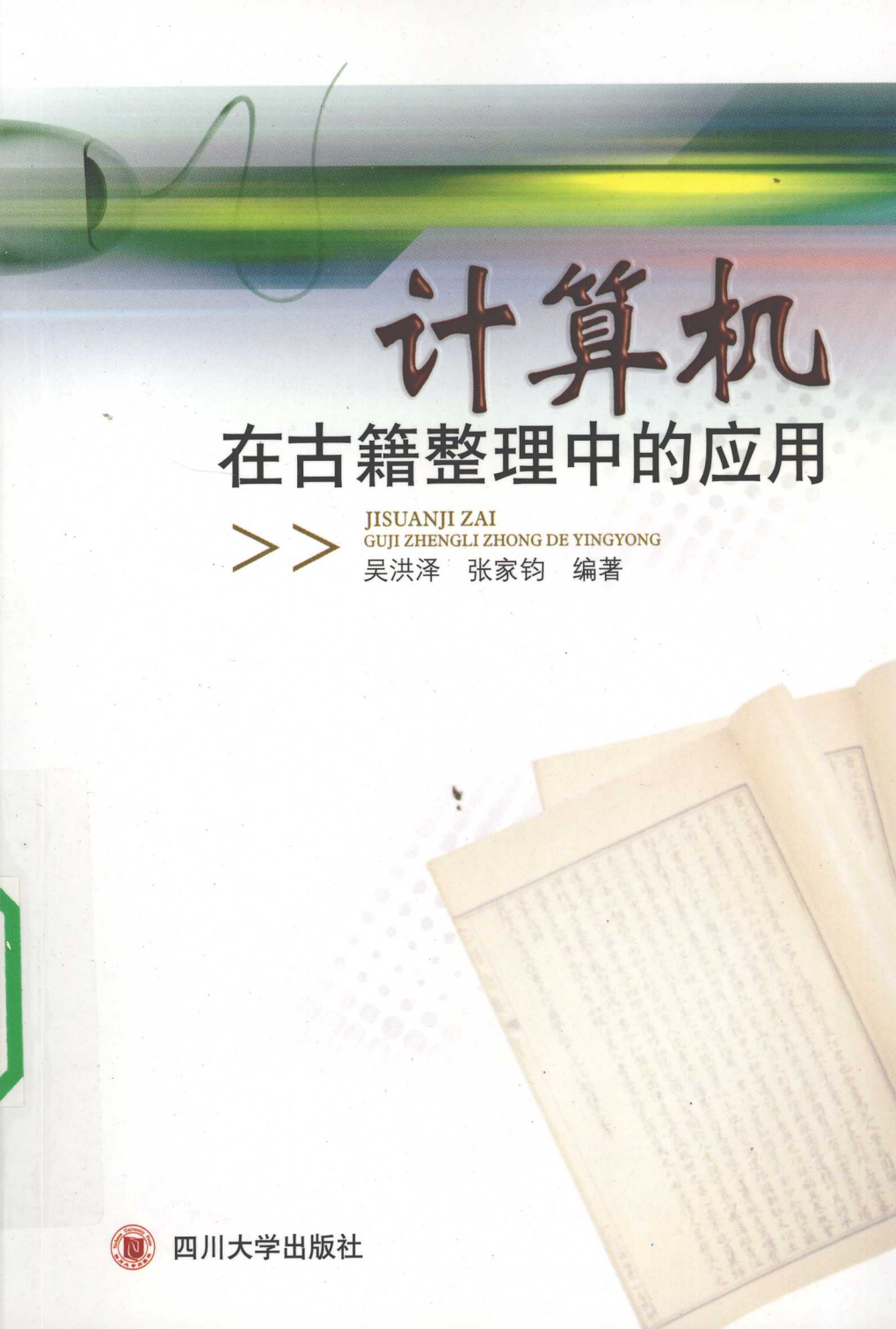




计算机

在古籍整理中的应用



JISUANJI ZAI
GUJI ZHENGLI ZHONG DE YINGYONG

吴洪泽 张家钧 编著

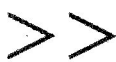


四川大学出版社

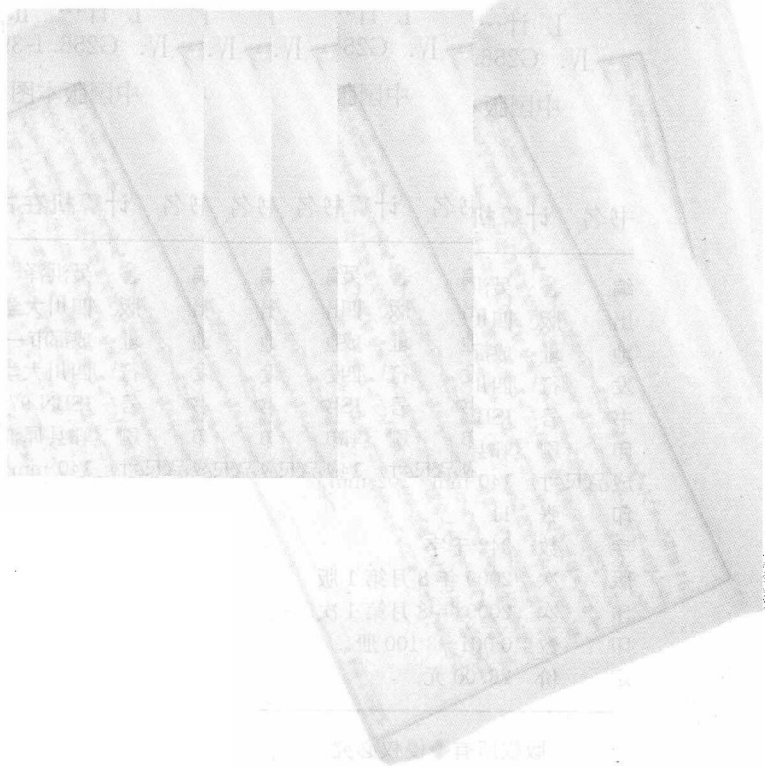
教育部人文社会科学研究“十一五”规划项目资助 批准号: 01JA870003

计算机

在古籍整理中的应用



JISUANJI ZAI
GUJI ZHENGLI ZHONG DE YINGYONG
吴洪泽 张家钧 编著



四川大学出版社

责任编辑:曾 鑫
责任校对:廖庆扬
封面设计:墨创文化
责任印制:李 平

图书在版编目(CIP)数据

计算机在古籍整理中的应用/吴洪泽, 张家钧编著.
成都: 四川大学出版社, 2009. 7
ISBN 978-7-5614-4508-2

I. 计… II. ①吴…②张… III. 计算机应用—古籍整理
IV. G256.1-39

中国版本图书馆 CIP 数据核字 (2009) 第 134156 号

书名 计算机在古籍整理中的应用

编 著	吴洪泽 张家钧
出 版	四川大学出版社
地 址	成都市一环路南一段 24 号 (610065)
发 行	四川大学出版社
书 号	ISBN 978-7-5614-4508-2
印 刷	郫县犀浦印刷厂
成品尺寸	140 mm×202 mm
印 张	11
字 数	242 千字
版 次	2009 年 8 月第 1 版
印 次	2009 年 8 月第 1 次印刷
印 数	0 001~3 100 册
定 价	26.00 元

◆读者邮购本书,请与本社发行科
联系。电话:85408408/85401670/
85408023 邮政编码:610065
◆本社图书如有印装质量问题,请
寄回出版社调换。

