

# 汉字信息处理

报版样张

1979 年 7 月 1 日

本刊是计算机—激光汉字编辑排版系统的试排样张。

由计算机总局主持，北京大学、新华社、山东省电子局、潍坊市电子局、潍坊电讯仪表厂、杭州五二二厂、天津红星厂等单位协作会战

## 计算机—激光汉字编辑排版系统主体工程研制成功

### 汉字编辑排版系统的工作流程和软件

**一、汉字怎样进入计算机** 本系统将采用两种汉字键盘。一种是大键盘，共四百键，每键九个字，可以脱机也可以联机。当联机使用时，每个键盘有一台汉字显示器，每击一键产生十四位代码直接进入计算机，并立即在显示器上显出这一汉字字形。另一种是中键盘，共二百五十六键，每键代表两个或三个字符（整字或组字部件），每个汉字由若干字符组成。输入一个汉字实际平均按键不超过三键。

**二、输出供校对用的小样** 输入计算机的文章经软件处理后在汉字印字机上输出小样，小样的汉字文字质量不高，只供校对修改用。

**三、通过汉字显示器联机修改文章** 校对人员可通过显示器修改文章。键盘上设有几十个功能键，可以方便地

增、删、改，或把文章的某一段移到另一段中去，修改后的结果立即在显示器上显示出来。有十台汉字显示器可以供十个校对人员同时使用。

**四、通过版面显示器作版面设计** 版面显示器能显示相当于一版参考消息的版面情况，编辑人员可以打入修改版面设计的命令，在软

件控制下很快在显示器上显示出修改的情况，一直到编辑人员满意为止。

**五、软件把输入的文章组成版面，形成照排信息，控制激光照排机工作** 软件分析输入的排版要求，确定每个汉字在版面上的位置形成照排信息，照排控制器按照排信息从字模存储器中取出所需的汉字

字模压缩信息，微程序汉字点阵生成器再把压缩信息复原成汉字点阵，提供给光调制器，控制激光扫描打点或不打点，在底片上形成所需要的版面。

上述全部工作都是在软件控制下进行的，整个软件系统包括分时操作系统、排版编辑程序、命令处理程序、终端程序等。

### 滚筒式激光照排机的工作原理

本系统的输出设备为滚筒式激光照排机。它利用激光束在底片上扫描打点，将点阵化的汉字按版面要求排出，排版幅面相当于两版《参考消息》的版面。排出一版所需的时间为一分四十秒。扫描线密度为每毫米二十九线。

滚筒式激光照排机的激光车上有四支激光管发出的四束光分别通过四个光调制器，再分别由四个聚光物镜会聚在四支光纤纤维端面上。通过光纤纤维的光由横移拖板上的物镜会聚在底

片上。在排版时主马达带动滚筒以均匀的速度转动。横移拖板以均匀的速度移动。使扫描头在底片上产生螺旋形的扫描线，底片旋转一周。横移拖板带动扫描头在底片上横向移过四条扫描线的宽度。照排控制器按编辑排版的要求，将汉字点阵信息以四行并行输出的方式，用四路信号分别控制四个光调制器按文字点阵使光束打开或关上，这样就通过扫描头在底片上打点扫描出版面。

实现汉字数字化存贮首先遇到的难题是“三多”：汉字的字数多，字体多，字号多。从印刷要求看，不仅要收七千左右的字，还得有各种不同的字体。从一般书报来说，字体就有书版宋、报版宋、标题宋、仿宋、楷体、黑体、长宋、扁宋、长黑、扁黑等十多种；字号就有小六号、六号直到初号，共十五种。

其次是文字质量问题，要保证字型美观、笔锋明晰。这就要求字模的点阵密度高，报版字的密度应在一毫米二十点以上，书版字的密度应在一毫米二十五点以上。我们所做的分解密度为一毫米二十九点，这就是说一个五号字，要由一万个点组成。在上述条件下如以小字号（小六号到三号）每种收七千字，大字号（小二号到特号）每种收四千字，总共要有六十五万个字头，这就需要二百亿位的存贮量。

