



## 第7章 汉字字形管理技术（二）



## 内容

---

- ❖ 字形的放大和缩小
- ❖ **TTF**字形技术
- ❖ 字库的设计和管理技术
- ❖ 汉字字库的应用



## 4. 字形的放大和缩小

---

- ❖ 汉字的字号
- ❖ 字形放大和缩小的基本原理
- ❖ 防止字形放大和缩小失真的措施
- ❖ 字形放大缩小的应用



## 4.1 汉字的字号

- ❖ 在计算机汉字照排系统中，至少容纳了十六种基本的字体。

字号	磅数	字心点阵	比例关系
七	6	56 × 56	小四号的一半
小六	7	64 × 64	四号的一半
六	7.875	72 × 72	三号的一半
小五	9	82 × 82	小二号的一半
五	10.5	96 × 96	六种字体的基本字号
小四	12	110 × 110	五号的1.146倍
四	14	128 × 128	五号的1.333.倍，大字体的基本字号
三	15.75	144 × 144	五号的1.5倍
小二	18	164 × 164	五号的1.708倍
二	21	192 × 192	五号的2倍，或四号的1.5倍
一	28	256 × 256	四号的2倍，或接近五号的2.5倍
小初	31.5	288 × 288	五号的3倍
初	35	320 × 320	四号的2.5倍
小特	42	384 × 384	四号的3倍
特	49	448 × 448	四号的3.5倍
特大	56	512 × 512	四号的4倍

注：字心点阵不包括字间距，比字身点阵略小，例如，五号字的字心点阵为96 × 96，字身点阵为108 × 108。



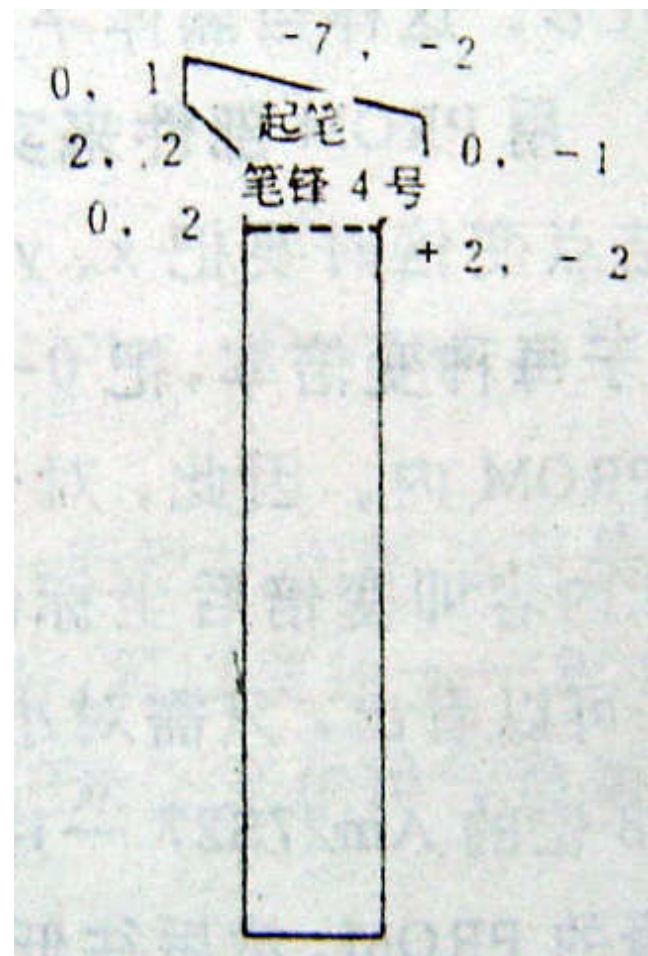
## 4.1 汉字的字号

- ❖ 书版宋体、报版宋体、标题宋体、仿宋体、黑体和楷体这六种字体，均以五号字为基本字号。
- ❖ 字模存储器中只存放五号字的压缩信息，其余字号都由五号字变倍而得。
- ❖ 基本字体还可以通过拉长和压扁的变倍方法变化出各种长字体和扁字体。



