

字形的“号制”“点制”与“米制”

张小衡

(香港理工大学中文及双语学系,香港九龙)

E-mail: xtzhang@polyu.edu.hk

摘要 号制和点制是计算机文字处理中表示字形尺寸的最常用标准。论文简单叙述这两种标准的历史和现状,给出一个含 Word2003 所有字号的“号-点-厘米”对照表,并介绍对照表数据的实验获取方法和多种用途。最后,在深入分析讨论号制和点制的种种缺陷的基础上提出字形度量全球统一使用米制的构想。

关键词 字形尺寸 号制 点制 米制

文章编号 1002-8331-(2006)10-0175-03 文献标识码 A 中图分类号 TP311

The Number, Point and Metric Systems of Font Size

Zhang Xiaoheng

(Dept. of Chinese and Bilingual Studies, Hong Kong Polytechnic University, Hong Kong)

Abstract: The point system and number system are the most commonly-used standards for font size measurement in China and the world. In this paper a brief introduction to the history and technical contents of both systems is given. Then a font size “number-point-millimeter” cross reference table with first-hand data from experiment is presented, followed by a description of some important applications of the table. Finally a unified metric system for font measurement is proposed based on an in-depth discussion on the deficiencies of the point and number systems.

Keywords: font size, number system, point system, metric system

1 前言

在计算机上,英文字形的尺寸大小(font size)常常以“点”(point)为度量单位来表示,而中文字形大小则常用“号”来说明。有时我们也用“号”来说明英文字形,用“点”来说明中文字形。因此,了解和研究号制和点制这两套字形度量标准及其相互关系对于语言文字的应用和计算机信息处理都具有重要的意义。这方面的知识有时甚至是必不可少的。例如,笔者前不久收到某论文集的入编通知,其中对于文章的排版格式做出这样的要求:“论文用 Word 编排,正文为中文宋体,字号 5 号……。文章内的图表、参考文献、作者简介要用 6 号宋体……。题目必须居中,用小 2 号宋体加粗,作者名字写在题目下面且居中,用 4 号楷体,……。”虽然我使用的英文版 Windows XP 和 Word 2003 提供所需的各种中文字体,但在字形大小方面只有点制可用,没有字号的选项。要在这种“点”的计算机环境中满足“号”的版面要求,就必须弄清字号和字点之间的对应关系。

关于字形号制和点制的讨论,我们在“中国期刊全文数据库”上仅检索到两篇学术论文^[1,2](不包括科普性短文)。由于写作时间较早,文章中有关 Word 97 等软件的内容现在看来显得过于陈旧,另一个美中不足的地方是许多重要数据的来源没有明确交待。本文以在“Windows XP+Word 2003”语言信息处理环境中获取的第一手实验数据为依据,较为全面深入地讨论字形号制和点制的关系、应用和发展等问题。

2 字号和字点简介

国际上,字形大小一般用“点”来度量,在中国内地,字形大小常用字“号”来表示。中文简体字版 Windows XP 和 Word 2003 既提供点制服务又提供号制服务。

2.1 字形的点制度量标准

“点制”是英文“point system”的译文,这里的“point”(简称 pt 或 p)在汉语中除了意译为“点”之外,还有音译为“磅”的。在西方印刷工业发展的初期,社会上使用的英尺英寸度量体制未能提供足够精确的度量单位来满足字形排版的需要,因此,1737 年法国巴黎的印刷字模铸造家 Pierre Fournier 发明了点制度量法,他的一个单位点相当于现在的 0.349mm(毫米)。大约 40 年以后,巴黎的另一个字模铸造家 Francois Ambroise Didot 向社会推介他的改良版点制度量法,Didot 的一个单位长度是 0.375 9mm。美国和英国则分别于 1886 和 1898 年决定以单位长度相当于 0.351 5mm(约 1/72 英寸)的点作为国内统一的字形度量标准^[3]。这个标准一直沿用至今,而且由于英美在世界政治经济和科技上的影响,尤其是其科技产品(如 MS Windows 和 MS Word 等)的作用,已经成为国际上最为通用的字形尺寸标准。下文所说的点制,如无特别说明,则是指英美点制。

字形大小为多少点,指的是该字形的满格字符高度有多长。在 MS Word 2003 上可用的字形点值是 1-1 638 之间可以被 0.5 整除的所有数字,即集合{1, 1.5, 2, 2.5, ..., 1 637, 1 637.5, 1 638}

基金项目:香港理工大学资助(编号:1-9827 A-PA0K)

