

维吾尔、哈萨克、柯尔克孜文键盘标准制定研究*

迪丽尼格尔·热夏提¹, 吾守尔·斯拉木^{1,2†}, 艾尼宛尔·托乎提^{1,2}, 米尔阿迪力江·麦麦提¹

(1. 新疆大学 信息科学与工程学院, 新疆 乌鲁木齐 830046; 2. 新疆大学 新疆多语种信息技术重点实验室, 新疆 乌鲁木齐 830046)

摘要: 为进一步统一与完善维哈柯文各种系统软硬件的键盘布局, 在传统键盘标准的基础上, 本文经过多次讨论及分析, 重新确立了符合GB12510-90标准的新的键盘布局. 并且按照这些新规定的原则制定出了新的pc键盘布局及数字键盘布局. 此外, 还有新出的各种嵌入式软硬件的键盘布局标准. 文中给出了这些新制定的键盘布局图, 并通过例子进一步验证了其有效性.

关键词: 键盘布局; PC键盘; 数字键盘; 嵌入式软硬件键盘

中图分类号: TP391.1 **文献标识码:** A **文章编号:** 1000-2839(2014)01-0102-06

Research on the Formulation of Standard Keyboard of Uyghur, Kazakh, Kirgiz

Dilnigar Reshat¹, Wushour Silamu^{1,2}, Aniwer Tohti^{1,2}, Miradiljan MeMet¹

(1. College of Information Science and Engineering, Xinjiang University, Urumqi Xinjiang 830046, China;

2. Key Laboratory of Xinjiang Multilingual IT, Xinjiang University, Urumqi Xinjiang 830046, China)

Abstract: In order to further unify the both software and hardware keyboard layout of various systems in Uyghur, Kazakh, Kirgiz, on the basis of the traditional keyboard standard, via great number of discussion and analysis reformulated the new layout principles in accordance with GB12510-90 and new PC keyboard layout, digital keyboard layout, in addition to the new standard keyboard layout of various embedded software and hardware. This article presents the new formulation of the keyboard layout, and its validity is further verified by examples. It will play an important role in promoting the development of Xinjiang minority information science.

Key words: keyboard layout; PC keyboard; digital keyboard; embedded hardware software keyboard

0 引言

新疆维吾尔自治区是多民族聚集区域, 其中维吾尔族、哈萨克族和柯尔克孜族约占总人口的60%. 维哈柯文无论在政府办公, 还是在人民日常生活中都有着广泛的应用. 因此进行维哈柯文信息化技术研究并且推出其标准, 对促进新疆的全面发展, 提高各民族的教育科学文化素质等都具有重要意义^[1]. 针对信息处理的应用需求及维哈柯文文字特点, 本文研究了信息处理中关键技术之一的“维哈柯文键盘制定标准”, 对新疆地区信息化技术的进一步发展具有一定的促进作用.

近十年来, 各种维哈柯文嵌入式系统和产品及其应用越来越多, 因此数字键盘的进一步规范化、标准化成了一件急需解决的重大问题. 故在这次制定的标准中我们还提出了维哈柯文数字键盘的布局标准.

据我们调研, 市场上的各种维哈柯文系统软件键盘布局基本上是一致的, 但也有一些差异. 并且初次制定的键盘布局标准有一些欠缺, 因此我们希望通过工信部的项目“维哈柯文软件标准体系研究及标准制定”重新讨论并进行修改, 制定发布了更加科学、合理、规范的维哈柯文键盘布局标准.

* 收稿日期: 2013-04-08

基金项目: 国家工信部项目(维哈柯文软件标准体系研究及标准制定-159016).

作者简介: 迪丽尼格尔·热夏提(1989-), 女, 维吾尔族, 硕士生, 研究方向为自然语言处理.

† 通讯作者: 吾守尔·斯拉木, 中国工程院院士, wushour@xju.edu.cn.

1 键盘标准的制定原则

标准键盘是计算机最基本也是最直接的外部设备之一。设计计算机键盘需要综合很多的因素,例如字母在书面表达中出现的概率、人体生理结构(肩膀距离、手指灵活程度)、左右手负荷、两手交替频率等。其中,各个字母在书面表达中出现的频率是一个关键的因素。

目前最常见且通用的布局方案是最初由Christopher Latham Sholes于1868年提出的Qwerty键盘。随后August Dvorak依据平衡双手负担、常用字母出现在中排、更有力的手指负担更多输入等原则,在1932年设计了新的键盘布局,从而提高了打字效率。Dvorak键盘布局理论指出:

- (1)双手交替打字,且右手和左手的负荷应该是相等的;
- (2)高频优先,最大化中间行(Home Row)的负荷,即中间行累计频率最高,其次是上面行和下面行;
- (3)两手顺序交替的频率最大化,并且相同手指的频率最小化^[2]。

近年来随着电子设备的不断普及及移动电子设备的发展,关于键盘设计又有人提出很多新的想法^[3]。然而,维吾尔文键盘布局国家标准早在1989年通过鉴定,1990年发布;哈萨克文键盘布局国家标准在1992年和哈萨克文字符集代码标准一起通过鉴定,国家标准局也已接受并通过了审批,由于排版印刷等原因拖了几年没能正式发布。当时同样讨论制定柯尔克孜文键盘布局标准,但是没有正式申报地方技术监督局审批发布。以上三种文字键盘布局标准已使用了20年左右,对开发应用的各种维哈柯文信息处理软件和系统的标准化、规范化起到了重要作用。但是,随着信息化技术的不断发展,统一规划及完善维哈柯文键盘布局已经成为社会各界迫切需求的重要任务之一。

总之,最新标准规定了面向GB 12050-1989^[4]、GB 21669-2008^[5]和GB 13000-2010^[6]中维吾尔文字符的通用键盘字母数字区布局;面向GB 21669-2008^[5]和GB 13000-2010^[6]中哈萨克文及柯尔克孜文字符的通用键盘字母数字区布局。除此还有,数字键盘还可以输入GB 12050-1989^[4]“信息处理信息交换用维吾尔文编码图形字符集”、GB 21669-2008^[5]“信息技术维吾尔文、哈萨克文、柯尔克孜文编码字符集”和GB 13000^[6]“信息技术通用多八位编码字符集(UCS)”中规定的全部维吾尔文字符;以及数字键盘可以输入的GB 21669-2008^[5]“信息技术维吾尔文、哈萨克文、柯尔克孜文编码字符集”和GB 13000^[6]“信息技术通用多八位编码字符集(UCS)”中规定的全部哈萨克文,柯尔克孜文字符等。

目前计算机键盘布局主要是参考英文字母,也就是说按26个字母来布局的。但是维吾尔文有32个字母,哈萨克文有34个字母,而柯尔克孜文中有30个字母。因此,制定这三个文字键盘布局时,还需考虑在一个键上放置几个字符,在逻辑上可以将这些字符分成不同的层。即引入主键盘与辅助键盘层等两层,采用换挡方式可以选择上下档对应层的字符。如下表1所示这两层的分配标准。

表 1 层的分配

层选	激活层
不用层选 (缺省层1)	层1(主键盘)
上档键选择(按shift键)	层2(辅助键盘)

在1989年发布的维吾尔文键盘标准方案考虑了以下三点因素:(1)重点考虑字母使用频率,使用频率高的字母安排在键盘中间位置;(2)尽量靠近英文键盘布局;(3)考虑十指操作方便。

当今信息技术的快速发展促进了少数民族信息化事业的发展。因此,考虑到社会各界对少数民族更加完善、合理、科学的信息技术标准的迫切需求。在国家工信部重点项目的多次讨论会,达成在保持上述三个原则的基础上对维哈柯文各类键盘标准原则进行了补充的共识,重新制定了以下几个原则:

1.1 重点考虑字母使用频率,使用频率高的字母安排在键盘中间位置

计算机键盘的每个键位都有其属性,而每种文字(音节文字)的音素或部件也各有其属性。如何将这两种属性恰到好处地使之“对号入座”,是键盘布局设计的基本要求^[7]。因此本文依照维哈柯文字符使用率高到低的顺序原则进行布局设计。

1.2 尽量靠近英文键盘布局

目前世界各个语言的键盘布局大都通用英文键盘布局,这样有利于记忆,并且在各个语言切换时也比较方便.

1.3 考虑十指操作方便

用户使用计算机越方便越好,而键盘作为最直接的外部设备,其布局对计算机的操作性能,以及用户对计算机的使用感受具有重大影响.

1.4 维、哈、柯文键盘布局尽量一致

维哈柯这三种文字差别不是很大,都属于阿尔泰语系中的突厥语系,而且在多民族居住的新疆是主要使用的少数民族语言,因此它们的键盘布局也应该要基本一致,这样对各民族间相互使用、交流与学习有一定好处.

1.5 数字键盘的字母按英文发音相似性安排在同一位置上

数字键盘布局键位不多,为了便于记忆和使用,按照国际通用性,应以英文字母作参考.

1.6 嵌入式硬键盘的布局跟pc键盘一样

由于嵌入式硬键盘本身就跟pc键盘一样,因此其键盘布局也完全按照pc键盘布局原则进行布局.