## 4.2 基本原理

- ❖ 字形在放大和缩小时，也分为规则笔画和不规则笔画这两种情况来处理。
- ❖ 可以编成号码，这些编号的笔锋是以精细的轮廓折线形式存在内存中的，如右图所示。





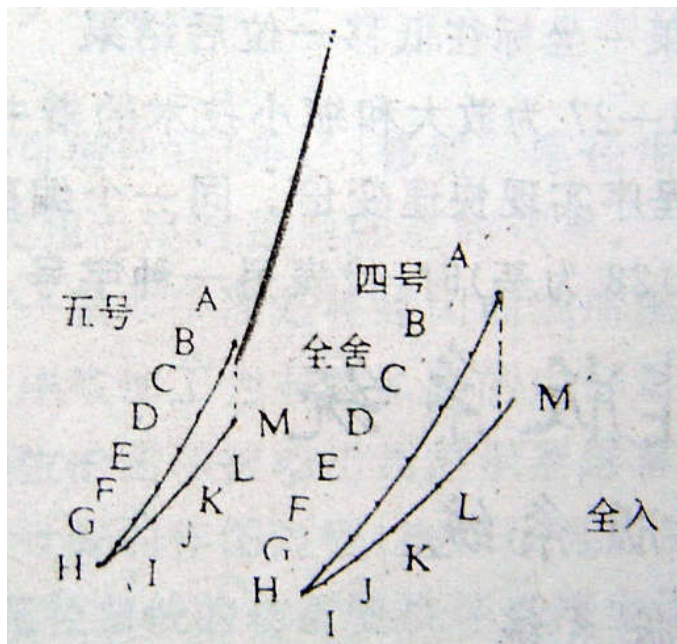
## 4.2 基本原理

- ❖ 无论使规则笔画还是不规则笔画，都是用轮廓折线表示的，它们很容易被放大缩小。
  - ✧ 例如，要使得图形放大 $r$ 倍（ $r$ 不一定是整数），只需把对应的每段轮廓折线放大 $r$ 倍，也即对每条折线的 $\Delta x$ 、 $\Delta y$ 值都乘以 $r$ 。
- ❖ 要得到高质量的放大或缩小的字形，需采取一系列的措施来防止在变倍过程中的失真。



## 4.3 防止字形失真的措施

### ❖ 1. 防止字形变倍时的舍入误差积累。



变倍时舍入误差的积累





## 4.3 防止字形失真的措施

❖ 结点变倍方法，用下面的公式获得变倍后的 $\Delta x'$ 、 $\Delta y'$ （变倍率= $r$ ）：

$$\begin{aligned} A'B' \text{折线的} \Delta x' &= B' \text{点的} x \text{坐标} - A' \text{点的} x \text{坐标} \\ &= (B \text{点} x \text{坐标} \times r)(\text{舍入}) - (A \text{点} x \text{坐标} \times r)(\text{舍入}) \end{aligned}$$

$$\begin{aligned} A'B' \text{折线的} \Delta y' &= B' \text{点的} y \text{坐标} - A' \text{点的} y \text{坐标} \\ &= (B \text{点} y \text{坐标} \times r)(\text{舍入}) - (A \text{点} y \text{坐标} \times r)(\text{舍入}) \end{aligned}$$

\* 这里 $(p)(\text{舍入})$ 表示 $p$ 经过四舍五入后取整数值。



## 4.3 防止字形失真的措施

### ❖ 2. 保证规则笔画宽度的一致性。

✎ 对于规则笔画横、竖、折的宽度部分，不采用结点变倍法，而按下述公式计算：

变倍后的宽度 = (变倍前宽度 × r)(舍入)

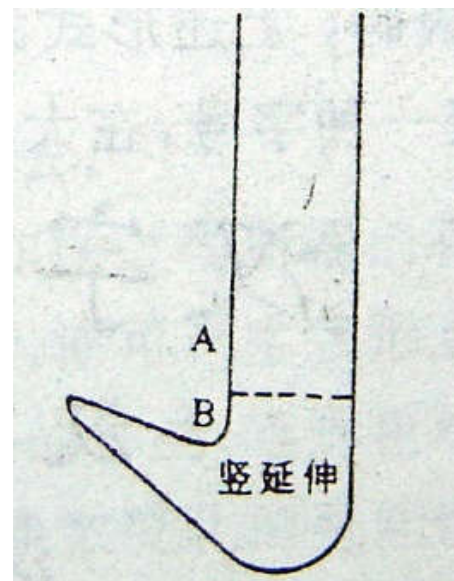


## 4.3 防止字形失真的措施

### ❖ 3. 规则笔画与不规则笔画衔接部分的处理

A点的A'点由直线变倍得到，而对应B点的B'点由结点变倍得到。这样，AB直线本来是直的，但变倍后的A'B'折线可能变成向左或向右倾斜一格的斜线。

可采取以下措施克服此失真：  
当延伸的第一笔原 $\Delta x = 0$ ，则变倍后永远使这一笔的 $\Delta x' = 0$ 。



笔画‘勾’的下延伸



## 4.3 防止字形失真的措施

### ❖ 4. 小号字横的宽度的控制和笔锋的细致描述。

✎ 字号缩小很小时，在变倍过程中会引起某些部分变得过分密集，尤其当一个字内横的数量很多时。



## 4.3 防止字形失真的措施

### ❖ 4. 小号字横的宽度的控制和笔锋的细致描述。

✎ 在横的压缩信息中，用两位二进制数细致刻画横宽。

编号	横宽
00	1.7格横宽
01	2格横宽
10	2.6格横宽
11	3格横宽

✎ 这里的格都是指五号字96 × 96网格中的格。当然，00和01编号在五号字时，只能都是两格宽，显不出差别，因为小数无法体现。但变成六号字时，01编号仍是两格宽，而00编号则变成一格宽。对于笔画很多，尤其横很多的汉字，00编号的横将使这一汉字在各种小字号下都保持清晰美观。



