

## PDF文档与多媒体电子图书

李 慧(武汉大学传播与信息学院)

选自: 计算机应用 2000

PDF技术及其  
开发应用PDF工作流程  
系统PDF与书刊出  
版PDF在编辑工  
作中的应用PDF文档与多  
媒体电子图书PDF及其在电  
子出版领域的  
应用

PDF到印刷机

爱克发推出适  
用于高阶印刷  
的PDF生产工  
具—Apogee  
CreatCID字库在  
PDF流程中的  
应用黑版的制作与  
应用自己组合PDF  
工作流程

回首页

**摘 要** 本文介绍了PDF文档的结构与生成工具,讲述了PDF格式的多媒体电子图书的制作、管理和版权保护问题,并提出PDF格式的电子图书应成为图书馆的收藏目标之一。

**关键词** PDF文档 多媒体电子图书 电子图书

PDF(Portable Document Format结构化文档格式)Adobe公司(<http://www.Adobe.com>)开发的一种电子文档格式,它可以将文字、字型、格式、颜色、图形图像、超文本链接、声音、动态图像等信息封装在一个文件当中。PDF从页面描述语言PS(PostScript)发展而来,具有与PS几乎相同的页面描述能力和相似的描述方法。但与PS不同的是,PDF除了能描述复杂版面外,还具有交互功能(如超链接、交互表单等)、页面随机存取及字体仿真描述等特性。因此,PDF不仅适合印刷出版,而且也适合电子出版。PDF和目前在电子出版中占统治地位的HTML语言相比也有自己明显的优势。HTML不能描述版面,而PDF能描述精美的版面;HTML页面中经常出现信息的不一致性(如不同平台,不同浏览器,不同尺寸的浏览器窗口看到的同一web页面呈现出不同样子),而在PDF中不存在这个问题;HTML页面在进行保存时需要另外保存图片,而PDF文档的保存则可以一步到位。PDF不仅具有PS和HTML两者的优点,还克服了它们的缺点,从而实现了纸张印刷和电子出版的统一。排版后的内容生成PDF文件,则能在交付印刷的同时,进行网络发行。

### 1 PDF的结构

PDF文档是通过一系列对象序列来构造的,PDF对象包括直接对象(Direct Object)和间接对象(Indirect Object)。间接对象是经过标识了的PDF对象,直接对象通常有布尔型(Boolean)、数值型(Number)、字符串型(String)、名字型(Name)、数组型(Array)、字典型(Dictionary)、空对象(Null)、流对象(Stream)等基本类型。一个PDF文档包含了一个或多个页面,而每一页都可以包含文档、图形和图像。一个典型的PDF文件包含4个部分,分别是:文件头、文件主体、交叉引用表和文件尾。文件头(Header)位于PDF文档的第一行,它表明了当前文件所使用PDF规范的版本;文件体(Body)中包括构成文件间接对象的一个序列,这些对象代表了文档中的各个组件,例如字体、页树和图像等等。

交叉引用表(Cross-reference Table)则是为了能对间接对象进行随机存取而设立的一个间接对象地址索引表,它给出了所有当前文件使用对象的入口地址,即字节偏移量,以便系统随机访问不同的对象;文件尾(Trailer)中包含了交叉引用表的起始地址和其中的对象总数、文档中Catalog(目录)对象的对象号以及加密等安全信息,根据文件尾提供的信息PDF的应用程序可以找到交叉引用表和整个PDF文件的目录对象,从而控制整个PDF文档。

### 2 用PDF文档生成工具制作多媒体电子图书

由于PDF文档具有的种种优点,使它成为出版界的新宠,目前网上的电子出版物大多采用PDF文档格式来发行。PDF版本的电子图书可以保留图书的原版原式,图文并茂,既可以直接打印,也可以下载保存,并具有强大的全文检索功能。

制作PDF格式的多媒体电子图书目前主要有三种途径:第一种途径是应用专门的PDF制作工具软件进行制作;第二种途径是将包含有Postscript页面描述语言的文件转换成PDF文档;第三种途径就是利用转换工具把其他格式文件(如DOC、HTML等)转换成PDF文档。

Adobe公司推出的Adobe Acrobat 4.0中集成了几个PDF文档的开发工具。其中,Acrobat PDF writer可将任何格式的文档转换成PDF文档,它是一个模拟的“打印”驱动程序,即在客户应用程序(如word97)中调用打印命令,以“打印”文件的方式创建PDF格式的文档。PDF writer的使用方法是:首先安装该软件,安装完成后在Windows打印机设置的列标上将增加PDF writer;然后启动应用程序并打开用于转换的文件,在菜单或工具栏中选择“打印”;在打印机名称菜单上选择PDF writer;接着选择目标目录并为PDF文档取名;在填写完相应的信息后,PDF writer将会自动生成PDF文档。Acrobat Distiller是一个可以用来制作和编辑PDF格式文件的高端PDF文档生成器,利用它可以将Postscript文件转换成PDF文档。Acrobat PDF maker则是一个系列工具,例如Acrobat PDF maker Forward 97可在word97中直接将编辑好的DOC文件转换成PDF文档,转换过程中PDF maker会将word文件中的一些信息和结构相应的转换成PDF文档中的信息和结构,比如将word中的头部转换成PDF书签,将word中的URL转换成PDF的网页链接等等。

