

波形编码技术

目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[规则](#)

[脉冲编码调制](#)

[过滤](#)

[采样](#)

[数字化语音](#)

[量子化和编码](#)

[压扩](#)

[A-law 和 u-law 压缩扩展](#)

[差分脉冲编码调制](#)

[自适应 DPCM](#)

[特定于 32 Kb/s 的步骤](#)

[相关信息](#)

简介

模拟通信已经很发达了，但是模拟传输还不是特别有效。当模拟信号由于传输损失而变弱时，很难将复杂的模拟结构从随机的传输噪音结构中分离出来。如果放大模拟信号，噪音也会放大，最终会导致模拟连接由于过于嘈杂而无法使用。只具有“一位”和“零位”状态的数字信号则更容易从噪音中分离出来。它们可以被无损放大。在长距离连接中，数字编码更不容易受到噪音损失的影响。此外，世界上的通信系统已转为使用一种名为“脉冲编码调制”(PCM) 的数字传输格式。PCM 是一种被称为“波形”编码的编码类型，因为它针对原始语音波形创建了一种编码形式。本文档概要介绍模拟语音信号到数字信号的转换过程。

先决条件

要求

本文档没有任何特定的要求。

使用的组件

本文档不限于特定的软件和硬件版本。

规则

有关文件规则的更多信息请参见“Cisco技术提示规则”。

脉冲编码调制

PCM 是 ITU-T G.711 规范中定义的一种波形编码方法。

过滤

将模拟信号转换为数字信号的第一步，就是过滤掉信号中的高频分量。这使得信号在下游更容易转换。大多数语音的能量都在 200 或 300 赫兹到 2700 或 2800 赫兹之间。针对标准语音和标准语音通信，建立了大约 3000 赫兹的带宽。因此，无需使用高精度的滤波器（这种滤波器非常昂贵）。从设备的角度来看，需要准备 4000 赫兹的带宽。这种频带限制滤波器用于防止失真（反失真）。在输入模拟语音信号的采样不足，即 Nyquist 准则中 $F_s < 2(BW)$ 时，就会发生这种情况。采样频率低于输入模拟信号的最高频率。这就会在样本的频谱和输入模拟信号之间造成重叠。而用于重建原始输入信号的低通输出滤波器则不足以检测到这种重叠。因此，它会创建一个来自信号源以外的新信号。这种在采样时创建错误信号的行为被称为“失真”。

采样

将模拟语音信号转换为数字语音信号的第二步，就是以固定的采样频率对已过滤的输入信号进行采样。这是通过脉冲振幅调制 (PAM) 过程实现的。此过程使用原始模拟信号，对具有恒定振幅和频率的脉冲序列的振幅进行调制。（请参阅图 2。）

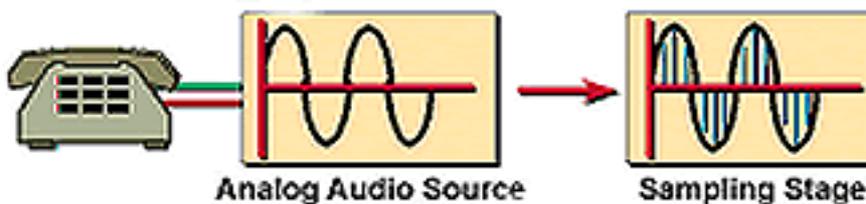
该脉冲序列以恒定的频率移动，称为采样频率。模拟语音信号的采样频率可以为每秒一百万次或每秒两到三次。如何确定采样频率？一位名叫 Harry Nyquist 的科学家发现，只要采样数量足够，就可以重建原始模拟信号。他确定，如果采样频率至少是原始输入模拟语音信号的最高频率的两倍，就可以使用低通滤波器在目标位置重建这个信号。Nyquist 准则的内容如下所述：