版权所有◆侵权必究

◆网址:www.scupress.com.cn

前 言

古籍，顾名思义，就是中国古代书籍的简称。古代的时间下限，一般认为是在 1911 年辛亥革命推翻清朝为止。尽管还有争议，但这种观点已为多数学者所接受。因此，从狭义的角度看，古籍就是指 1911 年前成书的图书。如果从广义的角度考察，则“古籍包括古书和古书以外的、未形成‘书’的其他古代文献，如甲骨刻辞、金石刻辞、简牍帛书、敦煌卷子等等”^①。所谓古籍整理，也有广义和狭义两方面的理解。大凡有关古籍各方面的学术工作，都可以称之为古籍整理，这是广义的，也是为大多数人所接受的观点。而狭义的、严格意义上的古籍整理，“就是对古籍的原文进行某种形式的整理加工，以便于人们阅读与研究”^②。本书所涉及的，不论是“古籍”还是“古籍整理”的概念，都是广义的，也就是针对产生在 1911 年以前的历史文献，包括纸本的以及其他形式的载体，人们利用计算机对这些历史文献的一切处理形式，都属于本书考察的范围。

古籍资源有多少？目前没有明确的统计，据《中国图书联合书目》统计约有 10 万余种，这一数字得到学术界普遍的认同。这些图书大多收藏于图书馆、博物馆中，一些珍贵的善本古籍，

① 刘琳《古籍整理学》，四川大学出版社，2003 年版，第 2 页。

② 同上。

收藏单位出于珍爱古籍的目的，使一般的读者很难接触到原本，即使有复制善本的缩微胶片，也不是轻易可以看到的。因此，保护与使用，就形成了一对矛盾。如何化解这个矛盾呢？就目前而言，没有比古籍数字化更好的解决方案了。

以全文检索对照原本图像的方式进行数字化的古籍，既可以起到保护古籍原本的目的，使其束之高阁，长久保存，尽量减少人为的损耗，又可以让使用者通过影像看到原本的面貌，通过全文版实现快捷方便的检索，可谓一举两得。同时，古籍整理专家应用计算机研究出的成果，本身也是数字化产品，可通过图书、光盘、网络等多种便捷的成果发布平台，以最快的速度让自己的科研成果与世人共享，从而避免了以前出一部书要等上十年八载的尴尬局面。所以，有人说，数字化是新形势下整理古籍的一种新方式，可谓一语中的。

自 20 世纪 80 年代以来，计算机在古籍整理中的应用越来越多，综合来看，最突出地体现在以下 3 个方面：一是古籍全文检索软件的开发，如《国学宝典》、《汉籍全文检索系统》、《文渊阁四库全书电子版》、《四部丛刊电子版》、《古今图书集成电子检索版》、《中国古籍基本光盘库》等，影响巨大，拥有大量的使用者；二是古籍处理基础软件的开发，如北京书同文公司开发的古籍数字化软件，可解决古籍扫描、汉字自动识别及校对等问题；三是古籍出版现代化与电子版古籍、古籍资源网络化的大爆发，把古籍资源的可利用度以及古籍整理研究成果的数量，推向了空前的高度。

数字化产品开发出来了，但使用者不一定用得上，这是购买力的问题，需要有关部门协同解决；而同样的产品，不同的人使

用，效果却不尽相同，这就是使用方法的问题了，可以通过学习来解决。

现在的年轻人，都是使用电脑的能手，但从事古籍整理的中老年专家学者，往往既没有上过专门的计算机课，甚至也没有经过短暂的培训班学习，虽然能用电脑编排古籍，但在使用电脑方面，属于入门级，包括本人在内，别说程序设计与编程，就是对计算机应用的诸多问题也不甚了然。一个个系统软件摆在面前，如何利用它去整理古籍，对我们来说也显得高深莫测。为此，在曾枣庄老师的建议下，从十多年前起，我开始进行利用计算机辅助古籍整理的尝试，学习各种应用软件的使用方法，其间也深深感到：软件研制固然不易，而要完全掌握一种系统软件的所有功能，并巧妙地加以运用，又谈何容易呢？比如说我们用得较多的方正系统，在对影印古籍的图像文件进行排版时，不少人是用维思、飞图等软件，一幅幅地修描、排版、添加页码。我们当初在制作《诸子集成补编》系列图书时，也是用维思软件做的，结果不是图像不正，就是单双页页码不能反扣，被印刷厂定为废品，只得重做。后来，我们利用书版软件的图片处理功能，将所有图像组成一个排版文件，既省时省力，又大大提高了质量，原本计划 10 个月完工，结果 5 个月就大功告成了。可见，软件的应用功能，有时也需要用户去“开发”的。

