数：分段大小和回声衰减系数主要影响算法的哪些指标?
  - 根据实验数据分析以下参数条件下，算法的透明性、稳健性和容量，可得出什么结论?
    - 第一：回声衰减系数为0.6，分段大小分别为：1024，256

回声隐藏

- 思考题
  - 根据实验数据分析以下参数条件下，算法的透明性、稳健性和容量，可得出什么结论?
    - 第二：分段大小为1024，回声衰减系数为0.6和0.3

回声隐藏

- 性能分析
  - 容量
    - 与具体环境参数有关，若载体为8000Hz采样语音，则容量约为30比特每秒或更少
  - 透明性
    - 调整参数可以获得较好听觉效果
  - 稳健性
    - 抗失同步攻击性能较好
  - 容量、透明性和稳健性难以兼得

回声隐藏

- 研究方向
  - 回声核
    - 多回声
    - 多位置
    - 分频段
  - 自适应调整衰减系数
    - 幅度
    - 功率谱
    - 掩蔽曲线

略过深入讨论

回声隐藏的改进-前后向算法（深入讨论）

- 普通核函数可表示为：$h[n]=\delta[n]+\alpha\delta[n-d].$
- 已知复倒谱为：$c_y[n]=F^{-1}\left(\ln F\left(y[n]\right)\right)$
- 核函数傅里叶变换为：$H(e^{jw})=1+\alpha e^{-jwd}.$
- 利用级数展开为：$\ln H(e^{jw})=\alpha e^{-jwd}-\frac{\alpha^2}{2}e^{-2jwd}+\frac{\alpha^3}{3}e^{-3jwd}-\cdots$
- 则核函数倒谱为：$c_h[n]=\alpha\delta[n-d]-\frac{\alpha^2}{2}\delta[n-2d]+\frac{\alpha^3}{3}\delta[n-3d]-\cdots.$
- 其回声位置处幅值为：α

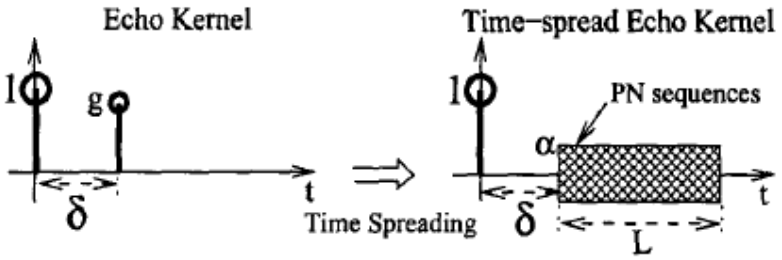
回声隐藏的改进-前后向算法（深入讨论）

- 前后向回声核函数表示为 $h[n]=\delta[n]+\alpha\delta[n-d]+\alpha\delta[n+d].$
- 其复倒谱为： $c_y[n]=F^{-1}\left(\ln F\left(y[n]\right)\right)$
- 同样利用级数展开核函数傅里叶变换，再做逆变换，则核函数倒谱为：

$$\begin{aligned}c_h[n]=&\alpha\{\delta[n-d]+\delta[n+d]\}\\&-\frac{\alpha^2}{2}\{\delta[n-2d]+2\delta[n]+\delta[n+2d]\}\\&+\frac{\alpha^3}{3}\{\delta[n-3d]+3\delta[n-d]\\&\qquad+\delta[n+d]+\delta[n+3d]\}-\cdots.\end{aligned}$$

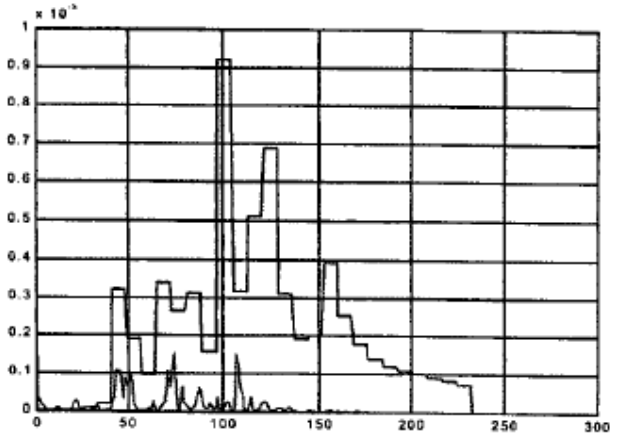
- 其回声位置处幅值为： $\alpha+\alpha^3+\alpha^5+\cdots$
- 即： $\alpha/(1-\alpha^2).$

