

汉字输入技术与应用研讨会论文集

纯数字编码的汉字输入法 将是今后键盘汉字输入法的主流 ——谈左右数码汉字电脑输入法及其键盘

郑岩松

【摘要】“易学”与“重码”是一对矛盾。人们是根据社会发展各个时期的不同要求，来决定如何解决这一矛盾。随着计算机、手机等设备的普及，输入法的“易学性”已是人们首要关心的问题；与此同时，由于手机、电话机发短信的业务已经展开，这些设备上的汉字输入的效率已越来越受到人们的关注；另外，人们在学会一种输入法后，对该输入法的部件或笔划在键盘上的分布已或多或少地形成了条件反射，再去学另一种的输入法，就会顾此失彼。所以，既能满足手机“上手即用”，又能满足计算机上输入汉字要每分钟20—80个汉字的纯数字编码的汉字输入法已成为社会的迫切需求。笔者发明的“左右数码汉字电脑输入法”能较好地满足这一要求。

现在社会已进入了数字化时代，数字键盘应用范围十分广泛，如电话机、手机、PAD等。汉字输入的数字化，已成为通讯、网络及金融商业领域的迫切需求。随着计算机应用的普及，掌握一门汉字输入法，已成为人们的必需技能。因此，既能满足手机等设备上使用要简单易学的要求，又能同时满足计算机上要求输入速度较快的纯数字小键盘输入汉字，已成为社会的迫切需求：

一、手机、信息电话机发短信和上网业务已或将得到蓬勃发展，它们的汉字输入已不仅仅局限于输入几个姓名，汉字输入的效率（速度）越来越受到人们的重视，而目前这些设备上使用的输入法的输入效率，将无法满足今后业务发展的需要。

二、银行的集约化办公，需要占用面积小键盘，它不仅要输入数字，也要输入汉字。

三、计算机已深入千家万户，掌握其基础知识，包括掌握一门汉字输入法，已成为人们生活与工作的必需技能。对绝大多数人来说，在“易学性”与“重码率”这一矛盾无法完全彻底解决的情况下，简单易学是它们首要关心的问题，其次才是输入速度（重码率）问题。其理由是，即使一种输入法其重码率再低、码长再短，也只是为快速输入提供一种可能，要真正做到输入速度快，必须还要同时具备三点：一是经过比较严格的手法训练；二是手指要灵活；三是要经常打字，否则难以形成条件反射。这些条件，对大部分人来说，是难以同时具备的，而实际上，大部分人都不可能经常有比较多的汉字再打，要形成快速的条件反射是不可能的。因此，每分钟能输入20—80个汉字也就足够了。这种的速度要求，为发明一个简单易学的编码提供了可能。

另外，人们在学会一种“形码”或“音形码”的输入法后，对该输入法的部件或笔划在键盘上的分布已或多或少地形成了条件反射，再去学另一种的输入法，就会顾此失彼。

因此，计算机上使用的与手机、电话机等设备上使用的应是一种的汉字输入法。也就是说，既能满足只含有小键盘设备，如手机、电话机，极其简单易（自学三分钟左右即可学会），又能满足计算机上输入汉字要每分钟20—80个汉字的纯数字编码的汉字输入法已成为中文信息处理领域极待突破的重大课题，它必将成为今后键盘汉字输入法的主流。

目前，数字编码方案有不少，但离能同时满足计算机、手机、电话机等设备同时使用的方案还有一定的距离。其中有一些方案把汉字中的“上下、上中下、左右、左中右、包围、半包围”结构的合体字分为两部分，然后对这两部分分别取码。采用这种方法，目的就是降低重码率。但是，在对它们进行分部时，产生了严重的“二义性”，致使用户无从下手，增加了学习难度。

笔者经过几年的努力，发明了“左右数码汉字电脑输入法”的汉字输入法，试图成为计算机、手机等设备都能使用的输入法。其技术方案如下：

左右数码输入法，利用数字来表示汉字的笔划、拼音和部件，其特征是将左右（含左中右）结构的汉字定义为“左右字”，再按书写顺序把“左右字”分为“左部”与“右部”；把“左右字”以外的汉字都定义为“整体字”；然后根据编码规则和数字与笔划、拼音及部件的对应关系，对“左右字”和“整体字”进行数字编码。其具体步骤如下：