正是因为在实际工作中不时有这样的小发现，我得到了老师们的信任，他们要求我将操作方法一步步地写出来，以供古籍整理专家们借鉴。为此，我申报了 2001 年度的教育部规划选题“计算机在历史文献信息化处理中的应用研究”，获得批准（批准号：01JA870003）。当时的软件，大多是在 DOS 环境下开

发的，进入 Windows 时代后，针对软件应用所写的篇章，许多只得重写。但是因为连续承担大课题的缘故，始终没能写成，只在 2003 年得到刘琳老师的厚爱，写成《古籍整理手段的现代化》一章，植入他的大著《古籍整理学》中。大课题不断地进行，而我在计算机方面的努力，也从数据库转为古籍书版制作与索引编制，对方兴未艾的古籍数字化问题，反而没有太多的精力去探讨了。课题研究屡兴屡废，始终没能完稿，深感惭愧！从 1992 年到现在，经历过从 DOS 到 Windows 的时代，感觉如同走马观花一般，记忆之城中，停留的尽是过眼烟云，计算机世界的变化实在太快了！

本书的最终完成，得益于曾枣庄、刘琳老师多年的鼓励，所以这本书是为像他们这样的专家学者写的，取其可资借鉴、便于操作而已。书中介绍的一些程序，得益于软件开发者的说明书；个别问题的探讨，参考了同行们的科研成果，在此深致谢意！本书的出版，四川大学出版社的曾鑫、庄剑先生付出了辛勤的劳动；吴洪锋、卢本丽代撰《计算机应用基础》一章，张家钧撰写了《网络古籍资源述略》一章，谨此致谢！

吴洪泽

2009 年 6 月 5 日

目 录

第一章 计算机应用基础	(1)
第一节 利用计算机辅助古籍整理的条件	(1)
第二节 计算机应用常识	(3)
一、计算机发展简史	(4)
二、微型计算机的构成	(4)
三、微机操作常识	(8)
四、Windows 操作技巧	(8)
(一) 程序安装	(8)
(二) 程序调用	(10)
(三) 常用快捷键	(12)
(四) 系统优化	(13)
(五) 系统备份	(16)
(六) 系统恢复	(20)
第二章 计算机在古籍整理中的应用	(25)
第一节 计算机与古籍整理的关系	(25)
一、古籍整理手段的更新势在必行	(27)
二、计算机应用在古籍整理中的优越性	(30)
第二节 计算机在古籍整理中的应用现状	(32)
第三章 古籍数字化	(37)
第一节 古籍数字化的历程	(37)

一、古籍数字化的几个阶段	(37)
二、古籍数字化概念的形成	(40)
第二节 古籍数字化述略	(42)
一、古籍数字化目的与原则	(42)
二、古籍数字化的基本方法	(44)
三、古籍数字化的现状与发展趋势	(48)
第三节 常用古籍数字化成果举隅	(50)
(一) 文渊阁四库全书电子版	(51)
(二) 《四部丛刊》电子版	(52)
(三) “中国基本古籍库”光盘工程	(52)
(四) 《国学宝典》	(53)
(五) 《中国历代基本典籍库》系列光盘	(54)
(六) 《汉籍全文检索系统》	(54)
(七) 《古今图书集成》全文电子版	(55)
(八) 中国教育科技数字图书馆	(57)
(九) 《龙语瀚堂典籍数据库》	(58)
(十) 《全宋诗》分析系统	(58)
(十一) 汉籍电子文献	(59)
(十二) 中华典籍网络资料库	(60)
(十三) 简帛金石资料库	(60)
第四章 汉字库及相关问题	(62)
第一节 古籍数字化的字库问题	(62)
第二节 汉字规范	(64)
一、异体字处理	(67)

