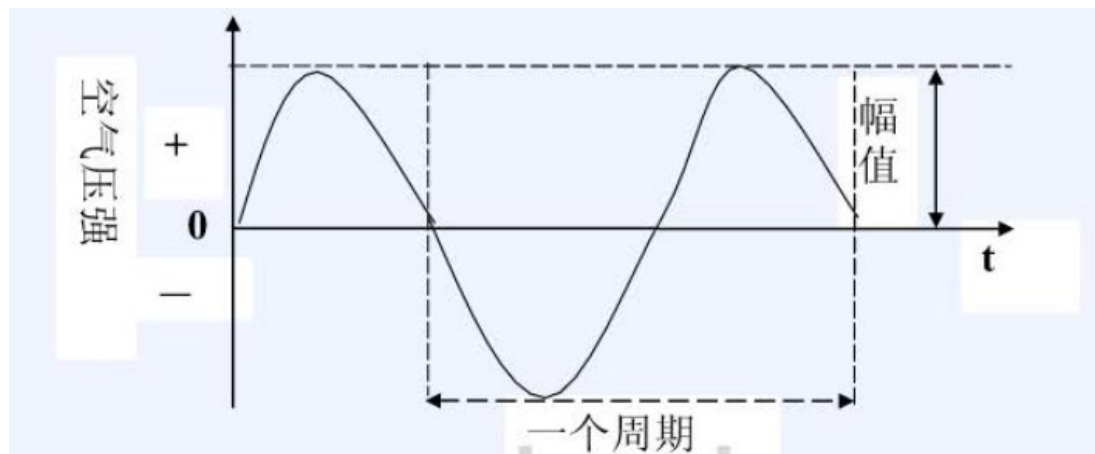


# 数字音频水印

# 人类听觉特点

## 频率

- 一个声源每秒钟可产生成百上千个波，通常把每秒钟波峰所产生的数目称之为信号的频率，单位用赫兹（Hz）或千赫兹（kHz）表示。
- 人耳能识别的声音频率范围大约在20～20kHz，通常称为音频信号



空气压强振荡的波形示意图

# 人类听觉特点

## 声压

- 从物理学的角度理解，大气压是空气分子的不规则运动及相互排斥所引起的。当空气中出现一种声音时，声音所产生的振动使空气分子在这个基础上产生有规律、有指向性的运动，改变了原来比较恒定的静压力，引起比原来静压力增高的量值就叫声压。
- 声波的震动可以使空气形成压缩状态和稀疏状态，从而造成原来大气静压力的增加或减少，所以声压的值可以是正值，有时也可以是负值。通常我们说的声压指的是它的有效值，所以实际上声压总是正值。声音产生的压力同声音的强度一样，变化范围极大。因此度量声压的大小，同样采用对数（数学里以 10 为底的对数，又叫常用对数）关系

# 人类听觉特点

听阈：

- 声音能被听见的最低声压
- 人耳对各频率声音所能听到的最小声压级所绘出的听阈曲线图中，2kHz~5kHz，人耳对这一频段的的声音最敏感

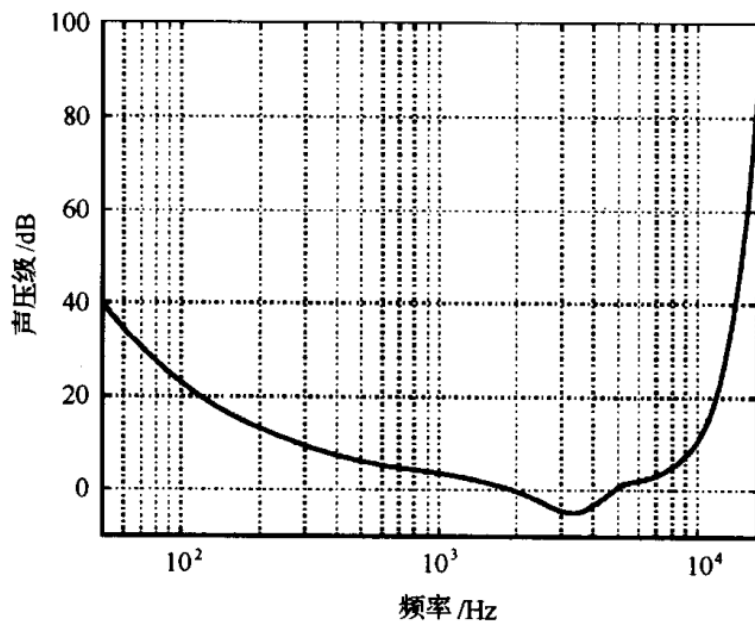


图 6-2 听觉系统在安静环境下的绝对听觉阈值

# 人类听觉特点

感觉域：

- 容忍的最高声压。当声压高到一定程度时，耳朵会出现不适感。对正常人而言，一般取120dB为不适阈，140dB为痛阈，且与频率无关。强度范围为-5dB～130dB

