

照片的平均大小是 200KB。(与以前一样,我们在整个讨论中都简单地假定 1KB = 8000 比特。)第二位用户 Martha 正从因特网(“云中”)向她的智能手机流式传输音乐。我们假定 Martha 正在听许多 MP3 歌曲,一首接着一首,都以 128kbps 速率进行编码。第三位用户 Victor 则正在观看以 2Mbps 编码的视频。最后,我们假设所有三位用户的会话长度是 4000 秒(大约 67 分钟)。表 7-1 比较了这三位用户的比特率和传输的总字节。我们看到这时流式视频消耗了最多的带宽,其比特率比脸谱和流式音乐应用的带宽大 10 倍。因此,当设计网络视频应用时,我们心中必须记住的第一件事是视频的高比特率需求。鉴于视频的流行性及其高比特率,也许不会对思科公司以下的预测感到惊讶 [Cisco 2011]:到了 2015 年,流式视频和存储视频将大约占全球因特网流量消费的 90%。

表 7-1 三种因特网应用的比特率需求的比较

	比特率	67 分钟传输的字节
Frank 脸谱	160kbps	80MB
Martha 音乐	128kbps	64MB
Victor 视频	2Mbps	1GB

视频的另一种重要特点是它能被压缩,因而要在视频质量与比特率间进行折中。视频是一个图像序列,图像通常以恒定的速率显示,例如每秒 24 幅或 30 幅图像。一个没有压缩、数字编码的图像由像素阵列组成,每个像素被编码为一定数量的比特来表示亮度和颜色。在视频中有两种类型的冗余,它们都可以用来进行视频压缩(video compression)。空间冗余是给定图像的内部冗余。从直觉上讲,一个主要由空白组成的图像具有高度的冗余,能够有效地压缩而不会明显降低图像质量。时域冗余反映一幅图像和后续图像的重复程度。例如,如果一幅图像和后续图像完全一致,没有理由对后续图像再进行编码;相反,在编码过程中直接指出后续图像是完全一样的则更为有效。今天商用的压缩算法能够将视频压缩为所希望的任何基本比特率。当然,比特率越高,图像质量越好,总体用户视觉体验也越好。

我们也能够使用压缩来生成相同视频的多重版本(multiple version),每个版本有不同的质量等级。例如,我们能够使用压缩生成相同视频的三个版本,速率分别为 300kbps、1Mbps 和 3Mbps。用户则能够根据他们的当前可用带宽来决定要观看哪个版本。具有高速因特网连接的用户可以选择 3Mbps 的版本;使用 3G 智能手机观看视频的用户可以选择 300kbps 的版本。类似地,在视频会议应用中的视频能被“动态”(on-the-fly)地压缩,以在会话用户之间给定的可用端到端带宽上提供最好的视频质量。

7.1.2 音频的性质

数字音频(包括数字化语音和音乐)的带宽需求比视频低得多。然而,数字音频具有自己独特的性质,当设计多媒体应用时必须考虑这些性质。为了理解这些性质,我们首先考虑模拟音频(由人和乐器所产生)是如何转换为数字信号的:

- 模拟音频信号首先以某种固定速率采样,例如每秒 8000 个样本。每个采样值是一个任意的实数。
- 然后每个采样值被“四舍五入”为有限个数值中的一个。这种操作被称为量化(quantization)。这些有限个数值(称为量化值)通常是 2 的幂,例如 256 个量化值。
- 每个量化值由固定数量的比特表示。例如,如果有 256 个量化值,那么每个值(因此每个音频采样)用一个字节来表示。所有样本的比特表示级联在一起就形成了该信号的数字表示。举例来说,如果一个模拟信号以每秒 8000 个样值采样,而且每个样本被量化并用 8 比特表示,则得到的数字信号的速率就为每秒 64 000 比