1.7 软键盘的布局保留pc的位置

除此之外,还可以实现一键一字母和字形自动选择的功能.

2 维哈柯文通用的pc标准键盘

2.1 标准维哈柯文pc键盘布局

图1~6是按照上述原则经过多次讨论与分析制定出的符合GB12510-90^[8]标准的维哈柯文pc键盘布局图.



图 1 维吾尔文pc键盘直档键(主键盘)



图 2 维吾尔文pc键盘按shift键时(辅助键盘)



图 3 哈萨克文pc键盘直档键



图 4 哈萨克文pc键盘按shift键时

2.2 验证字母频率的影响

从下面的例子可以看出本标准制定过程中,作为关键因素的字母频率对键盘布局的影响.表2所示是利用自制字符统计软件对维吾尔著名历史长篇小说“故乡-母亲”中字符出现的频率统计的结果^[9].这篇长篇小说分为三册,共有2 397 901个字符.



图 5 柯尔克孜文pc键盘直档键



图 6 柯尔克孜文pc键盘按shift键时

表 2 字符频率统计结果

序号	字母	出现次数	频率(%)	序号	字母	出现次数	频率(%)	序号	字母	出现次数	频率(%)
1	空格	380452	15.7348%	20	ش	35020	0.014436	39	«	1558	0.0641%
2	ى	288024	11.8923%	21	ي	34510	1.4295%	40	ف	1543	0.0631%
3	ا	176490	7.2903%	22	ك	31192	1.2901%	41	1	718	0.0293%
4	ه	119760	4.9448%	23	ء	31100	1.2871%	42	9	394	0.0160%
5	ن	118284	4.8860%	24	غ	29630	1.2255%	43	4	353	0.0144%
6	ل	113583	4.6884%	25	ز	29235	1.2053%	44	2	251	0.0102%
7	ر	92705	3.8273%	26	.	26613	1.1025%	45	3	236	0.0097%
8	ۇ	80199	3.3130%	27	ق	26169	1.0800%	46	8	157	0.0064%
9	ت	77708	3.2087%	28	چ	25163	1.0386%	47	5	151	0.0060%
10	د	70422	2.9085%	29	گ	20848	0.8620%	48)	133	0.0054%
11	ئ	68146	2.8157%	30	ء	17943	0.7399%	49	5	127	0.0052%
12	م	67879	2.8028%	31	-	17235	0.7125%	50	(121	0.0049%
13	قى	67146	2.7731%	32	ق	16236	0.6701%	51	6	113	0.0044%
14	ي	57785	2.3865%	33	ق	15458	0.6372%	52	7	101	0.0040%
15	ك	51666	2.1311%	34	خ	14150	0.5829%	53	0	88	0.0036%
16	ب	47881	1.9752%	35	ج	11807	0.4835%	54	ء	37	0.0014%
17	س	42647	1.7637%	36	!	5349	0.2210%	55	>	2	0.0001%
18	و	40013	1.6485%	37	?	2763	0.1158%	56	<	2	0.0001%
19	پ	38620	1.5966%	38	:	1983	0.0823%	57	...	2	0.0001%

将表2统计结果与图1、2对比可以看出,字符出现频率的高低在一定程度上决定键盘的布局.字符根据其使用频率的大小决定其在键盘中的位置:频率高的安排在中间位置,同排中优先排右键区,其次排左键区.出现频率最高的字母布局在键盘的最中间位置,这样用户手指操作更加方便.此外将有些使用率很低的字母安排在辅助键盘,也就是按shift键的同时才能显示出来.

3 嵌入式设备数字键盘标准

随着嵌入式设备的广泛应用,市场上出现了一些以数字键盘为主的,能处理维哈柯文信息的电子产品,例如手机、遥控器、数据采集器等.根据其采用的键盘布局类型,可以归纳以下几种:

(1) 按字母表的顺序安排;

特点:突出了维哈柯文的词典顺序.

缺点:记忆字母顺序不方便,容易出错.

(2) 跟PC键盘布局一致;

特点:跟PC一样,容易记忆.

缺点:有些键盘任务太大,八个键上输入不太方便.

(3)按英文字母音顺序安排.

特点:英文和维吾尔文在发音上比较相似,有些与英文发音不相同的字母根据维哈柯文新文字的特点安排,这样的键盘布局较方便记忆.

缺点:有些按键上安排的字母多,有些字母输入繁琐.比如要输入这个字母需要按5次6号键才能输入,如果用选择式输入只需要两次按键就可以;按键使用率不均匀.