作者简介:张小衡,男,香港理工大学助理教授,湖南大学兼职教授,英文专业学士,计算机专业硕士、博士,主要研究领域为计算语言学 and 计算机

201994-2(辅助语言教学)。Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

中的任一元素(这些规定可以直接在 Word 2003 上得到验证)。可见点制度量范围之广,取值之密。

2.2 字形的号制度量标准

根据中国国家标准 GB/T12200.2-94[4]第 4.1.2.9 节,字号就是“印刷体依据字体大小所编的号”。从历史上看,字号是印刷行业为统一字模规格而人为规定的尺寸标准。活字排版印刷技术发明以后,汉字大小的度量方法长期没有统一,直到 1858 年,美国人 William Gamble 根据当时美国铅字的尺寸为汉字制定了七种尺寸规格,即 1 至 7 号。后来人们又根据需要增加了一些新号,逐步发展成现在的字号体系^①。

中文简体字版 Windows XP 和 Word 2003 提供的字号选项由小到大是:八号、七号、小六号、六号、小五号、五号、小四号、四号、小三号、三号、小二号、二号、小一号、一号、小初号、初号。其中最小的八号字形高度约 1.76mm,最大的初号字形高度约 14.76mm(将在第三节说明)。方正电子排版系统还设有比初号更大的小特号(14.794mm)、特号(16.917mm)和特大号(19.726mm)^②。

顺便说明,“点制”和“号制”有人称之为“点数制”和“号数制”,本文不采用后种说法是出于以下几方面的考虑:(1)字号系统中的“特号”、“初号”、“小一号”、“小二号”等都不是严格的数量表达式;(2)就是严格按数量计算的“米制”、“尺制”等人们也很少称之为“米数制”、“尺数制”等;(3)点制的英语原文“The point system (of measurement)”并没有“数”的含义;(4)在语言表达上,“点制”和“号制”比“点数制”和“号数制”简练些。

3 号-点对照表及其制作方法

字形号-点对应关系的主要内容在于说明每种字号相当于多少点的尺寸。由于号制和点制之间不存在换算公式(这一点将在第 5 节说明),所以我们只好采用对照表的方法。

3.1 号-点对照表

表 1 是笔者用一手实验数据建立起来的“号-点-毫米”对照表。

表 1 字形大小“号-点-毫米”对照表

字号	点数/pt	毫米/mm
八号	5	1.76
七号	5.5	1.93
小六号	6.5	2.28
六号	7.5	2.64
小五号	9	3.16
五号	10.5	3.69
小四号	12	4.22
四号	14	4.92
小三号	15	5.27
三号	16	5.62
小二号	18	6.33
二号	22	7.73
小一号	24	8.44
一号	26	9.14
小初	36	12.65
初号	42	14.76

表中涵括了中文简体字版 Windows XP 和 Word 2003 上的所有字号,每种字号都附有相应的点(pt)数和毫米(mm)数,而且三项数据都以该尺寸排版。例如,从表中我们可以看到五号字的尺寸相当于 10.5pt 或 3.69mm,而且“五号”、“10.5”和

“3.69”这组数据就是用它们自身所表示的字形尺寸排版的。有了这个号-点转换表,要在各种非中文简体字版 Windows 上的点制字形环境中实现用字号说明的版面要求就显得轻而易举了。

3.2 号-点对照表的制作方法

上述“号-点-毫米”对照表的基本数据不是从其它文献抄拼而成的二手材料,而是笔者亲自在 Windows XP 和 Word 2003 上通过简单可靠的实验直接获取的,步骤如下:

(1) 在支持字号的中文简体字版 Window XP+Word 2003 环境中列出“八号、七号、小六号、六号、小五号、五号、...”等 Word 提供的所有字号名称,而且每个名称都用它自身所表示的尺寸规格设定字形大小,由此得出对照表中的第一列内容。然后将结果存为一个 Word 文件。

(2) 将步骤(1)所产生的文件拿到不支持字号的英文版 Window XP+Word 2003 环境中打开。用鼠标点击每一个字号名称,Word 工具列上的“Font Size”文字框中就会显示它们各自的字形点数。将这些数据写入对照表中的第二列,并用相应的尺寸设定字形大小。

(3) 按 $1\text{ pt}=0.351\ 459\ 8\text{ mm}$ 的精确转换公式(见下文第 5 节中的说明)计算出第三列的数据,并设定相应的字形大小。