## 4.4 字形放大缩小的应用

- ❖ 点阵方法描述的字形在缩放时采用插值加线或抽线的方法，会产生严重变形或以质量损失为代价。可以采用平滑技术使锯齿边的台阶缩小或用斜边代替锯齿。但是无论平滑技术如何好，所产生的字形因锯齿或折线而产生的不平滑感只能缩小而不能排除。



## 4.4 字形放大缩小的应用

- ❖ 轮廓矢量方法是为了改进点阵字形在缩放时会出现锯齿状轮廓而产生。在文字复原时，通过对向量的长度进行变倍放大或缩小，便可获得不同的字号。
- ❖ 如果在变倍时对x方向和y方向分别选不同的倍率，还可以获得变形长体字或扁体字。
- ❖ 轮廓字库在还原时需经特定的运算将这些用轮廓描述的字形转化成点阵方式，再由输出设备输出。
- ❖ 这个运算可以在输出设备内进行，也可在计算机上先转换成位图再输出。



## 1.4 字形放大缩小的应用

- ❖ 轮廓字库解决了点阵字库缩放后失真的问题。这种字形描述技术对汉字横平竖直的字形有很好的描述效果，但对于撇、捺、点，以及圆弧装饰效果的字体，即使是缩小直线距，仍然不能十分逼近字形，而且字体放大之后，轮廓仍然具有折线感。
- ❖ 由于近年来硬件的迅速发展，有能力用二次曲线或三次曲线代替直线描述字形轮廓的字库。





## 5. TTF 字形技术

---

- ❖ 什么是TrueType
- ❖ TrueType的基本原理
- ❖ TrueType字体文件结构
- ❖ TrueType的特点和优势
- ❖ TrueType的应用
- ❖ OpenType的出现



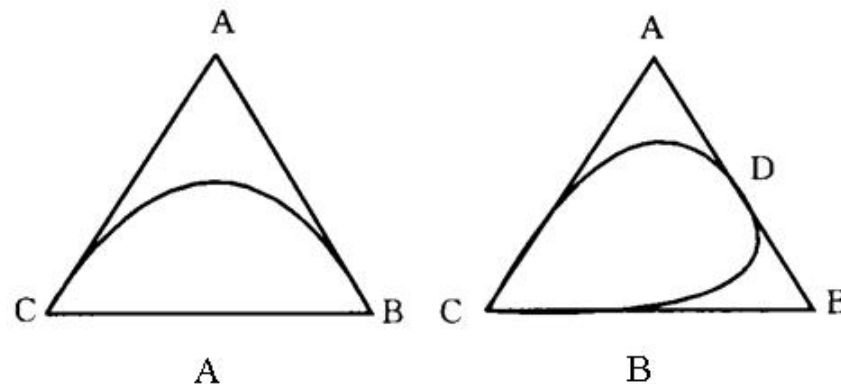
## 5.1 什么是TrueType

- ❖ TrueType是Apple公司和Microsoft公司合作开发的PDL（Page Description Language）语言TrueImage中的字形描述部分。
- ❖ TrueType有一套专门的指令集，用于完成字形的描述、特征参数的描述及其他一些描述。它用的轮廓描述曲线是二次B-样条（Bezier-spline）函数。



## 5.2 TrueType 的基本原理

- ❖ TrueType 字体对字形的描述是轮廓描述，在轮廓上有一串结点组成，这些结点有 **on-line** 和 **off-line** 两种属性来描述一个封闭的轮廓线。





## 5.2 TrueType 的基本原理

- ❖ TrueType对字形轮廓上的结点都给出一个编号，其顺序是按顺时针方向走一圈，即填黑的部分永远在其右边。
- ❖ TrueType提供了一个很大的指令集，用于特殊的字形特性的描述，所谓的Hinting技术，就是用TrueType指令集完成的。
- ❖ TrueType用一系列的轮廓点的坐标，构成字符的原始轮廓，用指令信息对原始轮廓进行修正。



## 5.2 TrueType的基本原理

- ❖ TrueType字库中的主要特征参数
- ❖ 在X方向上有5个关键距离值：
  - ☞ 字身最左边的起始点到字母主干的空间；
  - ☞ 字母主体部分的宽度；
  - ☞ 字身的宽度（西文字母字身不等宽）；
  - ☞ 字母主干（**stem**）的宽度；
  - ☞ 字母衬线（**serif**）的大小。
- ❖ 在Y方向上：
  - ☞ 字母横干（**crossbar**）的厚度；
  - ☞ 字母衬线的厚度；
  - ☞ 字母主体的高度；
  - ☞ 字母横干的高度。



## 5.2 TrueType的基本原理

- ❖ 在变倍时用数学方法将相同的特征参数值同等定义，在任何字号下都同等还原，这时考虑的是字体各特征点的形状，而不考虑坐标格的修整。
  - ✧ 宋体等笔画较细的字体，控制笔画的宽度
  - ✧ 黑体等笔画较粗的字体，还要控制笔画间的距离，如“量”字，横画需要保持均匀