# 人类听觉特点

音调：

- 对于频率低的声音，听起来感觉它的音调“低，而频率高的声音，听起来感觉它的音调“高”
- 但是音调与声音的频率并不成严格的正比关系，它还与声音的强度及波形有关
- 频率和音调都是对声音属性的一种描述，频率是声音的一种物理特性，是客观描述，而音调是人耳对声音的一种主观感觉。是主观描述。

# 人类听觉特点

掩蔽效应：

- 当人耳听到两个强度不同的声音时，强的声音的频率成分会影响人耳对弱的声音的频率成分的收听
- 频率掩蔽：一个强纯音（分贝高）会掩蔽在其附近同时发声的弱纯音。
- 时间掩蔽：强声音后面的弱音容易被掩蔽（滞后掩蔽50-200ms）：强声音前面的弱音也容易被掩蔽（超前掩蔽5-20ms）。

# 人类听觉特点

掩蔽效应：

- 噪音对单音的掩蔽：一个单音可以被以它为中心频率，具有一定频带宽度的连续噪音所掩蔽
- 噪音掩蔽的临界带宽：如果在这一频带内噪声功率等于该纯音的功率，这时该纯音处于刚能被听到的临界状态，则称这一带宽为临界带宽

音频中的信息隐藏，主要利用人耳听觉特性以及掩蔽特性，设计隐藏算法。如回声隐藏等



# 语音质量评价

语音的质量一般从两个方面来衡量

- 清晰度：衡量语音中的字、单词和句子的清晰程度
- 自然度：衡量通过语音识别讲话人的难易程度

语音质量的评价方法

## 1 ) 主观评价

- 由人对语音的质量进行评价，因为语音最终是由人来收听，因此主观评价应该是最符合实际的，是对语音质量的真实反映。可使用平均意见分（ MOS ）方法

## 2 ) 客观评价

- 以机器为主体对图像质量进行评价

# 语音质量评价

## 主观评价

- 真实：反映了人对语音质量的真实感觉
- 费时费力，灵活性不够，重复性和稳定性较低，而且受试听者的主观影响较大

# 语音质量评价

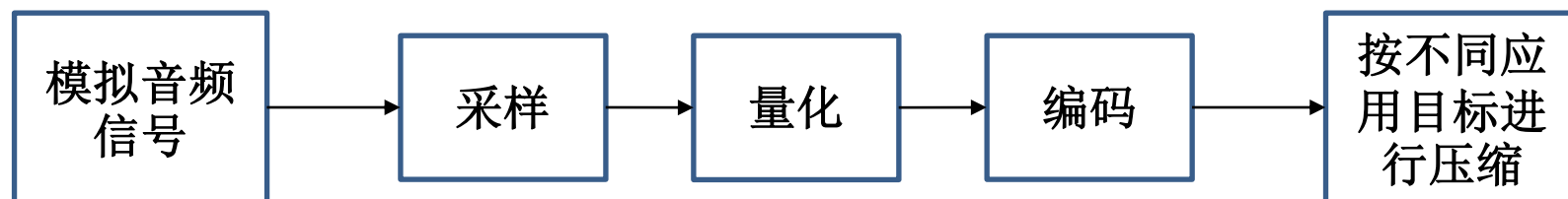
## 客观评价原理

对输入和输出语音信号进行分析和处理，从中提取出一些特征参量作为研究对象，最后设计一个“失真距离”，这个失真距离值跟提取出来的特征参量有关并由这些参量完全决定，以此“失真距离”值作为语音质量的客观评价值