最后再回到中文简体字版“Window XP+Word 2003”环境中核对数据。

4 号-点对照表的用途

号-点对照表的作用是多方面的。首先,该表提供较全面的“号-点-厘米”字形大小对应数据,并显示各种规格的印刷效果,有助于人们了解字形大小的不同描述体系以及它们之间的关系,有利于国际信息交流和学术合作。

在无字号选项的计算机环境中,号-点对照表可帮助用户实现用字号说明的排版格式。例如,由于采用 Unicode,MS Word 2003 的英文版、简体字中文版(内地)和繁体字中文版(台湾)都能处理包括中英文在内的多语字符集,此外,它们也都提供点制的字形度量服务,然而,字号选项却只有简体字中文版的 Word 提供。因此,如果要在 Word 2003 的英文版或繁体字中文版上从事涉及“字号”的格式排版,就可通过对照表将字号转换为点数来处理。

在有字号表示的排版环境中,号-点对照表可用于在相邻字号间“微调”字形大小。假如我们需要比小四号大,比四号小的字形。这在纯字号的环境中是无法满足的。但由于支持字号的计算机排版环境一般都同时支持点制,我们可以利用号点对照表来实现这种“微调”。从表中查得小四号相当于 12pt,四号相当于 14pt。在 Word 上,12pt 和 14pt 之间还有 12.5pt,13pt 和 13.5pt 三个数值可供选用。

为了进一步方便在无字号服务的环境中处理字号排版要求,可考虑给 Windows XP 和 Office 2003 附加安装多语用户接口(Multilingual User Interface)软件,这样可提供简体中文用户界面,并在字形大小选项中加入各种字号,结果就可以像使用简体字版 Word 一样来处理字号排版问题。但是这种方法也有缺点,就是多语用户界面软件需要额外购买,另外安装。如果电脑不是自己的,就是买了软件,恐怕也不能随便安装。另一种方法不需额外装软件,而是在 Word 上根据对照表提供的数据为每种字号建立以该字号命名的新“样式(Style)”。假设要在

英文版 Word 2003 的环境中建立宋体五号的样式,先在对照表中查得五号字形相当于 10.5pt,然后选择工具菜单 Format, Style and Formatting, New Style,就可以在“新样式”对话框里建立名称为“宋体五号”,字体为“宋体”,大小为“10.5pt”的字形样式供日后直接选用。如需要另一种字体的五号尺寸,可为其另建一个字样,也可先用已有的宋体五号处理后再将字形改为所需字体(大小不变)。

除 Word 之外,号-点对照表对于 PowerPoint,Excel 和 WWW 等其它计算机环境也照样有用。这里就不赘述了。

5 关于号制和点制的思考

号制和点制都有不尽如人意的地方。号制只限于表示字形大小,应用面比较窄。而且,从八号到初号总共只有 16 个选项可用,对于当今的科技信息时代来说实在是太多了。这个问题不是单靠增加新字号就可以解决的,原因在于号制的另一个缺陷:字号之间除了粗略的大小关系之外,并无规律可循。例如,我们知道五号字形比四号小,但是五号所表示的尺寸长度却难以从四号或其他号码的尺寸计算出来(注:尽管字号可以分为五个系列,但无整体规律^[6])。这种情况可以从表 2 得到证实。

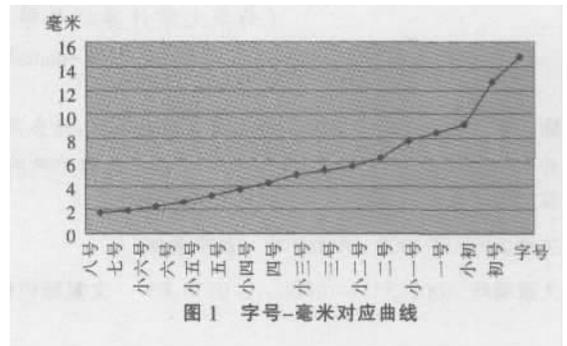
表中的第一、二列分别给出各种字号名称和相应的毫米尺寸。第三列表示相邻两号的长度差(mm),对应于某一个字号的长度差表示该字号与下一字号的尺寸差值。例如对应于八号字的长度差 0.17mm,表示该字号的尺寸长度与下一字号(七号)相差 1.93mm-1.76mm=0.17mm。第四列给出相邻字号的长度比。对应于某一个字号的长度比表示下面一行的字号与该字号的长度之比,例如第一行数据中的长度比 1.10 是七号的毫米数除八号的毫米数之商,即 1.93/1.76=1.10。