$$F_s > 2(BW)$$

F_s = Sampling frequency

BW = Bandwidth of original analog voice signal

图 1：模拟采样

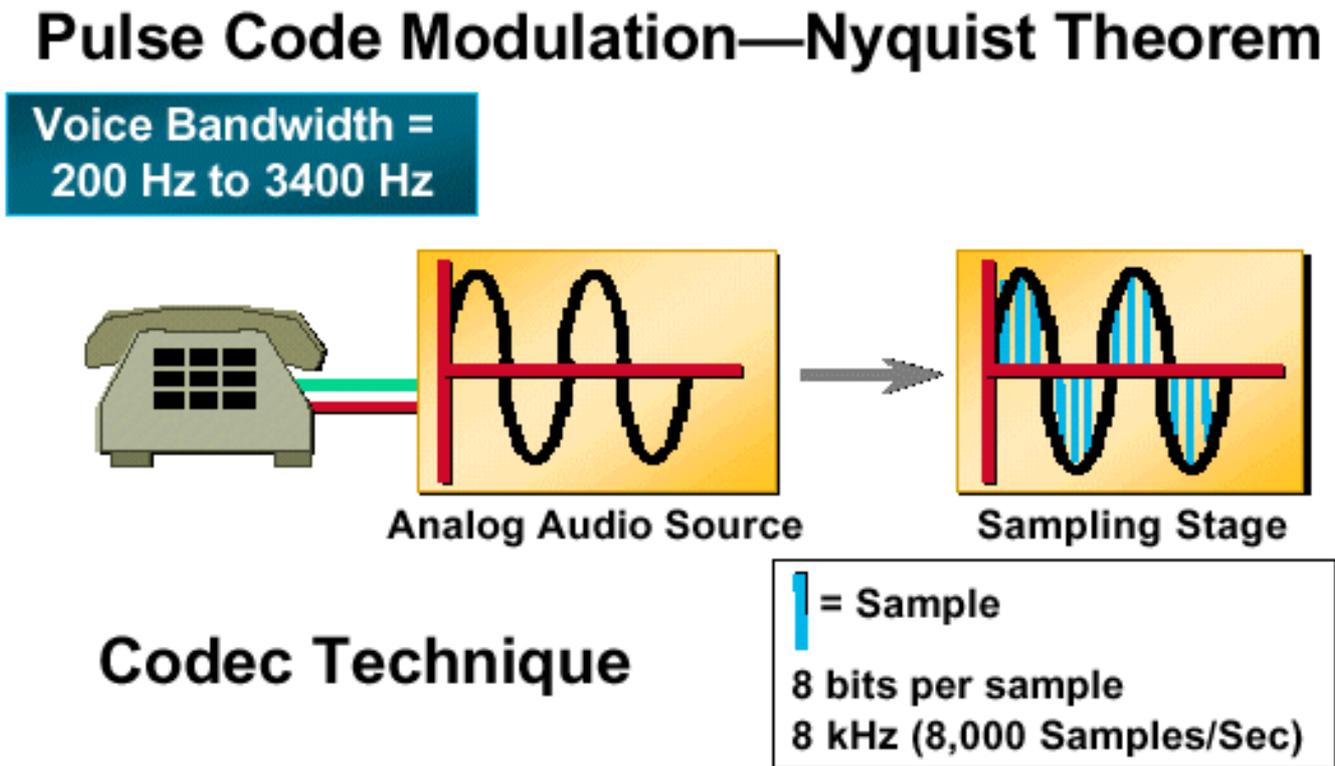


数字化语音

在对输入模拟语音信号进行过滤和采样（使用 PAM）后，下一步就是将这些样本数字化，为通过电话网络传输做准备。将模拟语音信号数字化的过程被称为 PCM。PAM 和 PCM 之间的唯一区别是，PCM 过程更进一步。PCM 使用二进制代码字对每个模拟样本进行解码。PCM 在源端有一个模数转换器，在目标端有一个数模转换器。PCM 使用量子化技术对这些样本进行编码。