回声隐藏算法改进-PN算法（深入讨论）



回声隐藏算法改进
-自适应算法（深入讨论）

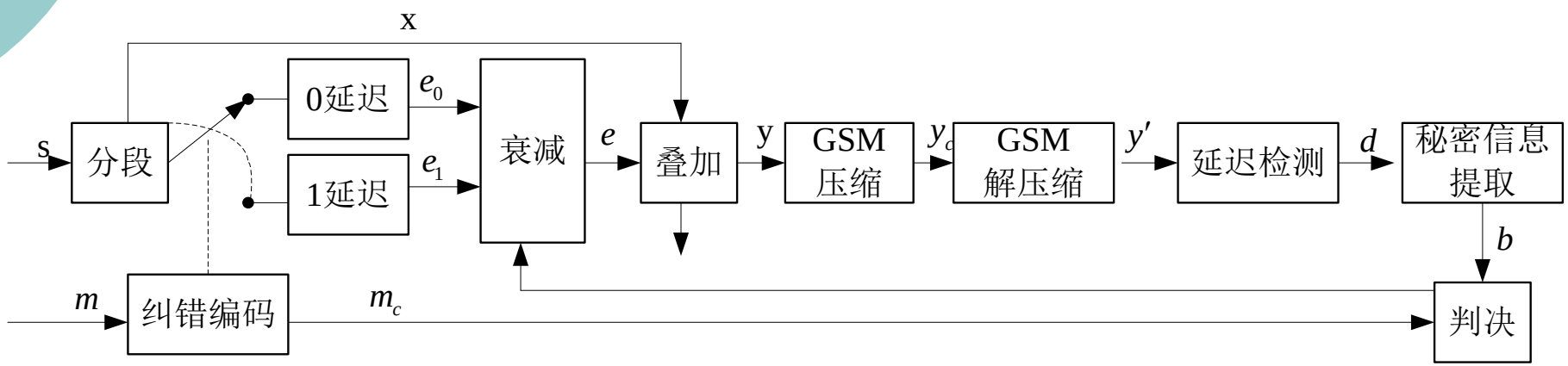
- 根据幅度
- 根据MPEG psycho-acoustic Model1



回声隐藏算法的改进
——多位置、合成分析

- 引入多组延迟，在多个位置检测回声。
- 隐藏后不立即输出秘密信息，经过信道后调整衰减系数，提取秘密信息，若能正确恢复，则输出，否则调整系数再次隐藏。

ABS算法流程


x
s
分段
0延迟
1延迟
e_0
e_1
衰减
e
叠加
y
GSM
压缩
y_c
GSM
解压缩
y'
延迟检测
d
秘密信息
提取
b
m
纠错编码
m_c
判决