表中第三列的数据显示,号制中各相邻两个字号之间的尺寸距离不恒等,而且变化也无规律。例如,小五号和五号相差 0.53mm,五号和小四号也相差 0.53mm,但小四号和四号的差值上升到 0.70mm,而四号和小三号的差值又下降到 0.35mm。第四列长度比数值的变化也是无规则可循的。图 1 以折线的形式描述字号之间的尺寸关系,显然这是一条无规则变化的曲线。因此,在字号和毫米之间不存在一个转换函数。字号内部及号-点之间也没有统一的转换公式。

表 2 相邻字号的长度差和长度比

字号	尺寸/mm	长度差/mm	长度比
八号	1.76	0.17	1.10
七号	1.93	0.35	1.18
小六号	2.28	0.36	1.16
六号	2.64	0.52	1.20
小五号	3.16	0.53	1.17
五号	3.69	0.53	1.14
小四号	4.22	0.70	1.17
四号	4.92	0.35	1.07
小三号	5.27	0.35	1.07
三号	5.62	0.71	1.13
小二号	6.33	1.40	1.22
二号	7.73	0.71	1.09
小一号	8.44	0.70	1.08
一号	9.14	3.51	1.38
小初号	12.65	2.11	1.17
初号	14.76		

着与其不一致的其他号制版本。例如,北大方正的一号字就定为 9.657mm^[6],不同于微软 Word 一号字的 9.14mm。号制的种种缺陷使得有些学者甚至建议电子排版采用点制而不用号制^[7]。



然而点制也并不完美。同“字号”一样,点也是为满足印刷排版的需要而建立起来的新度量体系。对于人们的认知活动来说,多一套体系意味着多一份负担。由于点是排版专用的度量单位,一点所代表的长度必须换算为人们所熟悉的一般度量单位才易于被理解。令人遗憾的是,较精确的转换因数相当冗长。一般教科书常将作为字形度量单位的一个点解释为“大约相当于 1/72 英寸的高度”^[7]。较为精确的数据是“1pt=0.351 46mm”^[8]。更准确的换算公式有 1pt=0.351 459 8mm,约等于 1/72.272 英寸^[9]和 1pt=0.996 264/72 英寸^[9]等等。如此种种,显然是不便掌握的。

点制的另一个缺点是,尽管上面介绍的英美点制影响最大,但它不是唯一的点制标准。欧洲还使用着法国的 Didot 点: 1pt=0.375 9mm。此外,还有 TeX 标准的 1pt=0.351 459 803 5mm, Postscript 的 1pt=0.352 777 777 8mm 和 l'Imprimerie nationale 的 1pt=0.4mm^[8]。这种多点制并存的现象也是不利于应用的。

号制和点制的种种缺点不能归咎于他们的设计者,如果考虑到早期铅字印刷时代的科技水平和社会需要,就应该说这两类字形度量标准在当时还是很完善的。例如,在使用传统字模的时代,最小的印刷体汉字是 7 号,因为再小就没法刻制字模了,就是能刻出来,也难以把字印清楚。而且,增加字号意味着需要更多铸字材料和更复杂的印刷设备。然而,现在是科技高度发达的计算机信息时代,我们完全有条件而且有必要考虑字形度量体制的进一步改善,以便更好地服务于社会。

笔者建议以毫米(mm)为基本单位来统一字形度量标准。技术实现上,精确度一般可定为 0.1mm,并将常用 mm 字形规格列为选项,其余的由用户根据需要来填写,类似目前 Word 点制的处理方法。这样,既能克服点制和号制的种种缺陷,又能保留他们原来的好处,大大简化字形处理。我们还可以进一步将排版印刷中使用的英制等其他各种尺寸说明都统一到米制上来。例如,把排版说明“论文中的中文摘要用宋体五号,英文摘要用 Times New Roman 9 点,版心尺寸 6.5*9.5 英寸,文章内的图表宽度不超过 12cm。”统一用米制表达成“论文中的中文摘要用 3.7mm 宋体字形,英文摘要用 3.2mm Times New Roman 字形,版心尺寸 165mm*240mm,文章内的图表宽度不超过 120mm。”这样显然简便多了。米制是全球通用的国际标准,是中国的法定计量体制,在欧洲和日本已经有人在排版中用米制代替点制^[7]。看来米制统一一点制和号制乃至所有长度度量标准将成为大势所趋。