## 客观评价

- 简单，可重复性强，与试听者主观感觉无关
- 无法达到与主观评价完全一致的效果

# 语音的基本表示

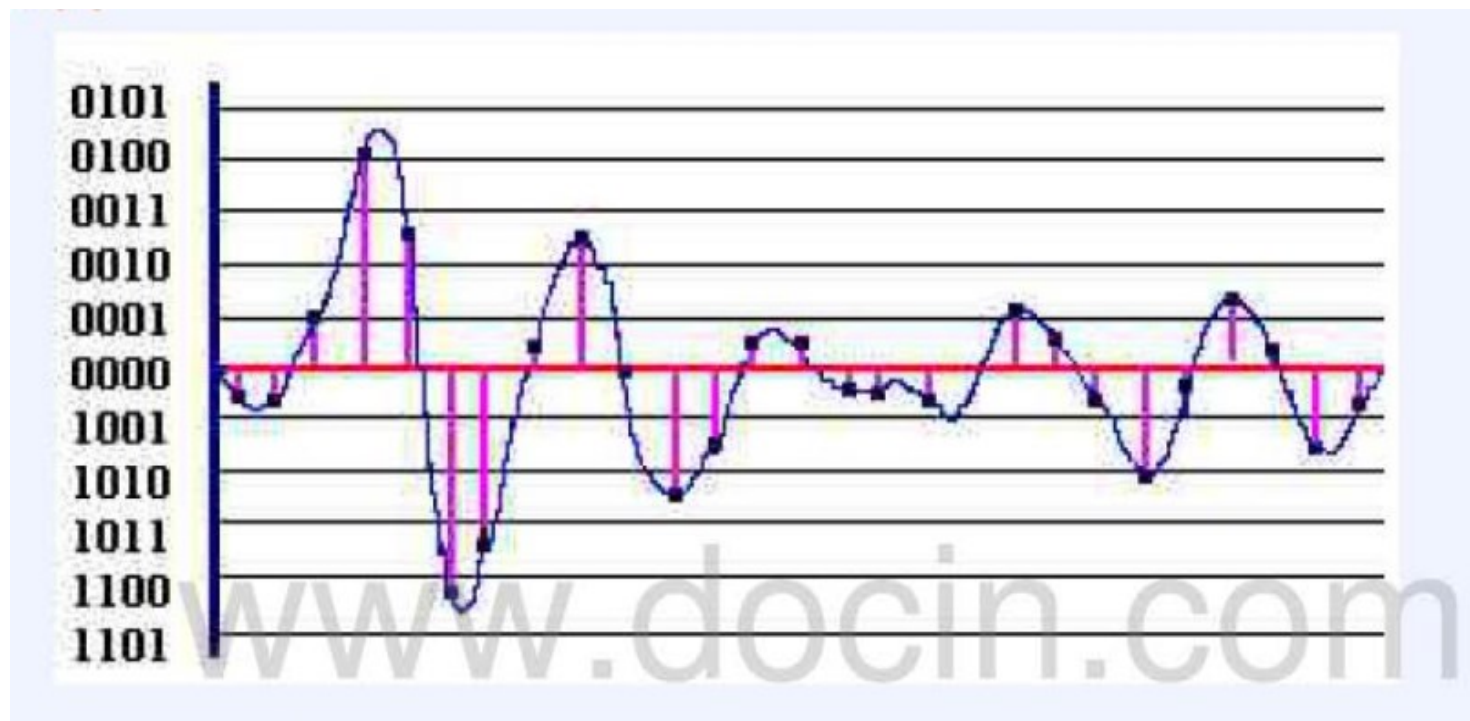


音频信息处理框图

# 语音的基本表示

## 1. 采样

采样就是在某些特定的时刻对模拟信号进行取值，如下图所示。采样的过程是每隔一个时间间隔在模拟声音的波形上取一个幅值，把时间上的连续信号变成时间上的离散信号。



# 语音的基本表示

## 2. 量化

每个采样值在幅度上进行离散化处理的过程称为量化。量化可分为均匀量化和非均匀量化。

均匀量化是把将采样后的信号按整个声波的幅度等间隔分成有限个区段，把落入某个区段内的样值归为一类，并赋予相同的量化值

非均匀量化是根据信号的不同区间来确定量化间隔。对于信号值小的区间，其量化间隔也小；反之，量化间隔就大。

量化会引入失真，并且量化失真是一种不可逆失真，这就是通常所说的量化噪声。

# 语音的基本表示

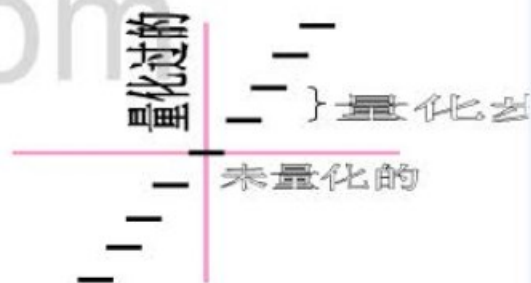
均匀量化



(1) 连续模拟声音信号



(2) 离散数据声音信号



(3) 幅值量化

# 音频水印

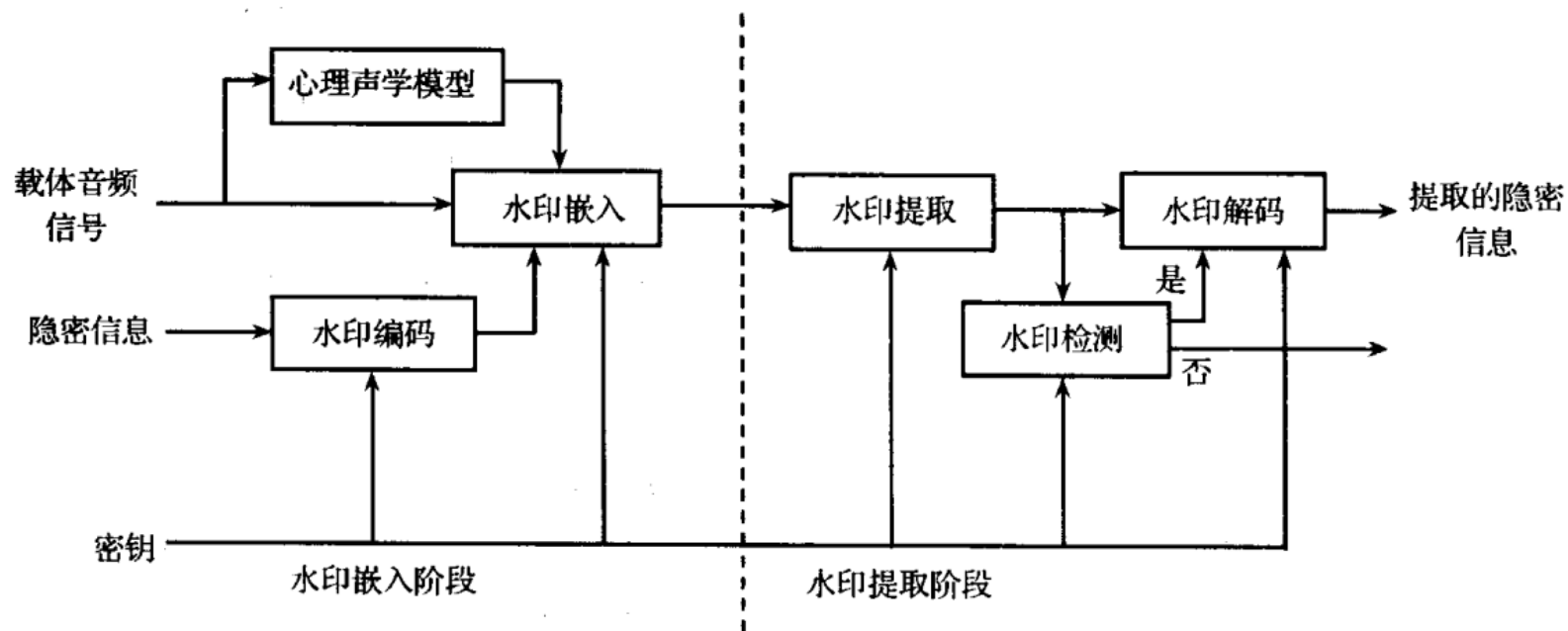


图 6-10 音频水印基本模型