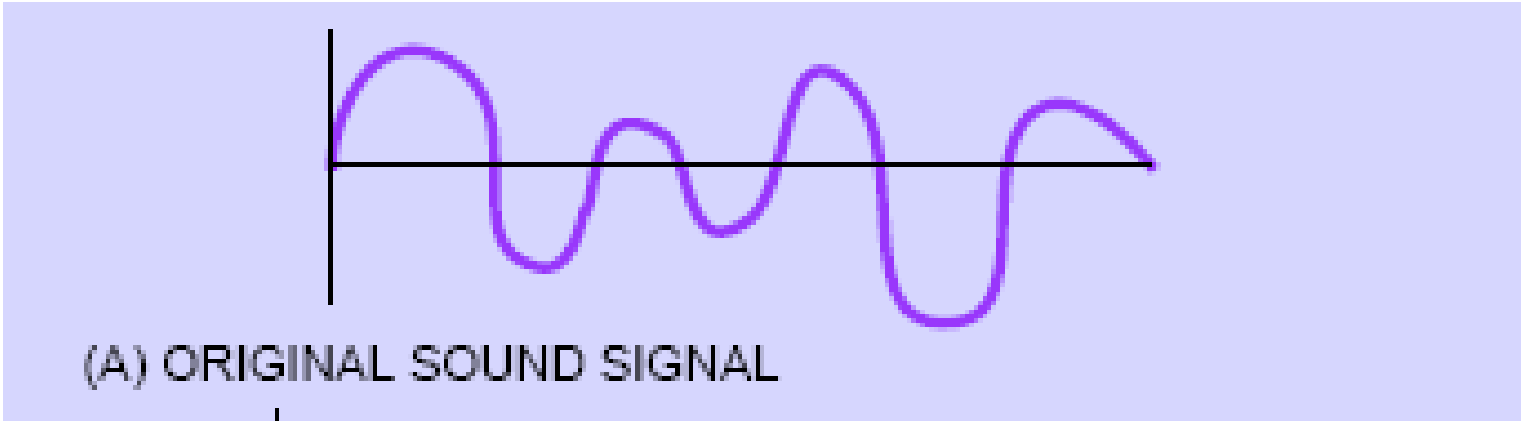
变换域音频水印

- 傅氏变换
- DCT变换
- 小波变换

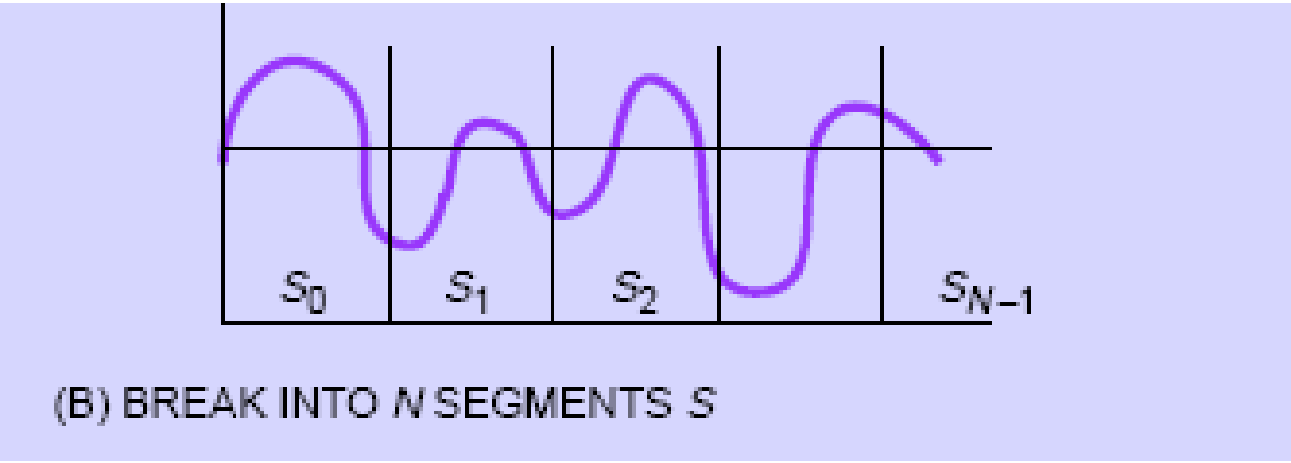
声音信号中的相位隐藏

- 人类的听觉系统特别敏感，声音信号中微弱的噪声都能够被察觉出来。
- 听觉系统对声音的绝对相位变化不太敏感
- 可以考虑在声音的相位中隐藏信息。

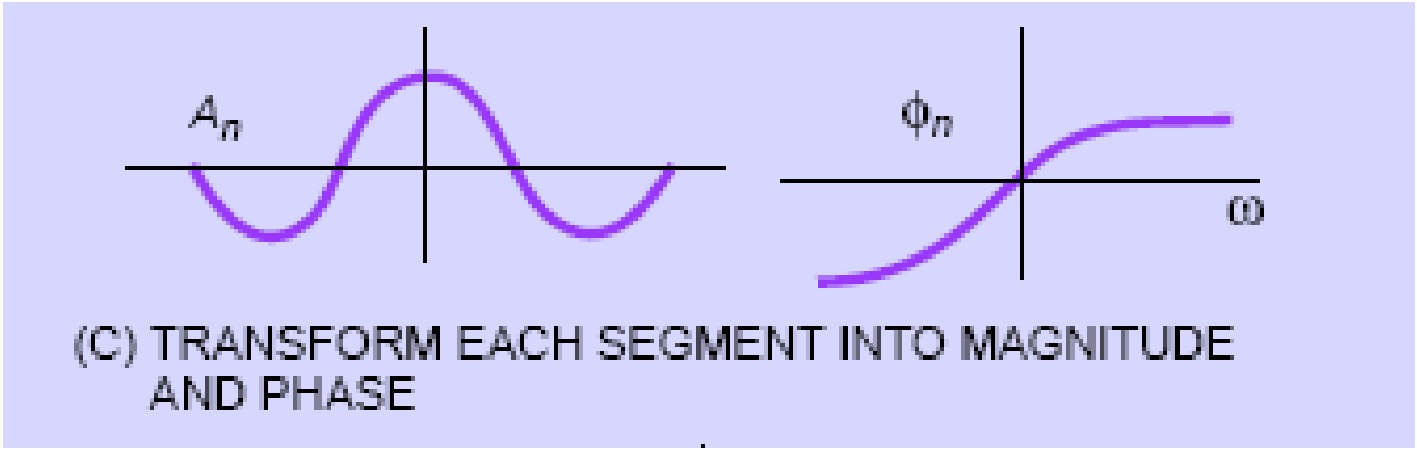
声音信号中的相位隐藏



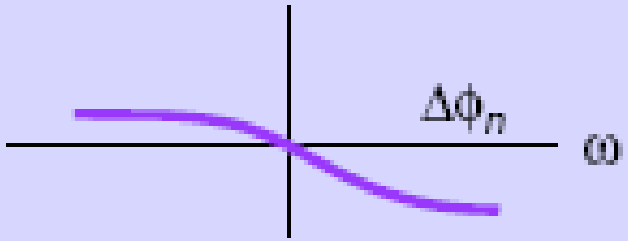
声音信号中的相位隐藏



声音信号中的相位隐藏

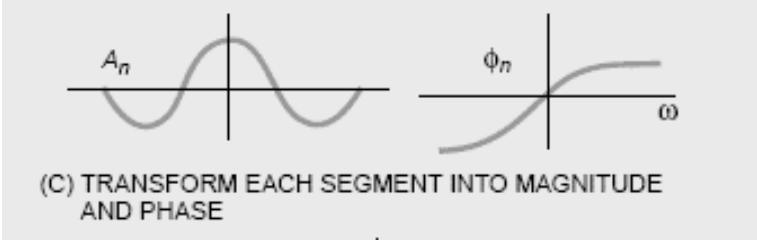

A_n
Φ_n
ω
(C) TRANSFORM EACH SEGMENT INTO MAGNITUDE AND PHASE

声音信号中的相位隐藏

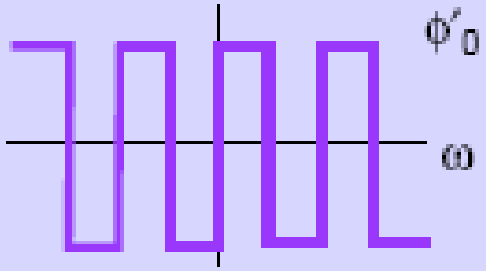


(D) CALCULATE THE PHASE DIFFERENCE BETWEEN CONSECUTIVE SEGMENTS $\phi_{n+1} - \phi_n$

声音信号中的相位隐藏

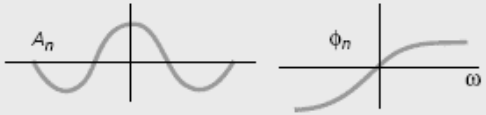