## 5.2 TrueType的基本原理

- ❖ TrueType轮廓结点到光栅点的对应
  - ✧ 光栅点落在轮廓内，黑点
  - ✧ 否则，白点
- ❖ 掉点（Dropout）现象。
- ❖ TrueType的指令集可以最大程度的避免掉点现象的发生。
- ❖ TrueType字形文件包括
  - ✧ 轮廓数据
  - ✧ TrueType指令写的程序



## 5.3 TrueType字体文件结构

- ❖ TrueType字体（TrueType Font, TTF）文件总体结构：

文件头（12 Bytes）
描述表目录（每个16 Bytes）
描述表数据





## 5.3 TrueType字体文件结构

❖ 文件头共12B:

Fixed	sfnt_version	//TTF文件版本号码 (0x00010000)
USHORT	numTables	//描述表数目
USHORT	searchRange	//描述表快速查找范围
USHORT	entrySelector	//描述表入口选择
USHORT	rangeShift	//范围调整

❖ 注：最重要的表项是numTables，其他一些数据，如searchRange、entrySelector和rangeShift都是通过numTables计算出来的。



## 5.3 TrueType字体文件结构

- ❖ 描述表目录。其目录个数由numTables决定，每个目录项16B:

ULONG	tag	//描述表名称（4 Bytes）
ULONG	checkSum	//描述表内容的校验和
ULONG	offset	//此描述表的位置偏移（从文件头开始）
ULONG	length	//描述表字节长度

- ❖ 根据描述表目录可以了解TTF文件中都包括哪些描述表，这些描述表在文件中的什么位置。



## 5.3 TrueType字体文件结构

- ❖ 常用的TTF文件描述表共有19个，其中有10个为每个TTF文件所必须包含的，其他的9个可根据需要任选。
- ❖ 10个必须的描述表：
  - 🔗 ‘cmap’ 字符代码到文字序号的映射表
  - 🔗 ‘glyf’ 文字轮廓数据
  - 🔗 ‘head’ 文件头信息
  - 🔗 ‘hhea’ 水平度量头信息
  - 🔗 ‘hmtx’ 水平度量信息
  - 🔗 ‘loca’ 字符轮廓数据位置索引
  - 🔗 ‘maxp’ 最大值描述
  - 🔗 ‘name’ 名字表
  - 🔗 ‘post’ PostScript打印机控制
  - 🔗 ‘OS/2’ OS/2和Windows度量信息



## 5.3 TrueType字体文件结构

❖ 以下9个可选的描述表：

☞	‘cvt’	控制值表
☞	‘fpgm’	定义程序区
☞	‘hdmx’	水平设备度量
☞	‘kern’	紧排控制
☞	‘LTSH’	线性门限
☞	‘prep’	CVT程序区
☞	‘WIN’	保留
☞	‘VDMX’	垂直设备度量
☞	‘FOCA’	保留



## 5.4 TrueType的特点和优势

### ❖ TrueType的特点有：

- ❧ TrueType既可作打印字体，又可以用作屏幕显示。
- ❧ 由于它是由指定字形进行描述，因此与分辨率无关，输出时总是按照打印机的分辨率输出。
- ❧ 无论放大或缩小，字符总是光滑的，不会有锯齿出现。
- ❧ 相对PostScript字体来说，其质量要差一些。特别是在文字太小时，表现得不是很清楚。



## 5.4 TrueType的特点和优势

- ❖ 由于TrueType技术本身的特点而使它具有以下优势：
  - ✧ 真正的所见即所得效果
  - ✧ 支持字体嵌入技术，保证文件的跨系统传递性
  - ✧ 操作系统平台的兼容性
  - ✧ ABC字宽值
  - ✧ 精确的字体轮廓描述
  - ✧ 品种繁多的字体家族



## 5.5 TrueType 的应用

### ❖ TTF字体的创建

✧ 使用已有的TTF字体编辑软件，如：Font Creator Program。

### ❖ TTF字体在面向对象程序中的引用：

✧ Windows使用LOGFONT结构定义字体文件的属性。此结构可以在程序中直接引用。



## 5.6 OpenType 的出现

- ❖ “前后端不一致”问题
  - ✧ 用TrueType字库排好版面后，后端是否有对应的PostScript字库？
- ❖ 1996年5月，微软和Adobe公司联合推出了一种新的通用字体格式——OpenType（简称OTF）。





## 5.6 OpenType 的出现

- ❖ OpenType集中了TrueType和Type 1两种字形描述技术的特点，是一个开放、无版权的字形描述技术。
- ❖ OpenType字体是TrueType字体的扩展，它在包含了TrueType的字体轮廓信息的同时，还可包含压缩字体格式（CFF）PostScript字体轮廓信息。



## 5.6 OpenType 的出现

- ❖ OpenType 字体格式能达到以下目的：
  - ✧ 更宽的多平台支持；
  - ✧ 更好的支持国际字符集；
  - ✧ 更强的字体数据保护；
  - ✧ 字体文件更小，有利于提高发送效率；
  - ✧ 为先进的字体印刷控制提供更多的支持。



