

特。通过音频扬声器播放,这个数字信号则能够转换回来(也就是解码),形成一个模拟信号。然而,解码后的模拟信号仅是初始信号的近似,并且声音质量也许有明显的下降(例如,高频的声音可能在解码信号中丢失了)。通过增加采样速率和量化值的数量,解码信号能够更好地接近初始的模拟信号。因此(与视频一样),在解码信号的质量和比特率与数字信号存储空间之间存在一种折中。

我们刚才描述的基本编码技术称为脉冲编码调制(Pulse Code Modulation, PCM)。语音编码通常采用 PCM,采样速率为每秒 8000 个样本,每个样本用 8 比特表示,得到 64kbps 的速率。音频光盘(CD)也使用 PCM,采样速率为每秒 44 100 个样本,每个样本用 16 比特表示;这样使得单声道速率为 705.6kbps,立体声速率为 1.411Mbps。

然而,PCM 编码的语音和音乐很少在因特网中使用。与视频一样,取而代之的是使用压缩技术来减小流的比特速率。人类语音能被压缩到小于 10kbps 并仍然可懂。一种接近 CD 质量立体声音乐的流行压缩技术是 MPEG 1 第 3 层,更通常的叫法是 MP3。MP3 编码器通常能够压缩为许多不同的速率;128kbps 是最常用的编码速率,并且能够产生非常小的声音失真。一种相关的标准是高级音频编码(Advanced Audio Coding, AAC),该标准已经随苹果公司而流行起来。与视频一样,能够以不同的比特率生成多重版本的预先录制的音频流。

尽管音频比特率通常比视频的比特率小得多,但用户通常对音频的小失误比视频的小失误更为敏感。例如,考虑在因特网上举行的视频会议。如果视频信号时不时地丢失几秒,该视频会议很可能继续进行而没有太多的用户抱怨。然而,如果音频信号经常丢失,用户就可能不得不中止该会话。

7.1.3 多媒体网络应用的类型

因特网能够支持各种各样有用的和娱乐性的多媒体应用。在本小节中,我们将多媒体应用分为三个大类:①流式存储音频/视频;②会话式 IP 语音/视频;③流式实况音频/视频。如我们很快将看到的那样,这些应用类型中的每种都有自己的服务需求和设计问题的集合。

1. 流式存储音频和视频

为使讨论具体化,我们这里聚焦流式存储视频,它通常结合了视频和音频组件。流式存储音频(例如流式音乐)非常类似于流式存储视频,尽管它的比特率通常要低得多。

在这类应用中,依赖的媒体是预先录制的视频(如电影、电视节目)预先录制的体育赛事或预先录制的用户生成的视频(如常在 YouTube 上看到的那些)。这些预先录制的视频放置在服务器上,用户向服务器发送请求按需观看视频。许多因特网公司今天提供流式视频,包括 YouTube(谷歌)、Netflix 和 Hulu。据估计,在今天的因特网接入网中流式存储视频构成了超过 50% 的下载流量 [Cisco 2011]。流式存储视频具有三个关键的不同特色。

- 流。在流式存储视频应用中,客户开始从服务器接收文件几秒之后,通常就开始播放视频。这意味着当客户正在从视频的一个位置开始播放时,与此同时正在从服务器接收该视频的后续部分。这种技术被称为流(streaming),它避免了在开始播放之前必须下载整个视频(并且引起一个潜在的长时延)。
- 相互作用。因为媒体是预先录制的,用户可以对多媒体内容进行暂停、重新配置前进、重新配置倒退、快进等操作。从一个客户提出这种请求到该动作在客户端表现出来,可接受的响应时间应该小于几秒。