量子化和编码

图 2：脉冲编码调制 - Nyquist 定理



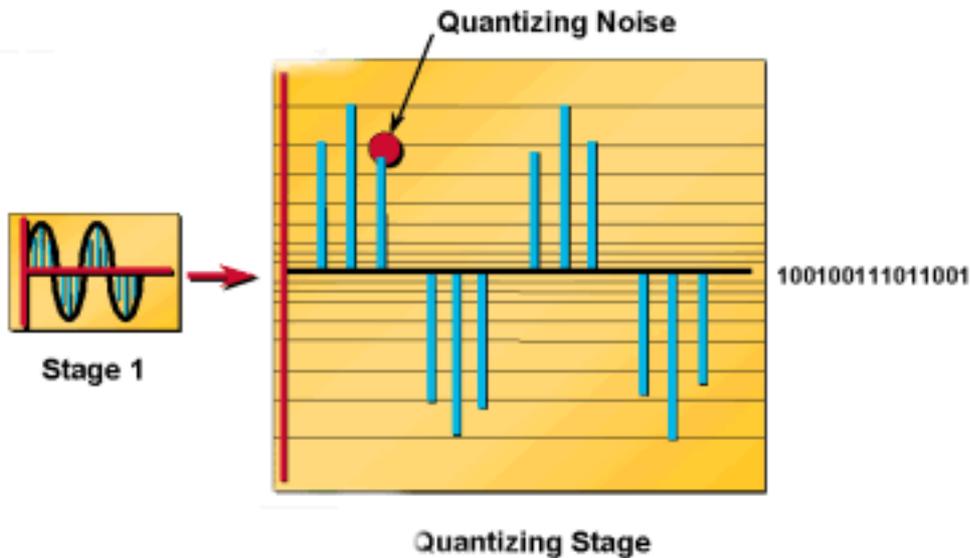
量子化是将每个模拟样本值转换为能够为其分配唯一数字代码字的离散值的过程。

当输入信号样本进入量子化阶段时，就会被分配到一个量子化间隔中。在整个输入模拟信号的动态范围内，所有量子化间隔之间的距离都是相同的（统一量子化）。每个量子化间隔都会被分配一个二进制代码字形式的离散值。使用的标准字大小是八位。如果对输入模拟信号以每秒 8000 次的频率进行采样，并且为每个样本都分配一个八位的代码字，则使用 PCM 的电话系统的最大传输比特率为 64,000 位/秒。图 2 显示了 PCM 系统的比特率是如何得出的。

每个输入样本都被分配了一个最接近其振幅高度的量子化间隔。如果为输入样本分配的量子化间隔与样本的实际高度不匹配，就会在 PCM 过程中引入一个错误。这种错误称为“量子化噪音”。量子化噪音相当于影响语音信号的信噪比 (SNR) 的随机噪音。SNR 用于衡量信号相对于背景噪音的强度。这个比率的测量单位通常是分贝 (dB)。以微伏为单位，如果流入的信号强度为 V_s ，噪音级为 V_n ，则以分贝为单位的信噪比 S/N 的计算公式为 $S/N = 20 \log_{10}(V_s/V_n)$ 。SNR 的测量单位为分贝 (dB)。SNR 越高，语音质量越好。量子化噪音会降低信号的 SNR。因此，如果量子化噪音增加，语音信号的质量就会下降。图 3 显示了量子化噪音是如何生成的。为了便于进行编码，一个 N 位的字会产生 $2N$ 的量子化标签。

图 3：模数转换

Pulse Code Modulation— Analog to Digital Conversion



降低量子化噪音的方法之一是增加量子化间隔的数量。当量子化间隔增加后（间隔增加会降低量子化噪音），输入信号振幅高度与量子化间隔之间的差别会减小。但是，随着量子化间隔的增加，代码字的数量也需要成比例增加。这个过程会引发其他问题，影响 PCM 系统处理更多代码字的能力。