## 6. 字库的设计和管理技术

---

- ❖ 汉字字库的基本要求
- ❖ 汉字库性能的评测
- ❖ 汉字库结构的设计
- ❖ 点阵字库的制作



## 6.1 汉字字库的基本要求

- ❖ 汉字库的设计应该追求三项目标：
  - ☞ 字形质量高
  - ☞ 读取速度快
  - ☞ 系统开销小
- ❖ 它们之间的关系是：随着字形点阵数的增加，字形失真就越小；但字形库容量会增大，也就意味着字形库成本增加，字形点阵存取速度下降。



## 6.1 汉字字库的基本要求

- ❖ 对于汉字显示来说，汉字字形库的速度可以估算如下：
  - ✧ 设汉字显示格式为12行×40字，显示器刷新速率为50周/秒，那么需要字库速度为：
$$12 \times 40 \times 50 = 24000 \text{字/秒}$$
  - ✧ 如果显示格式为25行×40字，则有字库速度为：
$$25 \times 40 \times 50 = 50000 \text{字/秒}$$
- ❖ 显然，字形点阵数越大，要求字库存取速度越高。事实上，由于显示器的分辨率所限，通常显示字库用16×16和24×24为宜。



## 6.1 汉字字库的基本要求

- ❖ 一般来说，打字机所需的字库速度不会高于显示器所需字库的速度。

记录方式	特点	记录速度
电子照相	超高速打印	1000—25,000字/秒
静电式	高速打印	数百—1000字/秒
喷墨式	中速打印	30—120字/秒
针式	中低速打印	10—50字/秒
感热式	低速打印	5—20字/秒
行式	高速打印	2000—5000字/秒

几种汉字打印机的特点及记录速度



## 6.2 汉字库性能的评测

- ❖ 汉字库的性能主要包括访问字库的速度 $v$ 和字库的内存开销 $m$ ， $v$ 和 $m$ 分别与汉字库的性能成正比和反比关系。
- ❖ 用 $k$ 参数来描述汉字库的性能：

$$k = (v - v_0) / m$$

式中 $v$ 为输出模块用此汉字库后所能达到的输出汉字平均速度， $m$ 为汉字库的内存开销， $v_0$ 为输出模块应具备的输出汉字的起码平均速度， $v_0$ 的取值决定了 $k$ 参数描述的有效范围。



## 6.2 汉字库性能的评测

- ❖ 进一步分析 $v$ 与 $m$ 之间的关系后，用下式表示：

$$v = m \times f(s) \times w$$

式中 $s$ 为GB2312的全集， $f$ 为确定其子集作为汉字库常驻内存部分的法则， $w$ 为其它因素。

- ❖ 欲提高 $k$ 参数，必须寻求 $f(s)$ 的最佳情况。由于中文信息处理系统的使用环境不是静态的，因此采用静态汉字库结构不可能进一步获得较高的 $k$ 参数，而必须考虑动态汉字库结构，以适应动态的使用环境。





## 6.3 汉字库结构的设计

- ❖ 输出模块实际是将汉字内部码转换成对应的汉字字形信息。可以用下面的公式表示这一过程：

$$y = f(x)$$

其中 $x$ 为汉字内部码， $y$ 为汉字字形信息， $f$ 为输出模块的转换处理。

- ❖ 汉字库在中文信息处理系统中占有相当重要的地位。



## 6.3 汉字库结构的设计

- ❖ 存储汉字库的介质有多种，如RAM、ROM、磁盘和磁带等。
  - ❧ 早期有一种基于RAM的汉字库
  - ❧ 现在的TTF汉字库通常是以文件的形式存在于磁盘上



## 6.3 汉字库结构的设计

### ❖ 静态汉字库：

- ❧ 全内存型字库 — 访问速度快，但占用大量的内存资源。
- ❧ 全外存型字库 — 响应速度较慢，只能满足一般用户的使用要求。
- ❧ 内外存结合型字库 — 可达到既减小内存开销，又保证访问速度的目的。实践证明具有较高的效率和较好的性能。



## 6.3 汉字库结构的设计

### ❖ 动态汉字库：

#### ☞ 准动态型字库

- ❖ 由静态汉字库中的内外存结合型字库改进而成。
- ❖ 访问这种汉字库的算法。
- ❖ 称不上真正的动态型字库。

#### ☞ 自适应型字库

- ❖ 分为内存字库和外存字库两个部分
- ❖ 内存字库中的内容由系统根据使用环境进行自动调整，是一种真正的动态型字库。

#### ☞ 多级型字库

- ❖ 多级型字库是对自适应型字库的改进。
- ❖ 设计了三级汉字库，分别是常用字库、动态字库和外存字库。



## 6.4 点阵字库的制作

- ❖ 要做到造字迅速、准确、直观和修改方便，必须使用计算机辅助设计。
- ❖ 关键要有点阵字形的显示，修改和存储，即要有一个造字软件。
- ❖ 造字时软件在终端屏幕上画出放大的栅格，然后利用键盘严格按字模移动光标打点、画线，最后完成一个汉字的输入。



