[19]中华人民共和国专利局

[11]公开号 CN 1069350A

[12]发明专利申请公开说明书

[21] 申请号 91104053.6

[51] Int.Cl[5]

G06F 3/023

[43] 公开日 1993 年 2 月 24 日

[22]申请日 91.8.4

[71]申请人 黄文智

地址 353204 福建省顺昌县高阳采育场南桥

[72]发明人 黄文智

[74]专利代理机构 福建省专利服务中心

代理人 丁秀丽

说明书页数: 13 附图页数: 1

[54]发明名称 结构型汉字编码输入法

[57]摘要

本发明涉及一种结构型汉字编码输入技术，特别涉及在计算机、打字机及类似装置上使用的汉字编码输入方式的文字处理系统。本发明是将汉字分解为独体字、部首和笔形三部分，利用 0—9 十个数字键对汉字进行编码。利用本发明方法不需经过专门的指法训练，就能实现对汉字及其词汇进行直观快速的编码输入，做到见字识码，会写就会输入，且简明易学，记忆量少，本编码的输入法对简体字和繁体字均可适用。

(BJ)第 1456 号

权　利　要　求　书

1、一种结构型汉字编码输入法，特别是在计算机、打字机及类似装置上使用汉字编码输入法的文书处理系统，在该文书处理系统中的键盘上至少包括下列操作键：

功能键，即用于决定操作时编码的输入方式的键；

数字键，即用于输入汉字编码的键（一般只采用0—9，10个数字键）；

选字键，即用于对提示行上的字进行选择的键（包括0—9，10个数字键以及26个相对于汉字读音的字母键）；

空格键，即用于要求继续提示的键；

清除键，即用于作废某一次字码输入的键；

对于错误码或不存在的字码的输入，系统会自动作废该次输入，并发出信号，等待新的输入，

本发明的特征在于，在对上述数字键进行操作时，所采用的汉字编码输入方法为：

A、将独体字的编码按照：第一码取该字的笔画，即字码，第二码补零，第三码取该字的起笔笔形代码，第四码取该字的收笔笔形代码，这样一种四位码的方式输入；

或B、将合体字的编码按照笔形的书写顺序“读”出组成合体字的各个部件（包括独体字、部首或笔形）的笔画（即

字码）或笔形代码组成的一组四位码作为该合体字的编码。

a、当一个合体字的部件超过四部时：按汉字的结构取前四部笔画编码，

b、当一个合体字的部件刚好四部时：按汉字的结构顺序依次取四部笔画，构成汉字的四位编码，

c、当合体字的部件为三部时：按汉字的结构顺序依次取三部笔画，第四码取该字末部的收笔笔形代码，如：

d、当合体字的部件为二部时：按汉字的结构顺序依次取二部笔画，第三、四码取末部的起笔及收笔笔形代码，

或C、将词汇编码按照如下方式输入：

a、双字词：取每个单字的前二码，组成四位码，如果词汇中的字为独体字，则按该独体字的笔画取两个相同的重码，

b、三字词：取前两个字的第一码，最后一个字的前两码，

c、四字词：每个字各取其第一码，组成四码，

d、多字词：取前四个单字的第一码，组成四码。

2、根据权利要求1所述的结构型汉字编码输入法，其特征在于：上述代码代表的笔形为：

0———L ノ ㇇ レ

9——丨 ㇕ ㇉ ㇋

8——㇏ ㇀ ㇏ ㇛

7——丿 ㇇ 彐 ㇉

6——乙 乚 ㇆ 一 了 ㇋ ㇉ 乞

3、根据权利要求1所述的结构型汉字编码输入法，其特征在于：重码的处理方法是在按本编码法输入编码后，再输入所需汉字的头一个拼音字母，此字即首先显示在屏幕正常的编辑位置上，不需作任何处理，即可正常继续键入。

结构型汉字编码输入法

本发明涉及一种汉字编码输入技术，特别涉及在计算机、打字机及类似装置上使用的汉字编码输入方式的文字处理系统。

随着微型计算机在各个领域的广泛使用，国内外的研究人员通过对微型计算机上流行的行编辑程序以及字处理系统进行改造，已将其用于对汉字的输入与编辑。在汉字输入方法上应用面较广泛的几种汉字操作系统中，主要配备了拼音、联系、区位、国标、电报、首尾、五笔字型等汉字输入方案。其中“五笔字型”是目前国内影响较大，应用面较广，公认使用较为方便合理的一种汉字输入法。虽然“五笔字型”汉字输入方法比较直观易学，输入速度也较快，但由于其必须使用二十五个字母键来表示１２５个字根，在每个键位上一般安排有２至６种字根，键入字根时，需要查找各基本字根所在的键位，操作者在指法上必须经过较长时间的专门训练、熟练记忆并掌握各字根所在的键位，才能达到一定的操作速度和实现盲打，训练时间较长，记忆量较大，不便于迅速普及推广。