除此之外,对于使用阿勒泰语系的语言,其基本字母根据其在单词中的位置不同而呈现出不同的形态.这给用户输入和显示带来了很大的困难,这也是设计键盘布局时需重点考虑的问题之一.为了解决这些问题,针对有字母变形的语言,设计了相应的变形引擎,引擎根据用户输入的字母位置信息,通过完整的变形规则,自动选择正确的形式输出显示,不需要麻烦用户自己选择字母形态,简化了输入过程,提高了速度.

经过对上述各种布局原则进行分析和讨论,同时参考其它形式的键盘布局,本文决定使用以下基本标准原则来布局数字键盘:

(1)维、哈、柯文键盘布局尽量一致;(2)与英文发音相似的安排在一个位置上;(3)对于英文里没有的字母,按它的源语言发音相似的字母安排在同一个位置,即遵守“一键多符”原则.

3.1 嵌入式设备数字键盘布局

根据上述原则,制定了以下符合GB12510-90^[8]标准的嵌入式设备数字键盘布局标准方案.本文选择图7~9中所示的“1-2-3”式维哈柯文数字键盘的键位布局为例.除此之外,“7-8-9”式的数字键盘布局也跟图2所示的字符键位一样,只是数字顺序不一样.



图7 维吾尔文数字键盘布局

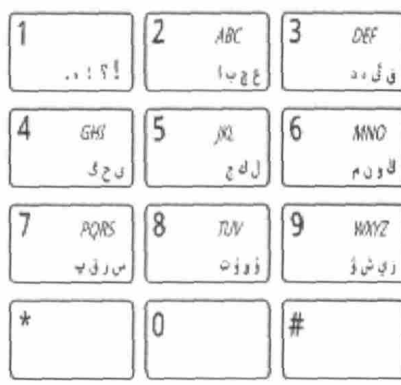


图8 哈萨克文数字键盘布局



图9 柯尔克孜文数字键盘布局

3.2 嵌入式设备硬键盘布局

除了数字键盘之外,在一些电子产品中也使用了嵌入式设备硬键盘.此类产品的键盘布局跟pc完全一样,也即用pc键盘的布局原则进行布局.

3.3 嵌入式设备软键盘布局

最新一系列产品上采用的不是物理键盘,而是通过软件生成的软键盘,比如:iphone、ipad、windows mobile等产品都采用着这样的软键盘.因此,根据市场需求,经对此类键盘布局讨论制定出了如下布局原则:(1)尊重pc键盘标准,并且保持基本一致;(2)通过软件扩展键盘区域,实现一键一字母输入方式,并自动选形;(3)保持原来的标点符号,修改等符号.

图10~12所示的就是根据上述原则设计的维哈柯文软键盘布局图.

4 结语

键盘作为计算机的主要输入设备,其设计与改进在计算机日益普及的今天尤为重要,随着信息时代

的快速发展,少数民族在政府办公、日常学习及生活中对信息化的需求也不断增强.键盘标准是计算机与信息领域中最基本的、也是最关键的标准之一.



图 10 维吾尔文软键盘布局



图 11 哈萨克文软键盘布局



图 12 柯尔克孜文软键盘布局

虽然维哈柯文键盘标准推出较早,但是因为标准并不完善、在市场上使用不统一等各种因素的影响,给维哈柯文信息化及标准化造成了一定的障碍.本文经过研究分析,制定出了统一的维哈柯文各类键盘的布局标准,对各民族共享信息时代的技术成果有重要的促进作用.

参考文献:

- [1] 张伟,任大明,许晓辉等.维,哈,柯文显示及排版技术研究[J].计算机应用研究,2008,25(3):836-844.
- [2] 田微,王江晴,朱宗晓等.女书计算机键盘布局与输入法的研究[J].中文信息学报,2010,25(5):124-126.
- [3] 张俊琳,崔勇,王弘毅.标准键盘布局比较与优化设计[J].计算机系统应用,2012,21(4):254-259.
- [4] 中华人民共和国国家标准.信息处理信息交换用维吾尔文编码图形字符集[S].GB 12050-1989.北京:中国标准出版社,1991.
- [5] 中华人民共和国国家标准.信息技术维吾尔文、哈萨克文、柯尔克孜文编码字符集[S].GB 21669-2008.北京:中国标准出版社,2008.
- [6] 中华人民共和国国家标准.信息技术通用多八位编码字符集(UCS)[S].GB 13000-2010.北京:中国标准出版社,2010.
- [7] 卢亚军.藏文计算机通用键盘布局与输入法研究[J].中文信息学报,2006,21(2):78-86.
- [8] 中华人民共和国国家标准.信息技术通用多八位编码字符集(UCS)[S].GB12510-90.北京:中国标准出版社,1991.
- [9] 明日科技.C# 技术大全[M].北京:人民邮电出版社,2011,650-652.

责任编辑:闫新云