为了减少汉字字模的存贮量，日本采用的办法是记录有笔划的黑段的起始位置和长度，这样可节省近一半的存贮量。日本经济新闻采用这种办法，用几十亿位的大磁盘只存贮了十多万个汉字字头。这远远满足不了我们的要求。而且大容量磁盘体积大，价格较贵，存取速度慢。能否使汉字信息大量压缩是整个系统的关键所在。日本电气公司和日本京都大学在七十年代初就进行这方面的研制，他们研制的汉字信息压缩技术虽然压缩倍数很高，但文字质量不好，因此未投入使用。

我们研制成功了一种确保文字质量的高倍数汉字信息压缩技术，可使每个五号字的信息量下降十二倍，即从一万位下降到平均八百位。这种压缩方法还允许文字变倍，能变大变小并保证质量。这样整体压缩倍数高达五百倍，只需四千万位的存贮量就能存下六十五万字头的全部信息。在照排过程中，有一个微程序汉字点阵生成器把汉字压缩信息高速复原成点阵。汉字信息压缩技术的研制成功给汉字照排系统开辟了新的途径，使得高速、廉价的先进汉字照排机成为可能。

### 第四代排字机

激光和计算机的结合给排版系统和信息处理带来新的突破

一九五八年美国发明了光机式西文照相排字机，在计算机控制下依靠光学和机械方法选取字模，逐字照相形成版面，这称为第二代排字机，六十年代在美国大量生产。由于汉字字数很多，照排技术要求更复杂得多，所以直到七〇年左右，日本才研制成功第二代汉字排字机，目前已经开始推广使用。一九六五年西德研制成功阴极射线管输出的第三代排字机，各种字模以数字化二进位的形式存放在计算机的外存贮器——磁盘中，

用扫描打点的方式在阴极射线管上显示出高质量的文字后照相。三代机的速度可以比二代机快十倍，适用范围广，七十年代已在美国和西德大量投产。日本引进西德和美国的技术于一九七五年初步研制成第三代汉字排字机，一九七八年已有少数几台在使用，估计今后几年内将逐步成熟和普及，三代机对底片要求高，对所使用的阴极射线管分因率也要求很高，这些困难在我国目前的技术条件下不易很快解决。第四代排字机采用

激光技术，用激光束在底片上扫描打点形成版面。由于激光很强，普通底片便能感光，还可利用转印的方式在普通纸上印出逼真的大样。更有意义的是用三到五瓦的强激光束直接打到版材上就能形成凹凸版面，免除了照相制版一大套工序，称为激光直接雕版，是今后的发展方向。目前世界上只有一家公司开始少量生产第四代西文排字机。日本正在研制第四代汉字排字机，但尚未成功。这里有一系列的困难需要克服，例如字模

存贮问题、文字变倍问题、激光只能逐线扫描不能逐字扫描的问题，等等。我们采取了一系列特别的方法，圆满地解决了所有困难。这标志着我国计算机编辑排版系统和汉字信息处理技术达到了一个新的水平。

名 称		研制成功的年代和国家	特 点
第一代	手动照排机	西文：一九四九年美国	效率很低，改版困难
第二代	光 机 式	西文：一九五八年美国 汉字：一九六九年日本	机械动作多，速度低 适用面窄，不易维修
第三代	阴极射线管	西文：一九六五年西德 汉字：一九七五年日本	速度快，对底片要求 高，不能出大样
第四代	激 光 照 排	西文：一九七六年英国 汉字：日本正在研制	速度快，适用面广， 还能直接雕版

[github.com/chenshuo/typeset](https://github.com/chenshuo/typeset)

<https://wangxuan.pku.edu.cn/>

汉字字模信息的存贮