本发明的目的在于提供一种操作更为简便迅度的汉字编码输入方法。

当我们对汉字字形加以分析，就能得出这样的结论，有的汉字可以分析出几个具有独立意义的汉字，有的汉字却不能再行分析，那些不能再行分析的具有完整形体和独立造字意图的汉字，我们称之为独体字。而那些由两个或两个以上的独体字按一定的组合规律组合而成的，有一定造字意图的汉字，称之为合体字。部首是指不具有独立造字意图的，但具有完整的形体，与独体字能够构成合体字的单位，部首中有的是独体字的变形，更多的则是纯粹的结构符号。当我们把部首和独体字称为字根时，也可以说汉字是以字根为单位，采用积木方式拼合起来的，而字根又由笔形构成。也就是说把每个字根机械地进行分析，就可得到若干的笔形，每个字根都具有一定数目的笔形，笔形的数量我们称之为笔画。

本发明就是利用了汉字组合的这一特征，设计出结构型汉字编码方案的。

本发明是这样实现的。

本发明涉及一种在计算机、打字机及类似装置上使用汉字编码输入方式的文书处理系统，在该文书处理系统中的操作键盘上至少包括下列操作键：

功能键，即用于决定操作时编码的输入方式的键；

数字键，即用于输入汉字编码的键；

选字键，即用于对提示行上的字进行选择的键；

空格键，即用于要求继续提示的键；

清除键，即用于作废某一次字码输入的键；

对于错误码或不存在的字码的输入，系统会自动作废该次输入，并发出信号，等待新的输入，

本发明特征在于：在对上述数字键进行操作时，所采用的汉字编码输入方法为：

一、将独体字的编码按照：第一码取该字的笔画，即字码，第二码补零，第三码取该字的起笔笔形代码，第四码取该字的收笔笔形代码。

上述编码中所提到的笔形是指构成现代汉字的各种点和线的形状。为了适合汉字编码技术的需要，我们将各种笔形概括为五种笔形，即横、竖、捺、撇、勾。依其使用频度，设立相应的代码为0、9、8、7、6。即下列笔形配上下列相应的笔形为：

0——横—— 一 乚 丿 乛 丶 乚

9——竖—— 丨 ㇇ ㇉ 亅

8——捺—— 乀 ㇏ 丶 ㇛

7——撇—— ノ ㇇ ㇋

6——勾—— 乙 乚 ㇆ ㇕ ㇇ ㇋ ㇉ 乙

由上述规则可见，下列独体字的编码为：

手：4079，笔画4，补0，起笔笔形代码7，收笔笔形代码9。

良：7088，笔画7，补0，起笔笔形代码8，收笔笔

形代码8。

二、将合体字的编码按照笔形的书写顺序“读”出组成合体字的各个部件（包括独体字、部首或笔形）的笔画数（即字码或代码，一般为0－9）的一组四位码作为该合体字的编码。

a、如果一个合体字的部件超过四部时，按汉字的结构取前四部笔画编码，例如：

喇：3242　　口－3，十－2，日－4，十－2

橙：4513　　木－4，𠂉－5，一－1，口－3

滚：3222　　氵－3，二－2，八－2，厶－2

b、如果一个合体字的部件刚好四部时：按汉字的结构顺序依次取四部笔画，构成汉字的四位编码，例如：

路：3433　　口－3，止－4，夂－3，口－3

统：3222　　纟－3，二－2，厶－2　儿－2

常：3233　　⺌－3，冖－2，口－3，巾－3

c、如果一个合体字的部件为三部时：按汉字的结构顺序依次取三部笔画，第四码取该字末部的收笔笔形代码，如：

程：5340　　禾－5，口－3，王－4，收笔“一”的笔形代码为0，

黄：4528　　艹－4，由－5，八－2，收笔“丶”的笔形代码为8，

d、如果一个合体字的部件为二部时：按汉字的结构顺

序依次取二部笔画，第三、四码取末部的起笔及收笔笔形代码，如：

取：6278　　耳－6，又－2，又的起笔笔形“7”代码为7，收笔笔形“㇏”代码为8，

三、词汇编码，在中文中，由单字可以组成成千上万条的词汇，且字和词是混合在一起使用的，为了体现中文汉字的这一特点，在以单字编码为基础上，结构型设立了词汇编码，使系统既能准确地处理尽可能多的单字，又能有效地处理常用词汇。为了相应地提高输入速度，所有词汇编码一律为等长四码，其码型与单字码完全相同。词汇编码按组成词汇的单字数，分为双字词、三字词、四字词和多字词四种情况，其输入方式如下：