(一) 异体字的定义	(67)
(二) 异体字规范	(68)
(三) 异体字处理的现状	(69)
二、繁简字处理	(71)
(一) 繁到简转换	(71)
(二) 简到繁转换	(77)
(三) 检查不规范繁简字	(93)
第三节 补 字	(96)
一、TrueType 造字程序	(96)
(一) 调用 TrueType 造字程序	(97)
(二) TrueType 造字	(98)
(三) 使用造字	(100)
二、女娲补字软件	(100)
(一) 女娲补字程序安装	(101)
(二) 启动女娲补字程序	(101)
(三) 造 字	(103)
(四) 补字编码区	(105)
(五) 转换补字	(106)
(六) 调用补字	(107)
(七) 备份补字库	(107)
第四节 汉字输入	(108)
第五章 古籍文本的编辑	(114)
第一节 Word 2003	(115)
一、Word 2003 应用基础	(115)

(一) Word 2003 的启动	(115)
(二) Word 2003 的工作窗口	(117)
二、Word 2003 文本编辑	(118)
(一) 建立文档	(118)
(二) 打开文档	(119)
(三) 文档管理	(120)
(四) 文档编辑	(121)
三、Word 2003 文档排版	(125)
(一) 字体与排法	(125)
(二) 页眉、页脚设置	(127)
(三) 页码设置	(128)
四、Word 2003 文档打印	(129)
(一) 页面设置	(129)
(二) 文档打印	(129)
第二节 方正书版 9.11	(129)
一、方正书版9.11应用基础	(130)
(一) 方正书版9.11的特点	(130)
(二) 方正书版9.11的文件系统	(131)
(三) 方正书版系统的工作流程	(132)
(四) 方正书版9.11的安装	(133)
二、编辑小样文件	(133)
(一) 方正书版9.11的启动与退出	(133)
(二) 方正书版编辑窗口	(134)
(三) 创建小样文件	(135)
(四) 打开小样文件	(136)

(五) 小样文件保存	(137)
(六) 设置编辑参数	(139)
(七) 文字输入	(140)
(八) 插入与改写	(140)
(九) 选择文字块	(141)
(十) 查找文字	(141)
(十一) 替换文字	(143)
(十二) 替换回车符	(144)
(十三) 删除文字	(144)
(十四) 移动文字	(145)
(十五) 复制文字	(145)
(十六) 消除误操作	(145)
三、多文件编辑	(146)
(一) 同时打开多个文件	(146)
(二) 合并文件	(147)
(三) 拆分文件	(148)
(四) 根据内容查找文件	(148)
四、符号输入	(150)
(一) 特殊符号输入	(150)
(二) 动态键盘输入	(151)
(三) 插入符号命令	(154)
(四) 字符拼合	(155)
第六章 古籍书版的制作	(157)
第一节 书版排版基础	(157)
一、BD 排版注解	(157)

(一) 使用 BD 排版注解须知	(157)
(二) 排版符号释义	(158)
二、BD 排版帮助	(159)
(一) 方正书版软件应用技巧	(159)
(二) 利用提示功能输入 BD 注解	(160)
(三) 利用模板功能输入 BD 注解	(161)
第二节 字体、字号的排法	(162)
一、汉字字体字号	(162)
(一) 汉体注解	(162)
(二) 使用外挂字体	(166)
(三) 繁简字转换注解	(166)
二、排数字及拼音	(167)
(一) 排数字	(167)
(二) 排外文	(168)
(三) 排拼音	(169)
三、变形字体的排法	(170)
(一) 排长扁字	(170)
(二) 文字倾斜	(171)
(三) 旋转文字	(171)
(四) 文字加粗或减细	(172)
四、字距的调整	(172)
(一) 空格注解	(172)
(二) 撑满注解	(173)
(三) 紧排注解	(173)
五、中文标点符号的排法	(174)

