

11 申请号 88102605.0 [51] Int.Cl.⁴
G06F 3/023

74 专利代理机构 广西壮族自治区专利服务中心
代理人 谭志涛

说明书页数: 3

附图页数:

[54]发明名称 普及型汉字编码“杞”码法

[57]摘要

本发明是一种普及型的汉字编码法,全部汉字笔形可分为横、竖、撇、捺、左向折、右向折六种,恰好为“杞”字的笔形,故称“杞”码法。其特点是把精选的常用汉字部首字母化,笔形数字法,构成输入码。

本“杞”码编码规则简单,重码较少,易学、易记、易用,定码速度快,不受读音限制,对任何汉字都能见字识码,不需专门训练便可很快学会并正确使用,因此是普及型的汉字编码法。

< 39 >

权 利 要 求 书

1. 一种普及型的汉字编码法，其特征在于编码原则是精选35个汉字中高频出现的部首字母化，“杞”码笔形取横1、竖2、撇3、捺4、左折5、右折6代码数字法。

2. 如权利要求1所述的“杞”码法，其特征在于所述的35个汉字部首及其代码为表3所示。

3. 如权利要求1所述的“杞”码法，其特征在于所述的汉字“杞”字笔形及其代码为表1所示。

4. 如权利要求1所述的“杞”码法，其特征在于所述的汉字书写笔顺基本规律为表3所示。

5. 如权利要求1—4所述的“杞”码法，其特征在于“杞”码取码是无表2所列常见部首的，按写字的笔顺依表1取五笔（前三笔加末二笔），不足五笔的全取；有表2所列常见部首的，先取相应部首代码，后按写字的笔顺依表1取四笔（前二笔加末二笔），不足四笔的全取。

普及型汉字编码（“杞”码）法

本发明属汉字编码方法，特别是“杞”（音jié）字笔形代码。其编码规则是：取汉字起笔的常用部首读音字母加四笔（前二笔加末二笔）“杞”字的笔形代码，如果汉字起笔不在所收的部首内，则取五笔（前三笔加末二笔）“杞”字笔形代码。“杞”字的笔形代码是指横1、竖2、撇3、捺4、左折5、右折6，这六个代码正好是“杞”字的笔形代码（简称“杞”码）。

目前的汉字编码输入方案很多，除“拼音码”和“五笔画”码容易掌握、规则简单外，其余的汉字编码都需经过较长时间的培训才能正确使用。但“拼音码”对读不准音和不会读的字输入存在较大的困难，且还要使用简码，人为地增加了使用规则，从而增加使用者的难度，重码又较多。王永民的“五笔字型”输入法分两种方案：

1、“五笔画”简易输入法，即每个字取前四笔和末笔的笔形代码，不足五笔的以“0”代，虽易学、易记、易用，适合于初学者使用，缺点是重码多；

2、“五笔字型”高效输入法，需学习熟记130个字根的位置和汉字的拆分规则，只供经培训的专业人员用。

以上这些汉字编码方案，远还达不到广泛、普及使用的要求。

本发明提供了一种普及型汉字编码法——“杞”码法，其编码规则简单，重码较少，易学、易记、易用，定码速度快，不受读音限制，对任何汉字都能见字识码，查字快速，具有小学三年级以上文化的使用者，很快就能熟悉并正确使用。

本发明的方法是将汉字部首字母化，即精选35个汉字中出现最多的部首，赋予其部首字母读音代码作码元（见表2汉字部首代码表）；汉字“杞”码笔形数字法，即把汉字的笔形分六种：横（包括提）、竖、撇（包括左向点）、捺（包括右向点）、左折、右折，

分别以1、2、3、4、5、6的代码表示（见表1汉字“杞”字笔形代码表）。

“杞”码取码方法：

1. 字的起笔（即书写时的第一笔）不在表2所列常见部首内时，按写字的笔顺依表1取五笔（前三笔加末二笔）“杞”字的笔形代码，不足五笔的全取。

例：部——41452，干——112，乏——3454，盛——13521；

2. 字的起笔在表2所列常见部首内时，先取相应部首代码，后按写字的笔顺依表1取四笔（前二笔加末二笔）“杞”字的笔形代码，不足四笔的全取。

例：萍——C4412，逛——G1154，马——M，伐——R1634；

3. 汉字的笔顺从习惯，一般以《辞海》为准，基本规律是：先横后竖，从左到右，从上到下，先外后里，封口在后，捺笔在后，点笔在后，先横后撇，左右先写，中间先写，横勾先写（见表3汉字书写笔顺基本规律表）；《辞海》没有明确的，参考《汉字正字小字汇》。有些字的笔顺不容易确定，如果你确定的笔形码里找不到要找的字，可以在另一笔形码里查找，字的笔形码可参考《汉字快速查字手册》。

附表如下：

表1 汉字“杞”字笔形代码表

代码	笔形名称	笔形	同类笔形
1	横、提	→	— ✓
2	竖、垂	↓	丨
3	撇、左向点	↙	ノ ノ ノ ノ
4	捺、右向点	↘	㇏ ㇏ ㇏
5	左折	↵	丿 ㇇ ㇈ ㇉ ㇊ ㇋
6	右折	↶	㇏ ㇐ ㇑ ㇒ ㇓ ㇔

表2 汉字部首代码表

表3 汉字书写笔顺基本规律表

代码	部 首		规 律	例 字	笔形码
B	贝 广 宀	1	先横后竖	干	112
C	艹 虫	2	从左到右	则	B25
G	广 彡			树	M5454
H	火 禾	3	从上到下	岚	S3634
J	钅 (金)			笔	Z3116
K	口 口	4	先外后里	区	1634
M	门 马 木 目	5	封口在后	图	K3544
N	女	6	捺笔在后	农	35334
R	亻 彳 日 日	7	点笔在后	戈	1634
S	纟 纟 石 山	8	先横后撇	厂	13
T	土 扌	9	左右先写	半	43112
W	王			炎	H3334
X	丨	10	中间先写	兆	36434
Y	衤 (衣) 鱼 月			非	22111
Z	足 (足) 竹 (竹)	11	横勾先写	皮	53254

本“杞”码的码长，最短一位，最长五位，平均4.86位，对《GB2312—80》的全部一、二级6763个汉字编码统计，使用码数共3497个，平均每个码与1.93个汉字相对应，重码较少，其中无重码字占总数的31.76%，一码对应五字以下占总数的78.6%，一码对应十字以上的小于10%；对于一码对应二字以上的重码字，作为计算机的汉字输入码时，可按高频先见或采取对号选择的方法得到。并与传统的笔划笔顺、部首、四角号码、汉语拼音等编码比较，具有简单、易学、易用、定码速度快、重码较少、查字快速的一种普及型汉字编码。适用于：

- 1、可供各类人员特别是非专业人员作为汉字输入码在计算机、电子打字机使用；
- 2、可供广大学生、干部、科技人员、工人，作为汉字检字法，用于各种字、词典等工具书作正文或索引的编排，提高查检效率；
- 3、可供图书情报部门用于中文文献字顺目录的组织。