a、双字词：取每个单字的前二码，组成四位码，如果词汇中的字为独体字，则按该独体字的笔画取两个相同的重码，如：

对待：2332　一定：1135　滋长：3344

b、三字词：取前两个字的第一码，最后一个字的前两码，（若最后一个字为独体字时，取相同码）组成四码，如：

新技术：2355　计算机：2642　微型机：3142

c、四字词：每个字各取其第一码，组成四码，如：

中外合资：4332　家用电器：3253　自动

控制：6232

d、多字词：取前四个单字的第一码，组成四码，如：

中国科学院：4353　　中华人民共和国：4225

本发明的结构型汉字编码输入方法是将汉字分析为独体字、部首和笔形三部分，由于只利用0～9十个数字键，因此不需要经过专门的指法训练，就能实现对汉字及其词汇进行直观快速的编码输入（即实现盲打），做到见字识码，会写就会输入，且简明易学，记忆量少，本编码的输入方法对简体字和繁体字均可适用。

采用本编码法出现重码字时，可以采用按其实用频度作分级处理的办法。即较常用的字首先显示在显示行的正常编辑位置上，若正好是用这个常用字，可继续输入下文，此字会自动显示在正常编辑位置上，而不影响输入速度。如果需要的是不常用的字，只需按与该字在提示行中的序号相同的数字键，所要的字即可显示在现行编辑位置上。重码时也可采用拼音字母进行选字，即当出现重码时，只需按与该字的第一个拼音字母相同的字母键，所要的字即可显示在现行编辑位置上，从而大大地缩短了重码的选择。

参见附图，这是作为一个实施例的操作键盘，采用结构型汉字编码输入法必须至少具有下述操作键并且按下述规程操作：

1、功能键（1），位于键盘的左边，用于决定操作时

编码的输入方法的键，因为一个文字处理系统可以同时采用几种汉字输入方法进行操作，所以必须通过功能键进行转换。

2、数字键（2），位于键盘的右边，用于输入汉字编码的键，键入一经开始显示器的正常编辑位置上立即显示与键入的编码相对应的汉字。

3、数字选字键（3），位于键盘的顶部，用于对提示行按频度顺序出现的字进行选择的键，由于在做了上述操作后，所要的汉字按频度顺序停留在提示行上，故需按与所要字在提示行中的序号相同的数字选字键。（包括0－9，10个数字键）。

或采用字母选字键（4），位于键盘的中央，用于对提示行中按拼音的第一个字母出现的字进行选择的键（包括26个字母键）。

4、空格键（5），位于键盘的下部，用于要求继续提示　　选字的键，若提示行所提示的五个或十个汉字中，没有所要求的汉字，这时须按下空格键，要求继续提示，然后用　　选字键输入。

5、清除键（6），键盘中的ESC键被定义为清除键，是用于作废其一次字码输入的键，即当操作者误打为另一组编码时，具有清除该编码的功能。

至于错误码或不存在字码的输入，系统会自动作废该次输入。

为了方便掌握汉字的部首及独体字，充分有效地使用本编码法，现列出《汉字部首及字例表》以及《独体字及其检码表》。

汉字部首及字例表

左部			右部					
笔画	形体	例字	笔画	形体	例字	笔画	形体	例字
二	亻	人伊华	二	刂(刂.刂)	刀别临坚归		廾	鹿
	讠	言试论		冂	兜假		冫	率兆
	阝	阡邦	三	勹	爪黎		廴	延
三	扌	手抚		彐	手归寻当		㠯	衰
	亻	攸修		㐅	场汤	五	𠂤	暇
	犭	犬独狠		丬	北		丗	塞
	饣	食饰馆		彐	舆	六	灬	米鬯卤
	忄	心忆惊	五	丰	拜			
	氵	水汝泛		东	练			
	丬	壮将					字头	
四	乍(乍)	爪窦印				笔画	形体	例字
	礻	示礼祖	中部			二	卜	卜卓卤
	手	拜看羊	笔画	形体	例字		冖	朱复
	卵	卵	二	丷(八)	八金办睿		勹	刀色急
五	𠂤	段	三	丰	用圣		冖	尔次
	衤	衣袂裙		彐	手秉		亠	文交
	钅	金针铁	四	丫	鬲岡		丷	八总卷
				夫	卷		工	夬侯

