

PCM音频编码

原创

Andy___Li

2017-12-27 17:30:00

17165

收藏 32

版权

分类专栏:

音视频

文章标签:

音视频

音频格式



音视频 专栏收录该内容

1 订阅

17 篇文章

订阅专栏

PCM语音编码

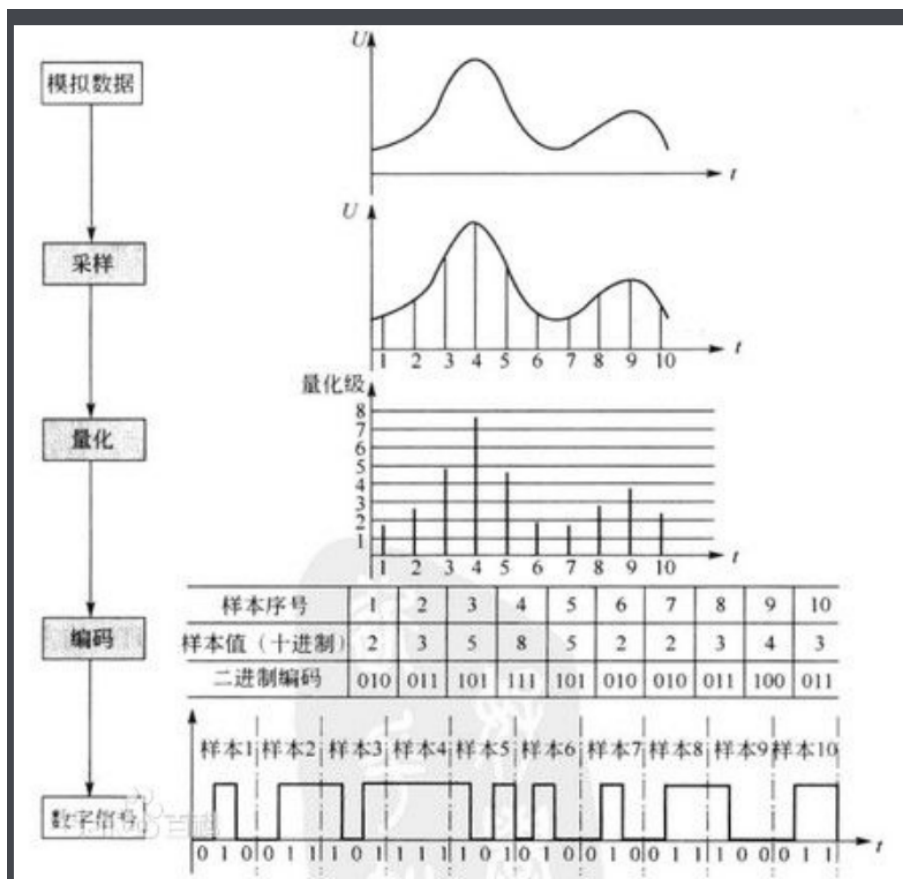
主要过程是将语音等模拟信号每隔一定时间进行取样，使其离散化，同时将抽样值按分层单位四舍五入取整量化，同时将抽样值按一组二进制码来表示抽样脉冲的幅值。也就是说语音信号最终以脉冲形式编码。

有一定电子基础的都知道传感器采集音频信号是模拟量，而我们实际传输过程中使用的是数字量。而这就涉及到模拟转数字的过程，下面将进行介绍。

1 PCM编码原理

PCM 脉冲编码调制是Pulse Code Modulation的缩写。脉冲编码调制是数字通信的编码方式之一。主要过程是将语音、图像等模拟信号每隔一定时间进行取样，使其离散化，同时将抽样值按分层单位四舍五入取整量化，同时将抽样值按一组二进制码来表示抽样脉冲的幅值。

模拟信号数字化必须经过三个过程，即抽样、量化和编码。



1.1 抽样

抽样是把模拟信号以其信号带宽2倍以上（奈奎斯特采样定理）的频率提取样值（采样率），变为在时间轴上离散的抽样信号的过程。

采样率：每秒从连续信号中提取并组成离散信号的sample:



Andy___Li

关注

7

如音频信号采样率为8000hz。

可以理解上图采样对应图中 那段电压随时间变化的曲线 为1秒 那下面那个1 2 3 ...10那就因该有1-8000个点，即将1秒均分为8000份，依次取出来那8000个点时间 对应的电压值。

1.2 量化

抽样信号虽然是时间轴上离散的信号，但仍然是模拟信号，其样值在一定的取值范围内，可有无限多个值。**必须采用“四舍五入”的方法把样值分级“取整”**，使一定取值范围内的样值由无限多个值变为有限个值。这一过程称为量化。

采样位数：指的是描述数字信号所使用的位数。

8位(8bit)代表 $2^8=256$ ，16 位(16bit)则代表 $2^{16}=65536$ ；

sample:

如音频传感器采集到的电压范围为0-3.3V，采样位数为8bit(位)

即我们把 $3.3V / 2^8 = 0.0128$ 即为量化精度。

我们把3.3v分成0.0128为步进的Y轴，如图3中的1 2 ...8就变成了0 0.0128 0.02563.3 V

比如某个采样点的电压值为1.652V(1280.128 与 1290.128之间) 我们对它四舍五入就为1.65V 对应 量化等级为128。

1.3 编码

量化后的抽样信号就转化为按抽样时序排列的一串十进制数字码流，即十进制数字信号。简单高效的数据系统是二进制码系统，因此，应将**十进制数字代码转换成二进制编码**。根据**十进制数字代码的总个数，可以确定所需二进制编码的位数，即字长(采样位数)**。这种把量化的抽样信号转换成给定字长的二进制码流的过程称为编码。

sample:

接着上面的1.65V 对应 量化等级为128。对应的2进制为10000000。即该采样点编码后结果为10000000。当然这是没有考虑正负值情况下的编码方式，而且编码方式种类非常多，需要具体问题具体分析。（PCM音频格式编码为A律13折线编码）

2 PCM音频编码

PCM信号未经过任何编码和压缩处理(无损压缩)。与模拟信号比，它不易受传送系统的杂波及失真的影响。动态范围宽，可得到音质相当好的效果。

2.1 PCM编码

编码上采用了的是A律13折线编码。

具体可参考：<http://blog.csdn.net/qingkongyeyue/article/details/52122486>

2.2 声道：

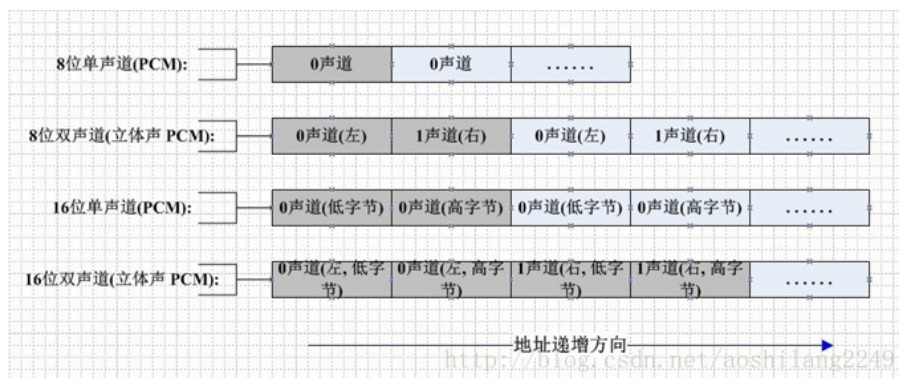
- 单声道声音文件，采样数据为八位的短整数（short int 00H-FFH）；
- 双声道立体声声文件，每次采样数据为一个16位（采样位数仍是8bit）的整数（int），高八位(左声道)和低八位(右声道)分别代表两个声道。

PCM的每个样本值包含在一个整数*i*中，*i*的长度为容纳指定样本长度所需的最小字节数。

首先存储低有效字节，表示样本幅度的位放在*i*的高有效位上，剩下的位置为0，这样8位和16位



的PCM波形样本的数据格式如下所示。



2.3 采样频率

人对频率的识别范围是 20HZ - 20000HZ, 如果每秒钟能对声音做 20000 个采样, 回放时就足可以满足人耳的需求.

- 8000hz 为电话采样。
- 22050 的采样频率是常用的。
44100已是CD音质, 超过48000的采样对人耳已经没有意义。

一般情况下, 一帧PCM是由2048次采样组成的。

2.4 采样位数

每个采样数据记录的是振幅, 采样精度取决于储存空间 (采样位数) 的大小:

- 1 字节(也就是8bit) 只能记录 256 个数, 也就是只能将振幅划分成 256 个等级
- 2 字节(也就是16bit) 可以细到 65536 个数, 这已是 CD 标准了;
- 4 字节(也就是32bit) 能把振幅细分到 4294967296 个等级, 实在是没必要了

如果是双声道(stereo), 采样就是双份的, 文件也差不多要大一倍.

2.5 比特率

码率是指经过编码后的音频数据每秒钟需要用多少个比特来表示

2.6 有损和无损

对于我们最常说的***“无损音频”来说, 一般都是指**传统CD格式中的16bit/44.1kHz采样率的文件格式**, 而知所以称为无损压缩, 也是因为其包含了20Hz-22.05kHz这个完全覆盖人耳可闻范围的频响频率而得名。

参考文献: <http://blog.csdn.net/ownWell/article/details/8114121>

显示推荐内容

打赏

文章很值, 打赏犒劳作者一下



优质评论可以帮助作者获得更高权重



评论



Andy__Li

关注

7