## 6.4 点阵字库的制作

❖ 点阵字形设计基础是点阵字稿的设计。在进行字稿设计时，应该充分注意下述六个方面：

- ❧ 结构端正、重心平稳
- ❧ 均匀足格、大小一致
- ❧ 偏旁组合、比例协调、上下对准、左右平衡
- ❧ 粗细黑白、灵活掌握
- ❧ 笔画舒展自然
- ❧ 部首归类，相对统一，灵活套用



## 7. 汉字字库的应用

---

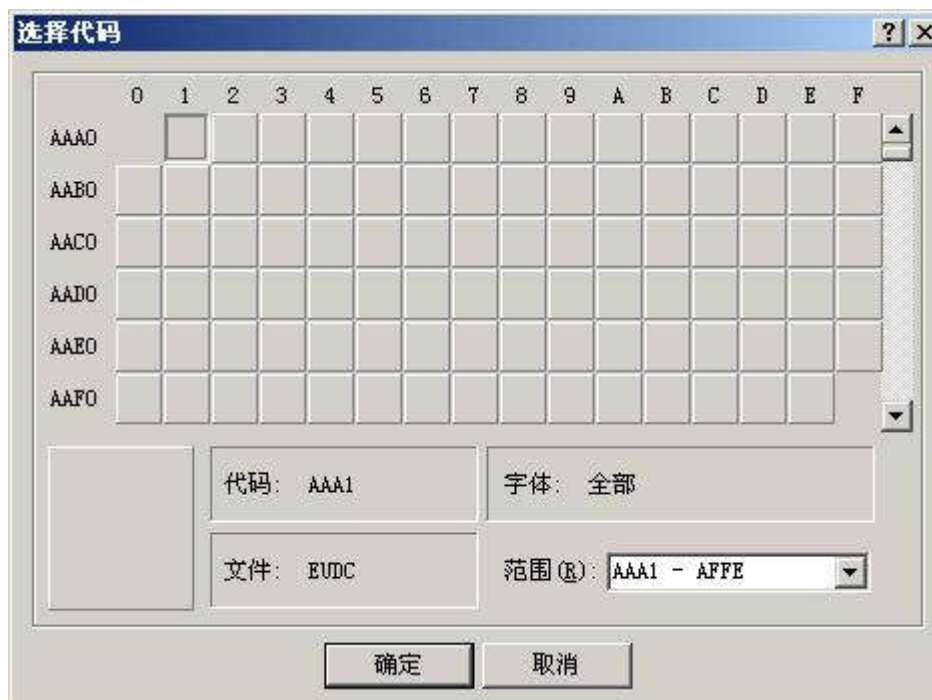
- ❖ 造字程序的使用
- ❖ 安装新字体



## 7.1 造字程序的使用

### ❖ 打开造字程序

🔗 Windows开始—程序—附件— TrueType造字程序（专用字符编辑程序）

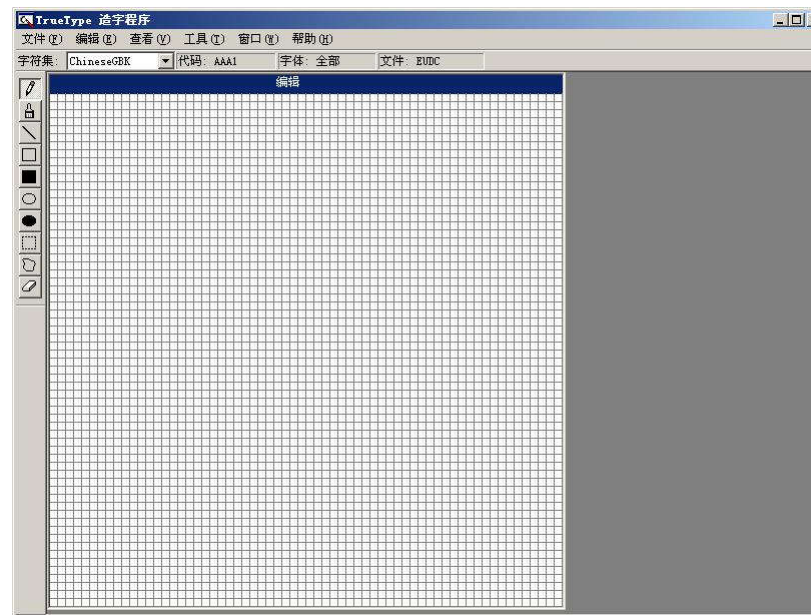






## 7.1 造字程序的使用

✧ 选择一个代码（例如，选定代码AAA1），进入编辑窗口





## 7.1 造字程序的使用

### ❖ 造字方法

- ❧ 第一种方法：利用窗口左侧工具栏中的前几种工具：铅笔、画笔、直线或者矩形等，直接在编辑区拖动绘出字形。这需要较多耐心，但可以创造出造型千奇百怪的字形。
- ❧ 第二种方法可以从几个字中各取出一部分来组合成新字，这样就能够很好的保持字体风格。
- ❧ 第三种方法是直接在画图程序或其他图形处理软件中绘制好需要的字，选中复制到造字程序的编辑窗口。这个方法可以同时做好几个字，组合出新意，甚至可以把你喜欢的图形或自己的相片做成字，可惜只能是单色图。