(C) TRANSFORM EACH SEGMENT INTO MAGNITUDE AND PHASE

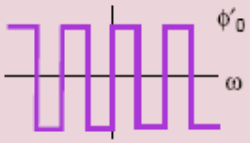


(E) FOR SEGMENT S_0 CREATE AN ARTIFICIAL ABSOLUTE PHASE ϕ'_0

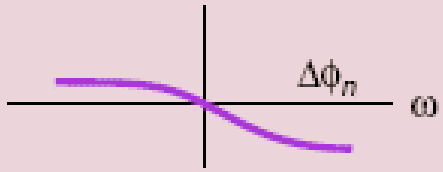
声音信号中的相位隐藏



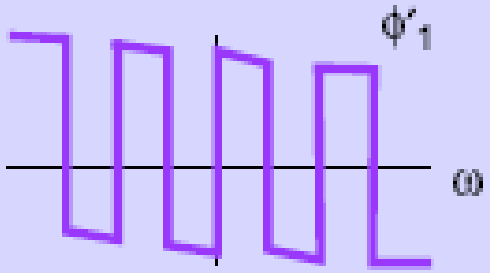
(C) TRANSFORM EACH SEGMENT INTO MAGNITUDE AND PHASE



(E) FOR SEGMENT S_0 CREATE AN ARTIFICIAL ABSOLUTE PHASE ϕ'_0

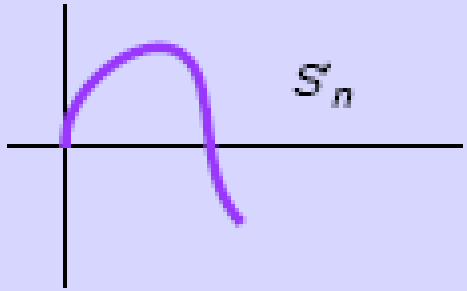


(D) CALCULATE THE PHASE DIFFERENCE BETWEEN CONSECUTIVE SEGMENTS $\phi_{n+1} - \phi_n$



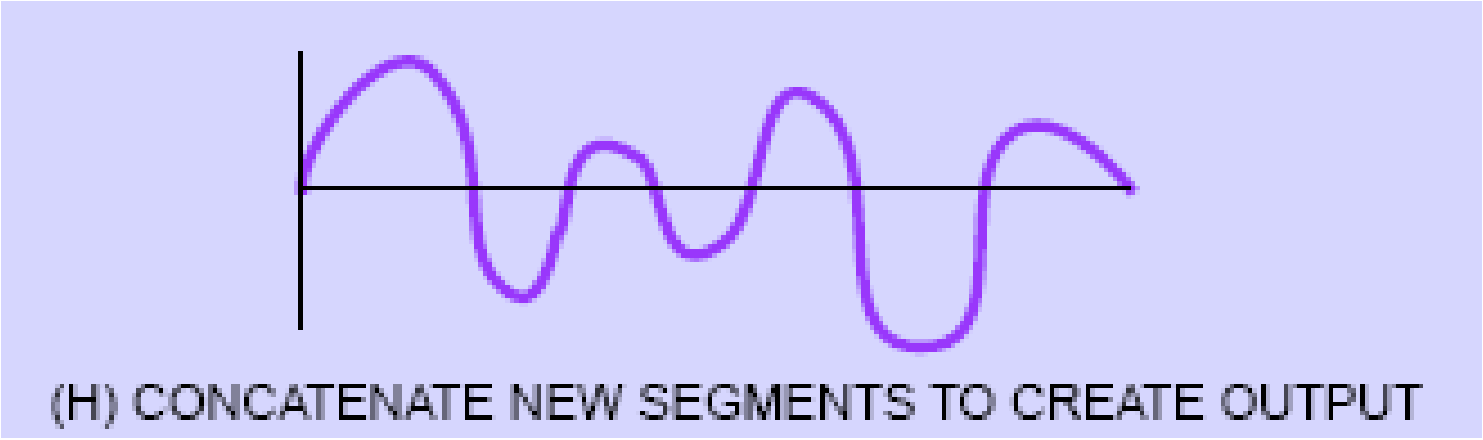
(F) FOR ALL OTHER SEGMENTS, CREATE NEW PHASE FRAMES $(\phi'_0+\Delta\phi_1)$