2 应用

分别使用 KFD、阈值算法、BCV (Between class variance) 算法对以下岩石节理裂隙彩色图像进行分割。



图 2 岩石节理裂隙彩色图像

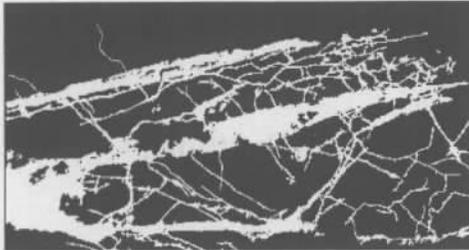


图 3 使用 KFD 分割的岩石节理裂隙



图 4 使用阈值算法分割的岩石节理裂隙



图 5 使用 BCV 算法分割的岩石节理裂隙

(上接 177 页)

6 结论

号制和点制是中文和英文字形尺寸的最常用度量方法。本文介绍了这两种方法的历史和现状,给出了一个由一手实验数据建立起来的“号-点-厘米”对照表,并介绍了对照表的用途。为了方便读者和有关专家检查数据的可靠性,文章还交待了实验的方法。最后,我们分析讨论了号制和点制的种种缺陷,并提出文字排版全球统一使用米制度量标准的初步设想,为语言文字现代化建设提供参考。(收稿日期:2005年9月)

参考文献

1.朱永和.排字的点数制与号数制探讨[J].编辑学报,1999,11(2):69-71

可以看出,使用 KFD 能更完整地分割出岩石节理裂隙,更好地去除噪声,达到了较好地分割效果。

3 总结

由于图像的多样性,使得每一种图像分割算法的适用范围很有限。岩石节理裂隙图像中,节理裂隙形状复杂、噪声很多,这样就使得一般分割算法的分割效果不太理想。而统计模式识别法在分割一幅岩石节理裂隙图像时,可以学习该图像的不同样本的特征,并且利用这些特征对图像中的每个像素进行分类,进而实现分割,所以该方法具有较广的适用性和较好的分割结果。

本文中,使用了基于核函数 Fisher 判别法 (KFD) 构造了统计模式识别的分类器。KFD 可以将高维的属性空间上的非线性判别转化为特征空间上的线性判别,而不需要知道从属性空间到特征空间的具体映射形式。

通过对岩石节理裂隙图像的分割,取得了良好的效果。

(收稿日期:2005年9月)

参考文献

- 1.张新峰,沈兰荪.图像分割技术研究[J].电路与系统学报,2004,9(2):92-99
- 2.Milan Sonka,Vacav Hlavac,Roger Boyle.Image Processing,Analysis, and Machine Vision[M].北京:人民邮电出版社,2002
- 3.边肇祺,张学工.模式识别[M].第二版,北京:清华大学出版社,2000:83-90
- 4.Anil K Jain,Robert P W Duin,Jianchang Mao.Statistical pattern recognition[J].IEEE Transaction on Pattern Analysis and Machine Intelligence,2000,22(1)
- 5.田有文,李成华.基于统计模式识别的植物病害彩色图像分割方法[J].吉林大学学报(工学版),2004,34(2):291-293
- 6.卢力,田金文,柳健.统计模式识别研究进展[J].军民两用技术与产品,2003(11):39-42
- 7.许建华,张学工,李衍达.应用核 Fisher 判别技术预测油气储层[J].石油地球物理勘探,2002,37(2):170-174
- 8.李映,焦李成.基于核 Fisher 判别分析的目标识别[J].西安电子科技大学学报(自然科学版),2003,30(2):179-182
- 9.Mjka S,Ratsch G,Weston J et al.Fisher discriminant analysis with kernels[J].Neural networks for Signal Processing IX,New York:IEEE Press,1999:41-48

- 2.刘岱伟,潘宝骏,阙少聪.科技书刊中英文字体字号的编辑处理[J].编辑学报,2001,13(1):23-24
- 3.McLean R.The Thames and Hudson Manual of Typography[M].London:Thames and Hudson Ltd,1980
- 4.国家技术监督局.GB/T12200.2-94 汉语信息处理词汇 02 部分:汉语和汉字[M].北京:中国标准出版社,1994
- 5.翟铭,杨新岚.当代排版技术概论[M].北京:印刷工业出版社,1994
- 6.楠天健.电子排版技术[M].北京:清华大学出版社,1995
- 7.Shelly G B,Cashman T J,Vermaat M E.Discovering Computers 2004: A Gateway to Information[M].Boston:Course Technology,2003
- 8.Vakulenko A.Difference between point systems.http://www.oberon-place.com/dtp/fonts/point.htm,2000
- 9.Clair K.A Typographic Workbook:A Premier to History,Techniques, and Artistry[M].New York:John Wiley & Sons,Inc,1999