汉字部首及字例表

笔画	形体	例字	笔画	形体	例字	笔画	形体	例字
二	丶	甬		血	血 罩		氏	旅
	丶	径 经		灬	竹笈笋		灬	火杰盡黑
三	才	才存在		甲	卑		小	心 慕
	犬(犭)	友求	七	車	橐 囊		尸	声 眉
	小	小 肖		美	养幞		小	亦赤兼变
	尸	后 卮	字 底				夂	夜
	灬	学	笔画	形体	例子	五	水(水水)	水隶皋眾瘝
	丷	前喜善	二	川	齐弗		内	禺 禹
	丑	韋五		冫	冬於头	六	与	焉鳥馬
	犬	发	三	壬	虐		丞	眾
	戈	尧		牛(牛)	文韋 外 年		氽	聚
四	主	青素		从	匆	字 框		
	丗	廿 昔黄		小	小 絃	笔画	形体	例子
	尹	者考		氵	爪 鼠	二	冂	迥
	日	冒	四	辰	長辰展		几	几风凰
	爫	爪觅爱		歹	賓步剷		冂	用 甩
五	夫	春 秦		辶	止 走	三	冂	向 奥
	四	网罕曼		四	临 象	四	口	粤
六	西	西要覃		日	齊			
	且	具 真		衣	衣表衰			

独体字及其检码表

一 画

一1000 乙1066 乚1066 乙1066 乀1088

二 画

二2000 匚2000 七2006 亏2006 厂2007

ナ2007 厶2008 十2009 丁2009 凵2009

刂2009 刁2060 乜2066 刀2067 力2067

匕2067 卩2069 冂2069 九2076 了2076

儿2076 几2076 乃2076 勹2076 冖2076

厂2077 八2078 又2078 人2078 入2078

乂2078 廴2078 冫2080 巛2088 卜2098

三 画

三3000 土3000 工3000 与3000 士3000

彑3000 万3006 尢3006 乡3007 才3007

屮3007 大3008 下3008 寸3008 丈3008

幺3008 于3009 干3009 亍3009 廾3009

丌3009 艹3009 忄3009 刃3060 也3066

飞3068 丹3068 阝3068 子3070 孑3070

亾3070 亾3070 厶3070 乇3076 凡3076

乡3077 弓3077 凡3078 么3078 及3078
久3078 夕3078 又3078 丸3078 勺3078
孑3078 夂3078 夂3078 矢3078 个3079
千3079 川3079 彳3079 女3080 亡3080
门3086 宀3086 广3087 之3088 义3088
巛3088 丫3089 口3090 上3090 马3090
□3090 彐3090 已3096 弓3096 巳3096
巳3096 尸3097 小3098 山3099 巾3099
弋3008

四 画

王4000 巨4000 廿4000 卄4000 屮4000
无4006 丐4006 屯4006 丙4006 乏4006
先4006 牙4007 毋4007 不4008 木4008
太4008 犬4008 专4008 瓦4008 歹4008
尤4008 戈4008 乘4008 不4008 中4009
车4009 丰4009 井4009 韦4009 卅4009
引4009 币4009 母4009 历4009 月4070
鸟4070 丘4070 壬4070 毛4076 片4076
氏4076 冗4076 今4077 勿4077 肖4077
长4078 乏4078 爪4078 丹4078 天4078
夕4078 手4079 升4079 斤4079 牛4079
予4079 币4079 方4080 方4087 户4087

心4088 为4088 火4088 卞4088 日4090
止4090 丑4090 曰4090 市4090 见4096
巴4096 少4097 尹4097 水4098 书4098
尺4098 聿4099 弔4099

五 画

本5000 世5000 甘5000 玉5008 东5008
母5008 术5008 丙5008 未5008 末5008
疋5008 木5008 米5008 玉5008 戋5008
卌5009 卅5009 疋5068 白5070 生5070
乍5070 矛5077 鸟5078 禾5078 令5078
皮5078 乐5078 瓜5078 斥5078 氐5078
癶5078 乎5079 主5080 必5087 永5088
四5090 目5090 且5090 田5090 由5090
业5090 皿5090 冉5090 凹5090 凸5090
吕5090 冊5090 囙5090 电5096 民5096
冎5096 艮5098 水5098 甲5099 申5099
聿5099