声音信号中的相位隐藏



(G) COMBINE NEW PHASE AND ORIGINAL MAGNITUDE TO GET NEW SEGMENT S'_n

声音信号中的相位隐藏



声音信号中的相位隐藏

- 嵌入
  - 将声音信号分段，分别进行DFT变换，得到每一段的幅频特性和相频特性。
  - 对第一个信号片段的相位进行修改，根据嵌入比特为“0”或“1”决定相位。并且在保持后面相位差不变的情况下修改绝对相位。

声音信号中的相位隐藏

- 嵌入
  - 结合原来的幅频特性和新的相位特性，计算傅立叶反变换，恢复语音信号。

相位隐藏步骤

1、 首先我们把一段长为 I 的语音信号 $s[n]$（$0 \le n \le I-1$）分割为 N 个片断，每个片段长为 K，其中一个片断可以表示为 $s_n[i], 0 \le n \le N-1, 0 \le i \le K-1, N = \left\lfloor \frac{I}{K} \right\rfloor$。

2、 对每一个片断作离散傅里叶变换，求取它们的相位、幅度谱。

$$FFT\left\{\begin{pmatrix} s_{0,0}, s_{0,1}, \ldots s_{0,K-1} \\ s_{1,0}, s_{1,1}, \ldots s_{1,K-1} \\ \cdots \\ s_{N-1,0}, s_{N-1,1}, \ldots s_{N-1,K-1} \end{pmatrix}\right\}, \quad A_n(k) = \sqrt{\mathrm{Re}\left(S_n(k)\right)^2 + \mathrm{Im}\left(S_n(k)\right)^2}$$

$$\phi_n(k) = \tan^{-1}\left(\frac{\mathrm{Re}\left(S_n(k)\right)}{\mathrm{Im}\left(S_n(k)\right)}\right)$$

3、 计算相邻两个分段的相位差，$\Delta\phi_n(k) = \phi_n(k) - \phi_{n-1}(k)$

4、 根据秘密信息构造第一个分段的绝对相位，隐藏 0 时，绝对相位为 pi/2，隐藏 1 时，绝对相位为-pi/2，得到 $\phi_0^{'}$

5、 利用步骤 3 中获得的相差矩阵，重新构造相位矩阵，$\phi_n^{'} = \phi_{n-1}^{'} + \Delta\phi_n$

6、 根据 $A_n(k)$ 和 $\phi_n^{'}(k)$ 作 IFFT，得到掩蔽载体

声音信号中的相位隐藏

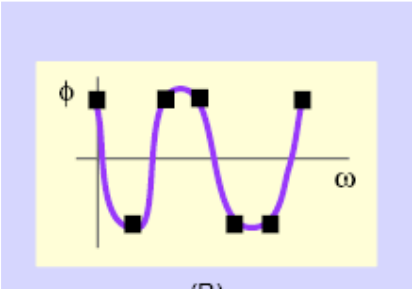
- 提取
  - 找到信号的分段，计算DFT，检测出初始相位，恢复秘密信息
  - 要求：信号同步

声音信号中的相位隐藏

- 思考题
  - 使用相位隐藏算法，若语音按照每段512个样点分割，那么100个分段最多隐藏多少比特秘密信息？为什么？

算法评估

- 1、样点的绝对相位发生了变化，但相邻片断间的相对相位保持不变，可以获得较好隐藏效果。
- 2、改变某些频率分量的相位，尽量使相位平滑，可以改善隐藏效果。



3、算法容量为8到32bps。

基于小波变换的音频水印算法

- 嵌入
  - 数字水印为一个随机信号
  - 选择适当的小波基对原始语音信号进行L级分解，在第L级的小波细节分量中嵌入水印
  - 水印嵌入算法

$$d_L'(i) = d_L(i)(1+\alpha x(i))$$

基于小波变换的音频水印算法

- 提取
  - 在水印检测端（作品所有者或第三方认证机构），原始的语音信号以及水印信号需要保留以备检测时用
  - 对L级分解的细节分量，利用原始语音信号找到隐藏了N个随机数的位置，求