在统一量子化中，SNR（包括量子化噪音）是影响语音质量的最重要因素。在整个输入模拟信号的动态范围内，统一量子化都使用相等的量子化级别。因此，低强度信号的 SNR 很小（低信号级别的语音质量），而高强度信号的 SNR 则很大（高信号级别的语音质量）。由于生成的大多数语音信号都是低强度的，因此获得更好的高信号级别的语音质量对于语音信号的数字化来说，是非常低效的方法。要改进较低信号级别的语音质量，需要将统一量子化（统一 PCM）替换为不统一的量子化过程，即压缩扩展。

压扩

压缩扩展是指首先在源位置压缩模拟信号，然后在该信号到达目标位置后将其扩展回原始大小的过程。术语“压缩扩展”是通过将“压缩”和“扩展”这两个术语合并得来的。在压缩扩展的过程中，输入模拟信号样本被压缩到对数段中。每个对数段随后通过统一量子化进行量子化和编码。压缩过程是一个对数过程。当样本信号变大时，压缩程度也会增大。换句话说，较大的样本信号被压缩的程度比较小的样本信号要大。这导致量子化噪音会随着样本信号的增大而增大。在输入样本信号的动态范围内，量子化噪音的对数式增加会让 SNR 在这个动态范围内保持不变。压缩扩展的 ITU-T 标准包括 A-law 和 u-law。

A-law 和 u-law 压缩扩展

A-law 和 u-law 是由国际电报电话咨询委员会 (CCITT) G.711 定义的音频压缩方案（编解码），将 16 位线性 PCM 数据压缩为 8 位的对数数据。

A-law 压缩扩展器

A-law 压缩是由以下方程定义的，它将线性样本值限制为 12 级位，其中 A 是压缩参数（在欧洲

A=87.7) , x 是要压缩的标准化整数。

$$F(x) = \begin{cases} \frac{A * |x|}{1 + \ln(A)} & 0 \leq |x| < \frac{1}{A} \\ \frac{\text{sgn}(x) * (1 + \ln(A|x|))}{1 + \ln(A)} & \frac{1}{A} \leq |x| \leq 1 \end{cases}$$

u-law 压缩扩展器

u-law (在本文中 u-law 和 Mu-law 表示相同的意思) 压缩是由以下方程定义的 , 它将线性样本值限制为 13 级位 , 其中 m 是压缩参数 (在美国和日本 m=255) , x 是要压缩的标准化整数。

$$F(x) = \frac{\text{sgn}(x) * \ln(1 + \mu|x|)}{\ln(1 + \mu)} \quad 0 \leq |x| \leq 1$$

欧洲和世界上其他国家或地区主要使用 A-law 标准。北美和日本使用 u-law。

A-law 和 u-law 的相似之处

- 都是对数输入/输出关系的线性近似值。
- 都使用八位代码字 (256 个级别 , 每个量子化间隔一个级别) 实现。八位代码字允许 64 千位/秒 (Kbps) 的比特率。计算方法是 , 用采样率 (输入频率的两倍) 乘以代码字的大小 (2 x 4 kHz x 8 位 = 64 kbps) 。
- 都将一个动态范围分成了 16 个分段 : 八个正分段和八个负分段。每个分段的长度都是前一个分段的两倍。在每个分段内都使用统一量子化。
- 使用相似的方法对八位字进行编码 : 第一位 (MSB) 标识极性。第二、三和四位标识分段。最后四位将低于 A-law 信号级别的分段进行量子化。

