

## 7.2 多媒体文件

在大多数系统中，普通的文本文件由字节的线性序列组成，没有操作系统了解或关心的任何结构。对于多媒体而言，情况就复杂多了。首先，视频与音频是完全不同的。它们由不同的设备捕获（视频为CCD芯片，音频为麦克风），具有不同的内部结构（视频每秒有25~30帧，音频每秒有44 100个样本），并且它们通过不同的设备来回放（视频为显示器，音频为扩音器）。

此外，大多数好莱坞电影现在针对的是全世界的观众，而这些观众大多不讲英语。这一情况有两种处理方法。对于某些国家，需要产生一个额外的声音轨迹，用当地语言进行配音，但是不包含音效。在日本，所有的电视都具有两个声道，电视观众看外国影片时可以听原声语言也可以听日语，遥控器上有一个按钮可以用来进行语言选择。在其他国家中，使用的是原始的声音轨迹，配以当地语言的字幕。

除此之外，许多在电视中播放的电影现在也提供英文字幕，使讲英语但是听力较弱的人可以观看。结果，数字电影实际上可能由多个文件组成：一个视频文件、多个音频文件以及多个包含各种语言字幕的文本文件。DVD能够存放至多32种语言的字幕文件。简单的一组多媒体文件如图7-3所示，我们将在本章后面解释快进和快倒的含义。

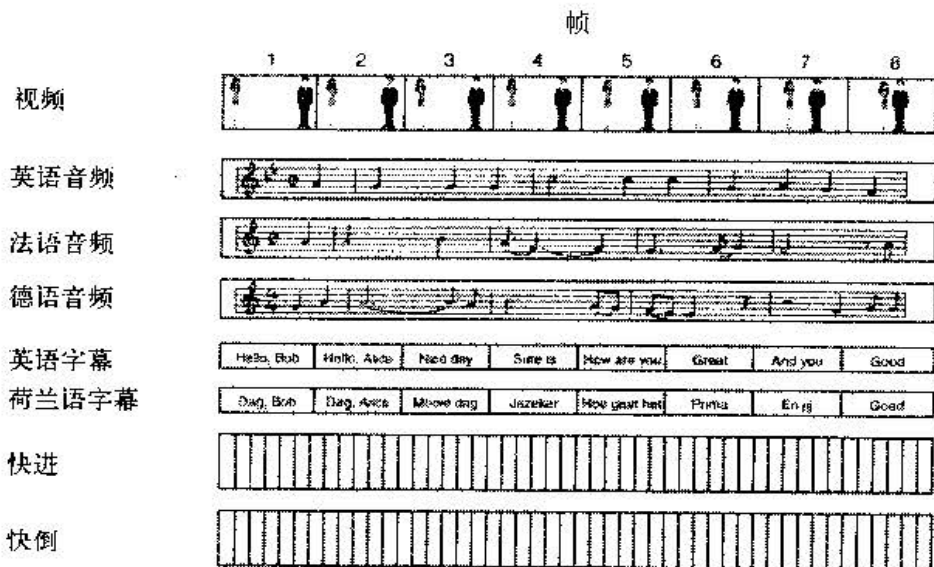


图7-3 电影可能由若干文件组成

因此，文件系统需要跟踪每个文件的多个“子文件”。一种可能的方案是像传统的文件一样管理每个子文件（例如，使用i节点来跟踪文件的块），并且要有一个新的数据结构列出每个多媒体文件的全部子文件。另一方法是创造一种二维的i节点，使每一列列出每个子文件的全部块。一般而言，其组织必须能够使观众观看电影时可以动态地选择使用哪个音频及字幕轨迹。

在各种情况下，还必须保持子文件同步的某种方法，这样才能保证当选中的音频轨迹回放时与视频保持同步。如果音频与视频存在即使是轻微的不同步，观众可能会在演员的嘴唇运动之前或之后才听到他说的话，这很容易察觉到，相当让人扫兴。

为了更好地理解多媒体文件是如何组织的，有必要在某种细节程度上了解数字音频与数字视频是如何工作的，下面我们将介绍这些主题。

### 7.2.1 视频编码

人类的眼睛具有这样的特性：当一幅图像闪现在视网膜上时，在它衰退之前将保持几毫秒的时间。如果一个图像序列以每秒50或更多张图像闪现，眼睛就不会注意到它看到的是不连续的图像。所有基于视频或影片胶片的运动图像系统都利用了这一原理产生活动的画面。

要想理解视频系统，最好从简单且过时的黑白电视开始。为了将二维图像表示为作为时间函数的一维电压，摄像机用一个电子束对图像进行横向扫描并缓慢地向下移动，记录下电子束经过处光的强度。在扫描的终点处，电子束折回，称为一帧（frame）。这一作为时间函数的光的强度以广播方式传播出去，