$$x'(i) = (d_L'(i) / d_L(i) - 1) / \alpha$$

  - 计算提取水印与原始水印的相关值，判断是否有水印信号存在

基于小波变换的音频水印算法

- 算法特点
  - 一方面语音信号遮盖了水印的影响，使其不易被发觉
  - 另一方面即使受到一定的破坏，只要语音信号有一定的可懂度，水印信号就可以检测出来

基于小波变换的音频水印算法

- 算法特点
  - 单比特算法
  - 非盲检测

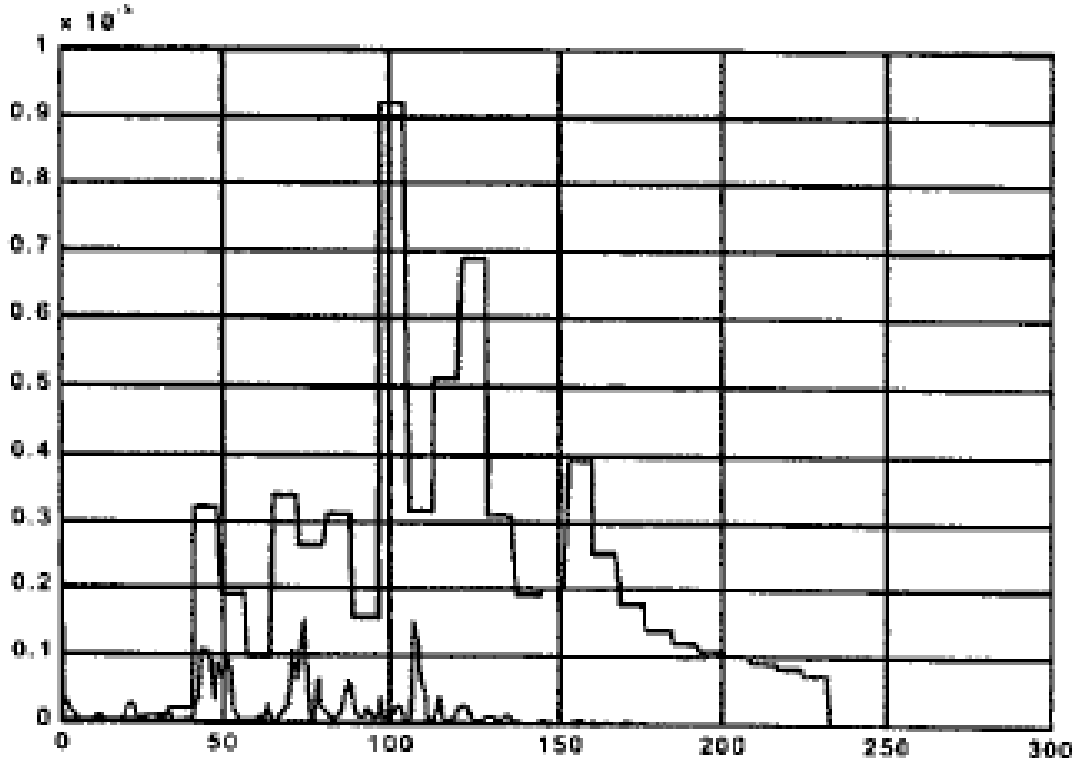
基于声音心理模型的信息隐藏算法

临界带和频谱

临界带（bark）	频率（Hz）			临界带（bark）	频率（Hz）		
	低频	高频	带宽		低频	高频	带宽
0	0	100	100	13	2000	2320	320
1	100	200	100	14	2320	2700	380
2	200	300	100	15	2700	3150	450
3	300	400	100	16	3150	3700	550
4	400	510	110	17	3700	4400	700
5	510	630	120	18	4400	5300	900
6	630	770	140	19	5300	6400	1100
7	770	920	150	20	6400	7700	1300
8	920	1080	160	21	7700	9500	1800
9	1080	1270	190	22	9500	12000	2500
10	1270	1480	210	23	12000	15500	3500
11	1480	1720	240	24	15500	22050	6550
12	1720	2000	280				

基于声音心理模型的信息隐藏算法

- 掩蔽曲线



基于声音心理模型的信息隐藏算法

掩蔽曲线计算方法

1、将频域映射到bark域，$z=13\arctan(0.00076f)+3.5\arctan\left[(0.000133f)^2\right]$，z取整为临界带编号。

2、计算bark域各子带能量。$E(z)=\sum_{\varpi=\varpi_L(z)}^{\varpi_H(z)}\left|X(j\varpi)\right|^2$，$\varpi_{L(z)}$和$\varpi_H(z)$为第z个bark域起止频率。

3、各子带内信号感知会受其他子带信号影响，可表达为：S(z)=E(z)*B(z)，B(z)为扩展函数，$B(z)=15.91+7.5(z+0.474)-17.5\sqrt{1+(z+0.474)^2}$

4、计算每个bark域的噪声特性因子$a(z)=\min\left[\frac{10\lg\left(G/A\right)}{S_{\max}},1\right]$，其中G、A分别为功率谱的几何平均和算术平均

5、根据噪声特性因子可得：$O(z)=(14.5+z)a(z)+5.5(1-a(z))$

6、计算每个子带的归一化阈值：$T_o(z)=\frac{1}{N(z)}10^{\lg S(z)-O(z)/10}$，与绝对阈值比较，较大的为听觉掩蔽阈值。

基于声音心理模型的信息隐藏算法

- 隐藏算法
  - 1、选取特定bark子带用于隐藏，例如15
  - 2、计算该bark域掩蔽阈值，T(15)
  - 3、统计掩蔽阈值下的频率分量个数，记为k
  - 4、根据秘密信息，调整前后k/2个频率分量的能量
  - 5、做逆变换重构语音