## 7.1 造字程序的使用

### ❖ 造字方法

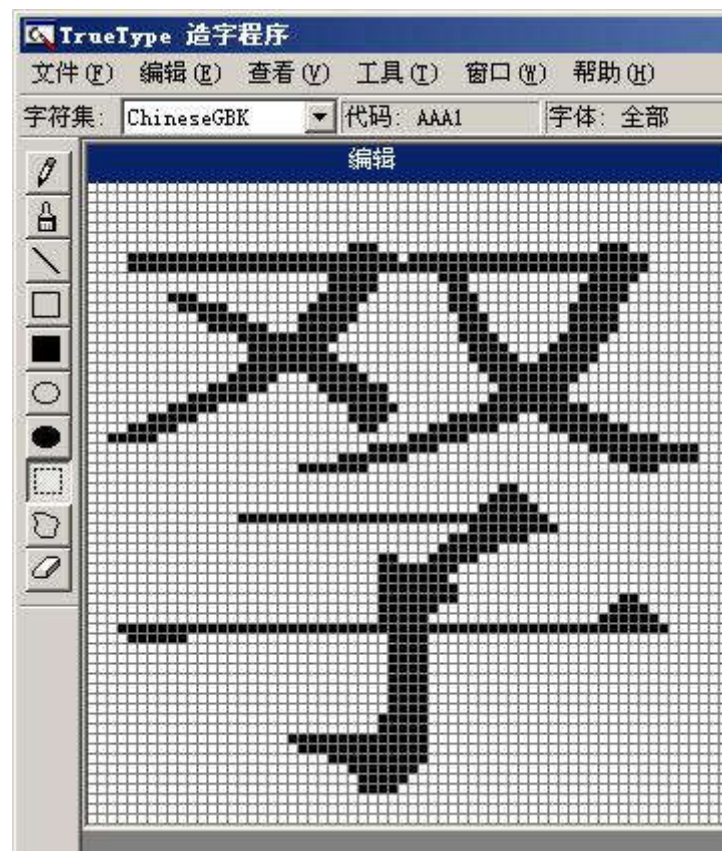
☞ 使用第二种方法，以“双字下面一个子字”为例

- ❖ 点击菜单栏的“编辑—复制字符”，打开复制字符对话框。
- ❖ 点击字体按钮，选择宋体。
- ❖ 在“形状”输入框输入“双”字，点击确定，造字程序自动返回编辑窗口，可以看到“双”字出现在编辑区。选取“双”字的全部字形并把它上压。
- ❖ 点击菜单栏的“窗口—参照”，打开参照对话框，同样设定字体为宋体，在“形状”输入框中输入“李”字，确定后会多出一个参照区，按照上面的选取技巧，选取其中的“子”字。把选定的“子”拖动到编辑区，并调整到适当大小与“双”字组成新字。
- ❖ 点击菜单栏的编辑—保存字符。一个新字就做好了。



## 7.1 造字程序的使用

✧ 新造出来的字



苏州大学：中文信息处理



## 7.1 造字程序的使用

### 保存与输入

#### ❖ 1. 间接法

- ❖ 点击造字程序菜单栏的“编辑—复制字符”，在“代码”输入框中输入AAA1，选中出现在“形状”输入框的字，按“**Ctrl+C**”复制，再按“**Ctrl+V**”粘贴到你需要的地方。而且系统会自动将字体大小调整合适。
- ❖ 另一种方法是打开字符映射表程序。点击“开始—程序—附件—系统工具”，运行“字符映射表”程序。从字体下拉列表中选择“所有字体（专用字符）”，可以看到我们刚刚造好的字。双击需要复制的字符，该字就会出现在“复制字符”框，单击“复制”按钮，然后到需要输入的地方粘贴即可。



## 7.1 造字程序的使用

### 保存与输入

#### ❖ 2. 直接法

- 第一种方法可以采用内码输入法，或称作区位输入法。在系统中添加好内码输入法，在需要输入的地方按照刚才选定的代码进行输入即可。
- 第二种方法是进行“输入法链接”。在造字程序中完成造字保存后，点击菜单栏的“编辑—输入法链接”，在弹出的外码对话框中分别输入各种输入法的外码，每输入一种外码都要按空格键确认，然后单击注册按钮。以后就可以用已链接的输入法输入了。这种方法只对“全拼”、“微软拼音”等少数几种输入法有效。





## 7.2 安装新字体

- ❖ 在网上下载需要安装的字体，一般是一些后缀名为“.ttf”文件。
- ❖ 将新的字体文件直接拷贝到Windows的字体目录里面（一般在C:\Windows\Fonts目录中）即可。