(一) 标符注解	(174)
(二) 对开注解	(175)
(三) 全身注解	(175)
第三节 版面控制	(176)
一、排版参数设置	(176)
(一) PRO 文件的建立	(176)
(二) 定义版心参数	(179)
二、多文件组合排版	(181)
三、定义页码	(182)
(一) 页码注解	(182)
(二) 页号注解	(183)
(三) 排暗码	(185)
(四) 不排页码	(185)
四、书眉的排法	(185)
(一) 定义书眉格式	(186)
(二) 书眉的排法	(187)
五、边文与边栏	(188)
(一) 边文注解	(188)
(二) 背景注解	(190)
(三) 边栏、边注注解	(191)
六、注文排版	(192)
(一) 设置注文格式	(192)
(二) 脚注的排法	(194)
(三) 割注注解	(196)
七、标题排版	(197)

(一) 定义标题格式	(197)
(二) 标题的排法	(198)
八、强制换页的方法	(201)
第四节 正文排版	(202)
一、位置控制	(202)
(一) 指定排版起始位置	(202)
(二) 文字居中	(203)
(三) 文字居右	(204)
(四) 左边缩进	(205)
(五) 右边缩进	(206)
(六) 左右同时缩进	(207)
(七) 字符对齐	(207)
(八) 上下注解	(208)
二、行处理	(209)
(一) 行宽控制	(209)
(二) 空行的方法	(210)
(三) 调整行距	(211)
三、保持行距一致	(211)
四、上下移动文字	(212)
(一) 行中并列多行内容	(213)
(二) 消除单字行	(214)
(三) 多行文字撑满对齐	(214)
五、块处理	(215)
(一) 段首处理	(215)
(二) 段首缩进	(215)

(三) 防止拆页的方法	(216)
(四) 指定不排版内容	(216)
第五节 划分版面	(216)
一、分栏的排法	(216)
(一) 分栏注解	(217)
(二) 另栏注解	(217)
(三) 对照注解	(218)
二、划分排版区域	(218)
(一) 方框注解	(219)
(二) 分区注解	(220)
(三) 多页分区注解	(221)
第六节 图表框线排版	(222)
一、插图排版	(222)
(一) 图片注解	(223)
(二) 图说注解	(224)
(三) 插入随文图片注解	(224)
(四) 插入 EPS 注解 (PS)	(226)
(五) 插入注解 (CR)	(226)
二、框线与底纹	(227)
(一) 长度注解	(227)
(二) 画线注解	(228)
(三) 斜线注解	(229)
(四) 线号注解	(229)
(五) 着重注解	(230)
(六) 排底纹	(230)

(七) 排附加符号	(231)
三、表格排版	(232)
(一) 表格注解	(234)
(二) 表格内容排版	(236)
(三) 无线表注解	(237)
(四) 排族谱世表	(238)
第七节 自动排目录	(239)
(一) 目录定义注解	(239)
(二) 自动目录登记注解	(240)
(三) 自动生成目录	(241)
第八节 图书索引编制	(244)
一、用方正书版系统排索引	(245)
(一) 索引点注解	(245)
(二) 用排序功能排索引	(249)
二、排四角号码索引	(250)
(一) 中文索引编制系统	(250)
(二) 自制四角号码索引	(252)
第七章 影印古籍	(253)
第一节 影印古籍的方法	(253)
一、古籍影印简介	(253)
二、古籍影印的新技术	(255)
三、古籍影印的方式	(255)
第二节 古籍影像处理	(257)
一、古籍图像化	(257)