基于声音心理模型的信息隐藏算法

- 算法特点
  - 1、透明度高
  - 2、容量在每分钟30-50比特范围内
  - 3、鲁棒性强：
    - 抵抗MP3压缩
    - 同步要求低

扩频隐藏原理

- 原理:

扩频通信的抗干扰能力

扩频隐藏步骤

1、 将语音信号 $s[i]$ 分段为 $s_n[j]$，并作正交变换（T）得到， $S_n[k]$

2、 根据秘密信息，选择 PN 序列 pn_k，经过衰减系数加权，得到：$W = \alpha\, pn_k$

衰减系数的选择需要权衡透明性和鲁棒性。

3、 将调制后的秘密信息叠加到正交变换系数上， $S_n^{'} = S_n + W$

4、 作正交逆变换得到隐蔽载体， $s_n^{'} = T^{-1}\{S_n^{'}(k)\}$

基于MATLAB的算法分析

- 1、正交变换的选择

  DCT、DWT、FFT
- 2、嵌入位置的选择
- 3、衰减系数的确定

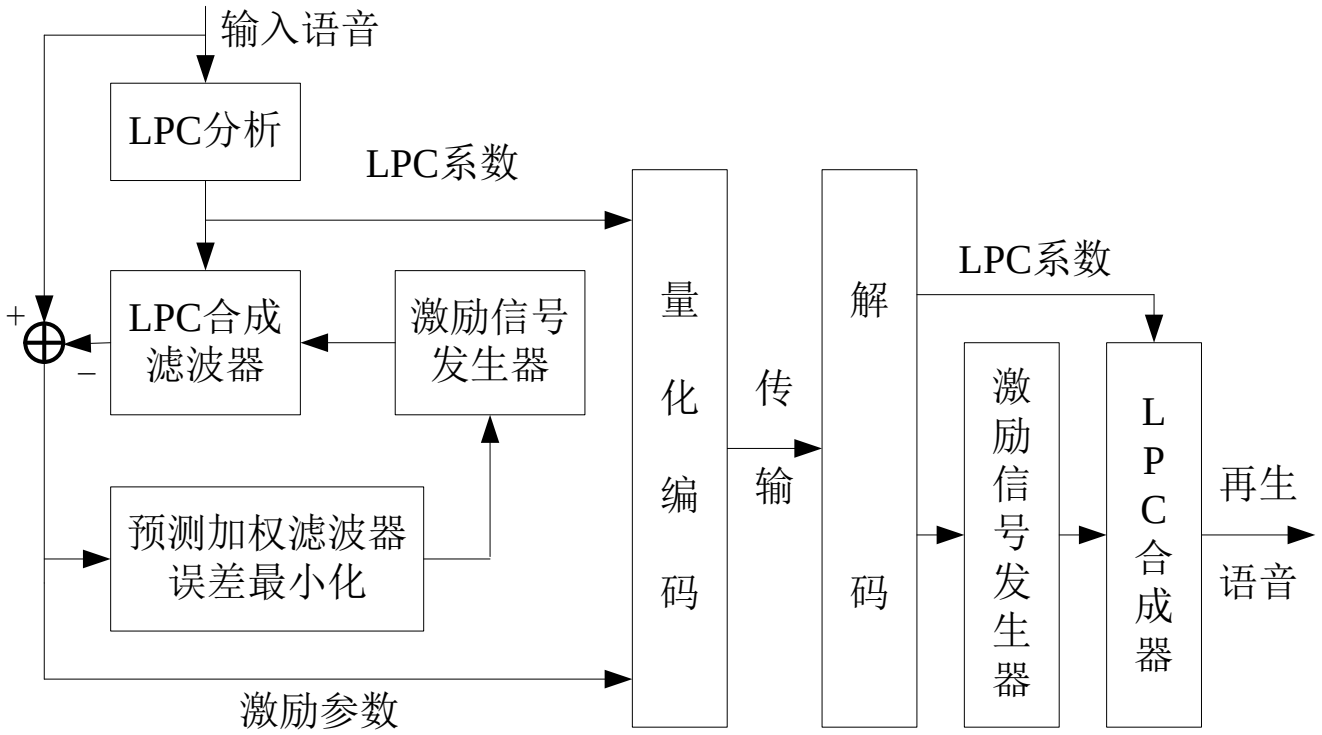
前沿问题

- 合成分析在信息隐藏中的应用
- 抗低比特率压缩编码算法的信息隐藏算法
- 信息隐藏中的同步问题

ABS语音编码算法

- ABS (Analysis-By-Synthesis)
- 将语音综合算法引入编码器，在编码器中产生与译码器端完全一致的合成语音，将此合成语音与原始语音进行比较，根据某种预定误差准则，对各个参数进行计算和调整，使得合成语音和原始语音之间的误差达到最小
- 在编码端就能知道解码端信号的效果