## 7.2 安装新字体

- ❖ 如果上述方法不行，则在控制面板中打开字体程序。







## 7.2 安装新字体

- ❖ 文件菜单选择“安装新字体”
- ❖ 在弹出的“添加字体”的对话框中，选择字体文件存放的路径，新的字体就会出现字体列表框中，选择所需要的字体，并且在下面选“将字体复制到**Fonts**文件夹”这个选项。当安装完成之后就会在字体菜单里面看到新安装的字体了。



## 7.2 安装新字体



安装新字体对话框

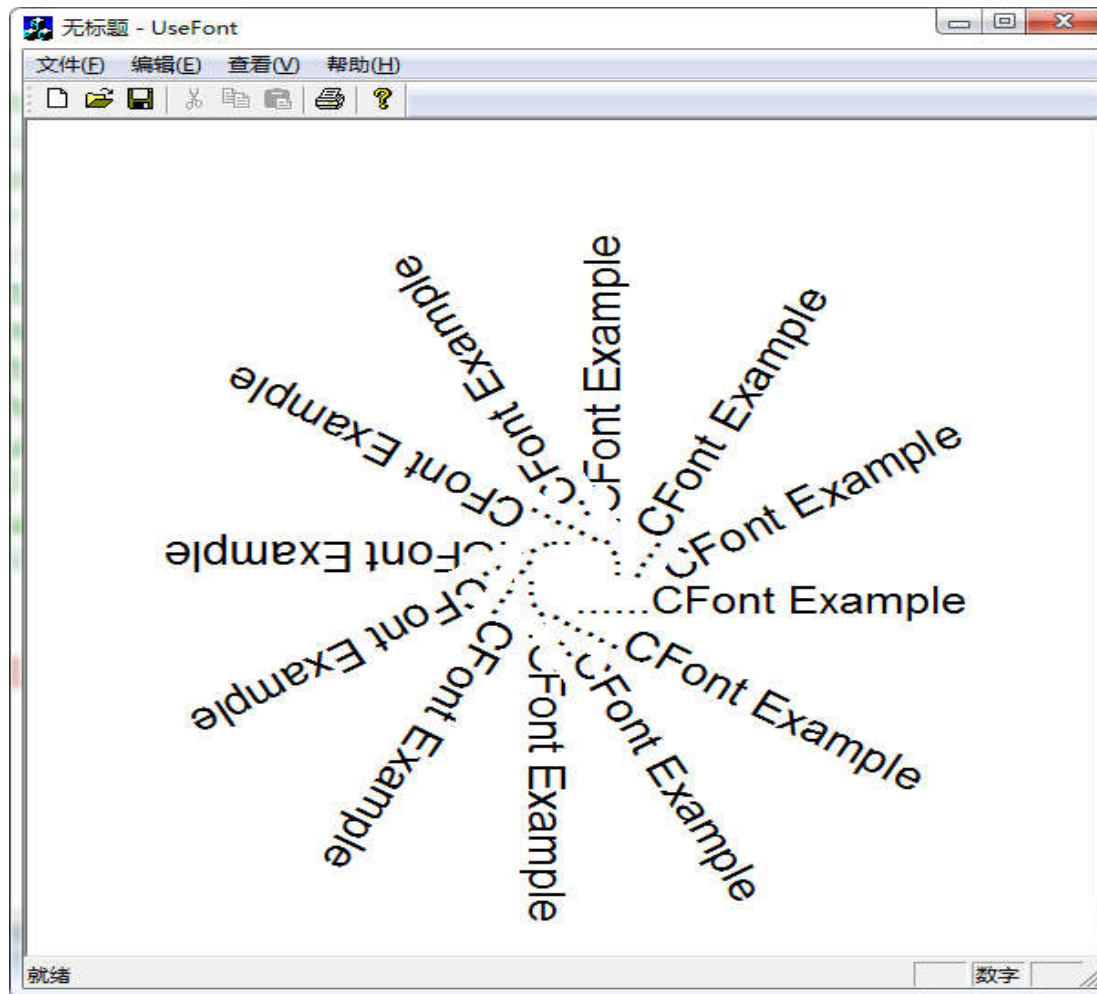


# Windows程序使用字体

```
CFont *pOldFont;
for(int i=1 ;i<=12;i++)
{
    font1.CreateFont(
        40,                // nHeight
        0,                 // nWidth
        i*300,             // nEscapement
        0,                 // nOrientation
        FW_NORMAL,         // nWeight
        FALSE,             // bItalic
        FALSE,             // bUnderline
        0,                 // cStrikeOut
        ANSI_CHARSET,      // nCharSet
        OUT_DEFAULT_PRECIS, // nOutPrecision
        CLIP_DEFAULT_PRECIS, // nClipPrecision
        DEFAULT_QUALITY,   // nQuality
        DEFAULT_PITCH | FF_SWISS, // nPitchAndFamily
        "Arial");
    pOldFont=pDC->SelectObject(&font1);
    pDC->TextOut (300,300,".....CFont Example");
    pDC->SelectObject (pOldFont);
    font1.DeleteObject ();
}
```



## 使用字体的效果





# 作业

---

❖ P.160, 3-12