### 3 PDF格式多媒体电子图书的阅读与管理

目前PDF格式的多媒体电子图书多采用Acrobat Reader 4.0进行阅读。Acrobat Reader 4.0是一个免费软件,可直接从网上下载,其下载地址为<http://168.160.233.14/utility/ar40chs.exe>。该文件是4.0版本的简体中文版,文件大小为14MB。Acrobat Reader既可以独立运行,也可以作为浏览器的插件(plug-in)对PDF文档进行在线阅读。如果你的计算机中已经安装了Netscape浏览器,则在安装Acrobat Reader 4.0时安装程序会自动将其作为插件插到Netscape中;如果你所使用的是IE浏览器,则应该在程序安装完成后,将目录“\Adobe\Acrobat\4.0\Reader\browser”中的nppdf32.Dll文件拷贝到IE的插件目录“\Internet Explorer\Plugins\”中,并重新启动计算机,这样就可以在浏览器中像浏览HTML网页一样方便的浏览PDF文档了。

Acrobat Reader 4.0为用户提供了多种形式的浏览窗口,有“单页面”、“页面+书签目录”和“页面+书页图标”等显示方式,即可以全屏浏览,也可以单页浏览,它还提供了强大的翻页功能,可利用菜单、滚动条、右键菜单、方向键等进行翻页,可选择首尾、前后及指定页号翻页;Acrobat Reader 4.0的页面多级缩放功能使用户可以按12.5%到1600%的比率进行缩放,以选择合适的大小进行阅读;同时它还提供了复制的功能,点击工具栏中含“T”字样的“TextSelectTool”按钮,通过拖拉鼠标选定欲复制的文本内容,就可以进行复制了。需注意的是,如果在制作PDF文档时设定了“不允许选择文本和图形”的安全保护,则不能够复制文本。

Acrobat Reader 4.0是我们阅读PDF格式的多媒体电子图书的较好选择,但如果我们欲对其中的内容进行编辑和修改则应该使用Adobe公司的另一产品Adobe Exchange,利用它我们可以在文档中添加超文本链接、书签、小书页图标、表等,也可以在文档中插入、删除、复制或移动一个或多个页面。

目前国内许多图书馆都开辟了自己的多媒体阅览室,笔者认为PDF格式的多媒体电子图书应成为其中主要的图书来源之一。图书馆可以从网络上下载或者进行购买,对其进行著录和标引之后组织成电子书库提供给读者,也可以将自己的馆藏图书采用图像扫描的方式制作成PDF格式的电子出版物,抑或是将已经电子化了了的普通格式的电子图书转换成PDF格式,充分发挥PDF文档体积小、易于传输与储存的特点,为读者提供大量的、内容丰富的、图文并茂的电子书籍。总之,在信息技术飞速发展的今天,纸制图书应该逐渐退出其主导地位,图书馆应该更多的将目光投注在电子图书之上,突破传统的图书组织方法和模式,将电子图书的收集、整理与组织纳入日常工作之中。PDF格式的电子图书因其应用的广泛性应该成为电子图书的主流。

#### 4 版权问题

由于PDF文档易于传输,且在获得了免费阅读软件后可进行随意浏览,给PDF格式电子图书的版权保护带来了极大的挑战。针对这一问题,Adobe公司推出了面向出版商和发行商的安全出版系统。该系统由两部分组成:Adobe PDF Merchant和Acrobat Reader的Web Buy插件。PDF Merchant是作为服务器端的程序,Web Buy插件作为客户端程序,通过两者配合工作来确保网上出版物免遭盗版。PDF Merchant为准备传送的PDF文档加“锁”,当读者试图利用Acrobat Reader阅读上了“锁”的文件时就会看到一个说你无权阅读该文件内容的信息,并询问你是否愿意购买文件内容,在你点击购买之后,PDFMerchant就会直接将你带到出版商或发行商的站点上,完成购买后,网站就会配合Acrobat Reader将文件内容和阅读许可发送的你的计算机上。

当然,该系统只是针对出版商和发行商的,图书馆该如何保护自己的权益、堵住盗版的漏洞,目前还没有现成的解决方案,但是上述的安全系统为我们提供了思路。图书馆也可以仿照其方法,为PDF格式的电子图书加“锁”,当网络用户将其下载并进行阅读时,要求用户输入相关信息(如借书证的号码等)并承诺不将此书用于商业用途,然后将内容显示出来。

#### 5 结束语

PDF文档在电子文档的发行和数字化传播方面大行其道并不是偶然的,而是有一系列的优势作为其后盾的。随着因特网上越来越多PDF文档的出现,其范围已不再局限于电子图书,许多产品说明、电子报刊、公司广告、网络资料以及电子邮件都被制作成PDF文档,因此,图书馆应该对这一新事物给予充分重视,将之纳入自己的信息资源体系中来,探讨其搜集、组织、管理及提供利用的方式、方法,以更好的为读者服务。

#### 参考文献

- 1 杨道良.面向对象的中文PDF阅读器的设计与实现.计算机应用,1999
- 2 徐 鹏,等.多国语言PDF软件包的设计-Java技术在电子商务领域的应用.微型机与应用,1999(2)
- 3 肖 明.PDF文档的阅读与生成工具.电子与电脑,2000(1)(责任编辑:滕代娣)