ABS语音编码算法



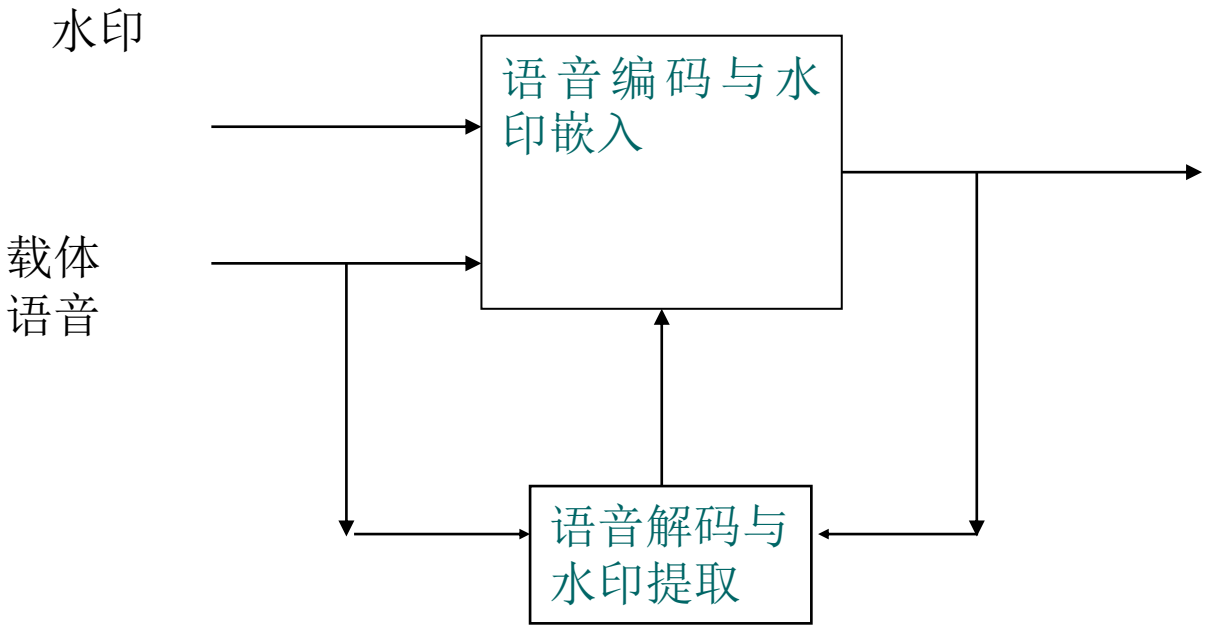
基于LPC的合成 - 分析系统结构

ABS语音编码算法

- 该基本原理可以引入信息隐藏之中，在语音信息隐藏编码方案中，引入语音合成器，将原始语音和嵌入后的隐蔽载体语音码流分别解码、合成复合后语音
- 根据一定的误差准则，对所有的可能嵌入方法所得到的隐蔽载体语音与原始语音比较并进行误差计算，确定出误差最小的一种嵌入方法，作为本帧入选的嵌入方案，其嵌入结果作为隐蔽载体语音码流输出

基于ABS的信息隐藏算法

- 在信息隐藏端就能够知道监听端的效果以及接收端的误码率



前沿问题

- 合成分析在信息隐藏中的应用
- 抗低比特率压缩编码算法的信息隐藏算法
- 信息隐藏中的同步问题

抗低比特率压缩编码算法
音频水印的分类

- 在原始音频信号中嵌入
- 在音频编码器中嵌入
  - 这种方法稳健性较高，但需要复杂的编码和解码过程，运算量大，实时性不好。

抗低比特率压缩编码算法
音频水印的分类

- 在压缩后的音频数据流中直接嵌入
  - 这种方法避免了复杂的编解码过程，但稳健性不高，而且能够嵌入的水印容量不大（压缩域数字水印）。

抵抗GSM压缩编码的语音隐藏算法

- GSM：码速率为13kbit/s 的，带有长时预测环节的规则脉冲激励线性预测编码器(RPE-LTP)
- 隐藏方法
  - 压缩编码后的码流中隐藏
  - 原始语音中隐藏

抵抗GSM压缩编码的语音隐藏算法

- 方案分析
  - 压缩编码后的码流中隐藏
    - 优点：抗GSM压缩编码；可采用ABS算法；
    - 问题：隐藏容量小；不实用，需要在手机内部嵌入模块。

抵抗GSM压缩编码的语音隐藏算法

- 方案分析
  - 原始语音中隐藏
    - 优点：可以在手机之外加模块
    - 问题：要求隐藏算法抗GSM压缩编码；

相邻分段能量比算法

- 基于语音能量比的隐藏算法
  - 可以有效抵抗GSM压缩编码
  - 统计压缩编码前后分段能量比，90%小于1.5
  - 调整相邻分段能量比，使之大于1.5，则90%的分段所隐藏的信息能够被正确提取

前沿问题

- 合成分析在信息隐藏中的应用
- 抗低比特率压缩编码算法的信息隐藏算法
- 信息隐藏中的同步问题

语音同步问题研究

- 在PSTN网、GSM网中的语音传输，语音信号经过数模、模数转换，同步信息丢失，要求一种模拟同步的算法
- 同步算法的思想：利用噪声的自相关和互相关性。在接收端可以利用滑动相关来进行同步检测