# 音频水印算法

时域音频水印算法：

把水印信息直接嵌入到原始音频信号的采样点幅值上

变换域音频水印算法：

先对原始音频信号实施某种变换运算（如 DCT 变换、DFT 变换、DWT 变换等），然后根据人类听觉系统特性，通过修改某些变换系数的方式来嵌入水印信息。

压缩域音频水印算法：

直接把水印信号添加在经过压缩编码的音频信号中。

# 音频水印算法

时域音频水印算法：

易实现，速度快，嵌入信息量较大，鲁棒性普遍较低。

变换域音频水印算法：

抵抗干扰和抗恶意攻击的能力较强，鲁棒性好。但是是计算量较大

压缩域音频水印算法：

能够有效地避免压缩算法编解码的复杂过程。

# 时域音频数字水印算法

## 最不重要位（LSB）算法

任何形式的秘密数据（水印）都可以看作是一串二进制位流，而音频文件的每一个采样值也是用二进制数来表示的。该方法是用代表秘密数据的二进制位来替换音频信号的采样值的最不重要位（一般为最低位），以达到在音频信号中嵌入秘密数据的目的，提取水印时只要把相应的最低位取出来，就能恢复出嵌入的秘密信息。该方法是一种盲水印算法，在提取水印的过程中不需要原始音频信号的参与。

# LSB算法

优点：秘密数据嵌入和提取算法简单，速度快：音频信号里可编码的数据量大。

缺点：对某些信号处理技术比较敏感，也就是说对一些信号处理的鲁棒性较差。

# LSB算法