六 画

西6000 耳6000 臣6000 亙6000 襾6000
兆6006 米6008 页6008 而6009 兩6009

自6070 血6070 舟6070 自6070 臼6070
6076 乒6077 舟6078 乓6078 舛 6078
州6089 曲6090 田6090 虎6090 鬼6097
虫6098 艮6098 隶6099

七 画

酉7000 臣7000 頭7008 雨7008 更7008
甫7008 豕7008 车7009 免7076 龟7076
身7077 多7077 采7078 長7078 鬼7078
言7080 良7088 粥7088 串7099

八 画

非8000 亞8000 事8006 雨8008 豕8008
兔8078 門8096

九 画

面9000 閒9079 為9088

十 画

倉1070 閂1096

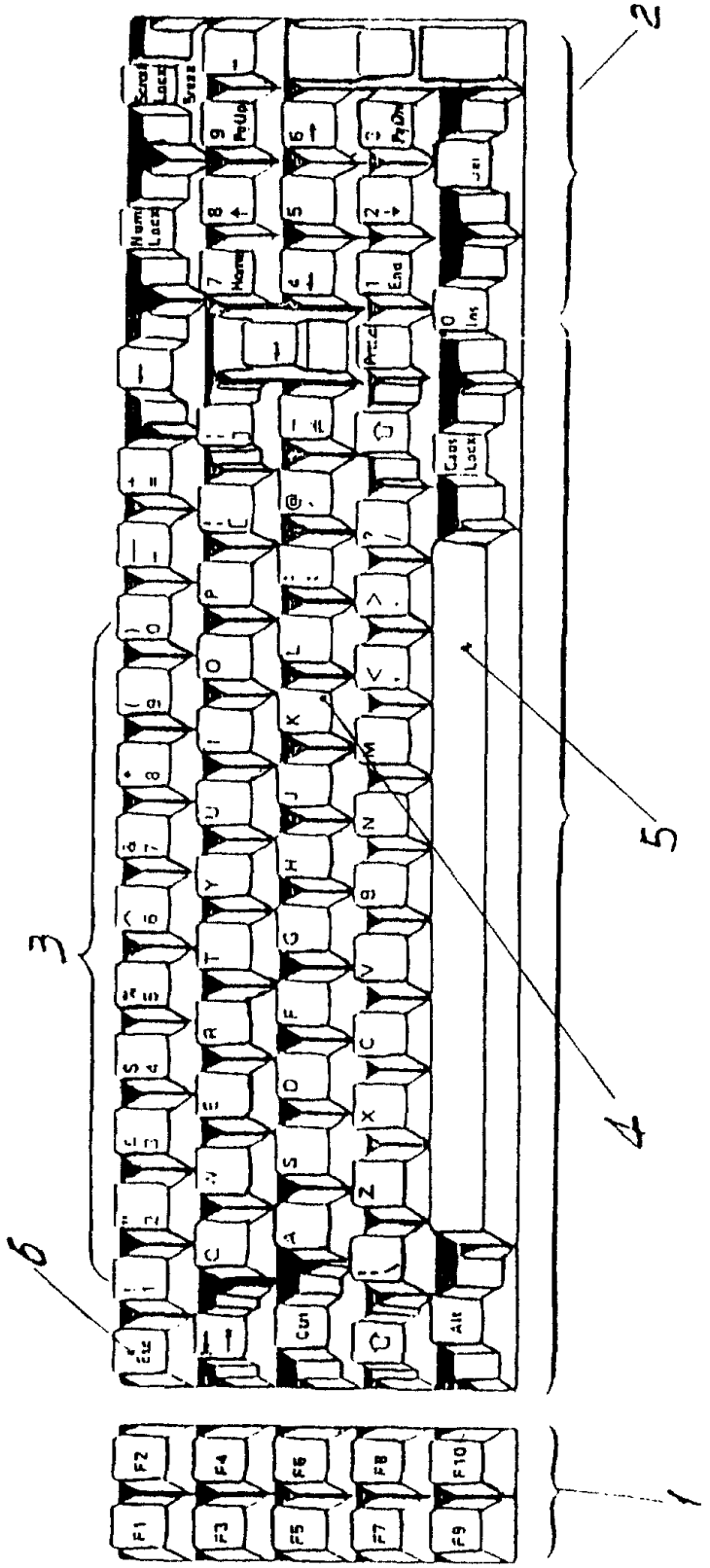

图 1