A-law 和 u-law 的不同之处

- 不同的线性近似法导致了不同的长度和斜率。
- 八位代码字中各个位的位置与不同分段之间的数值分配 , 以及分段内的量子化级别都是不同的。
- A-law 提供的动态范围比 u-law 更大。
- u-law 针对低级别信号提供的信号/失真性能比 A-law 要高。
- 对于统一 PCM 对等值 , A-law 要求使用 13 位数据。对于统一 PCM 对等值 , u-law 要求使用 14 位数据。
- 国际连接需要使用 A-law , 使用 u-law 的国家或地区负责 u-law 到 A-law 之间的转换。

差分脉冲编码调制

在 PCM 过程中 , 输入样本信号之间的差别很小。差分 PCM (DPCM) 用于计算这种差别 , 然后传输这种小差别信号 , 而不是整个输入样本信号。因为各个输入样本之间的差别比整个输入样本要小

，所以传输所需的位数也相应减少。因此传输语音信号所需的吞吐量也可以减少。使用 DPCM 可以将语音传输的比特率降至 48 Kbps。

DPCM 如何计算当前样本信号与上一个样本信号之间的差别？DPCM 的前一部分的工作方式与 PCM 非常像（这也是它被称为差分 PCM 的原因）。系统以固定的采样频率（输入频率的两倍）对输入信号进行采样。然后使用 PAM 过程对这些样本进行调制。从此时开始，DPCM 过程将接管剩余的部分。样本输入信号被储存在预报器中。预报器提取存储的样本信号，然后通过一个差分器将信号发送出去。差分器将前一个样本信号与当前的样本信号进行比较，然后将两者的差别发送到 PCM 的量子化和编码阶段（在这个阶段可以进行统一量子化处理，也可以使用 A-law 或 u-law 进行压缩扩展处理）。进行量子化和编码处理后，差别信号被传输到最终目标位置。在网络的接收端，整个过程正好相反。首先对差别信号进行取消量子化处理。然后将此差别信号添加到预报器中存储的样本信号中，发送给低通滤波器，重建原始输入信号。

DPCM 是降低语音传输的比特率的好办法。但是，它会带来一些其他问题，可能会影响语音质量。DPCM 对前一个样本输入信号与当前样本输入信号之间的差别进行量子化和编码处理。DPCM 使用统一量子化方法对差别信号进行量子化处理。统一量子化会生成一个 SNR，小输入样本信号的 SNR 较小，大输入样本信号的 SNR 则较大。因此，信号强度越高，语音质量就越好。这种情况的效率非常低，因为人的语音生成的大部分信号强度都很小。因此需要着重提高小强度信号的语音质量。自适应 DPCM 就是为了解决这个问题而开发的。

[自适应 DPCM](#)

自适应 DPCM (ADPCM) 是 ITU-T G.726 规范中定义的一种波形编码方法。

ADPCM 可以适应在 DPCM 过程中生成的差别信号的不同量子化级别。ADPCM 如何适应这些量子化级别？如果差别信号较低，ADPCM 就会增大量子化级别的大小。如果差别信号较高，ADPCM 就会减小量子化级别的大小。因此，ADPCM 会根据输入差别信号的大小调整量子化级别。这就会在差别信号的动态范围内生成一致的 SNR。使用 ADPCM 可以将语音传输的比特率降至 32 kbps，是 A-law 或 u-law PCM 的比特率的一半。与 A-law 或 u-law PCM 一样，ADPCM 生成的也是“长途电话质量”的语音。编码器必须带有反馈环路，使用编码器输出位来重新校正量子化装置。

[特定于 32 Kb/s 的步骤](#)

采纳为 ITU 标准 G.726。

- 将 A-law 或 Mu-law PCM 样本转换为线性 PCM 样本。
- 计算下一样本的预测值。
- 测量实际样本与预测值之间的差别。
- 将差别编码为四位数据，然后发送此数据。
- 将四位数据反馈给预报器。
- 将四位数据反馈给量子化装置。

[相关信息](#)

- [语音技术支持](#)
- [语音和统一通信产品支持](#)
- [Cisco IP 电话故障排除](#)
- [技术支持 - Cisco Systems](#)