一、首先将汉字分为“左右字”与“整体字”，并按书顺序，把“左右字”分成“左部”与“右部”。

(1)、“左右字”的定义：汉字中左右（含左中右）结构的合体字叫做“左右字”。按书写顺序，把“左右字”中先写到的部分叫做“左部”，余下来的叫做“右部”。如：

左右字	左部	右部
湘	彳	相
明	日	月

(2)、“整体字”的定义：“左右字”以外的汉字都叫做“整体字”。

二、左右数码键位示意图

7 卅 木 p q r s	8 t u v	9 wx y z
4 ˘ g h i	5 乙 j k l	6 口 日 m n o
1 一	2 a b c	3 丿 d e f

三、取码规则：

(1)、对于“左右字”，首先取该字拼音的第一个字母，其次取该字“左部”的第一、二个笔划或部件，再取该字“右部”的第一、二个笔划和最后一个笔划或部件，共六码，不足六码的，有多少取多少。

(2)、对于“整体字”，首先取该字拼音的第一个字母，其次取该字的前四个笔划或部件和最后一个笔划或部件，共六码，不足六码的，有多少取多少。

左右数码汉字输入法的特点分析：

一、把左右结构（含左中右）的汉字定义为“左右字”，再按书写顺序把“左右字”分为“左部”与“右部”；把其余的汉字都定义为“整体字”。这样的汉字分类法，具有以下优点：

1、在对“左右字”的分部时，不会产生“二义性”，为左右数码输入法的简单易学打下坚实基础。

笔者曾参与过原名“阴阳码”（现改名为“二笔码”）的发明工作，对“阴阳码”把汉字的合体字，主要是把上下、上中下、包围、半包围结构的汉字分为“阴半”与“阳半”的优缺点有着切身的体会：

(1)、在二十世纪九十年代初，由于计算机还很不普及，人们比较注重编码的重码率，因此，对合体字这样的分部，并在“跳跃技术”的帮助下，对降低重码率起到了重要作用。这与时代的要求是合拍的。

(2)、由于对上下、上中下、包围、半包围的汉字进行了分部，有相当一部分的汉字一时难以分清“阴部”与“阳部”，存在着不少的“二义性”，如：“爱、爵、磨、看、掌”等，人们要经过一段时间的练习后，才能掌握其正确的分部法。这就大大增加了学习难度。

(3)、有些独体字与上下结构的合体字难以分清。如：“产，表”等。

如今，计算机、手机等设备已进入千家万户，汉字编码首要解决应该是“易学性”的问题。要做到易学易用，就必须对汉字的定义和分类清楚明了，在实际使用过程中不能有二义性，规则或约定要少，尽可能减少记忆量等。从这一思路出发，笔者经过长期的研究发现，“左右”（含左中右）结构的汉字，其“形状”最容易被人们辨认，在划分“左部”和“右部”时不产生“二义性”，并且具有较好的离散能力。笔者把这一认识用于编码的实践，比较好地解决了现行一些技术方案中对汉字划分的“二义性”严重、规则或约定多的问题。

2、有较好的编码离散能力。

把汉字分为“左右字”与“整体字”，不仅有简单易学，而且也起到了离散重码的作用。以下以国标字集中“木”部汉字为例，说明左右数码离散重码，提高编码唯一性和输入效率的特点：

在限定每个汉字最大码长为五码且都用五种笔划作为码元的情况下，采用现有技术的前4末1五笔画数字编码时：

构：12344 树：12344 枚：12344 梗：12344 枢：12344

桧：12344 校：12344 根：12344 板：12344 杖：12344

其数字编码完全相同，这些字都会同一个编码。

按照左右数码取码规则，即按书写顺序，取“左部”的前二笔划，“右部”取“首次末”三个笔划，上述汉字的数字编码是：

构：12354 树：12544 枚：12314 梗：12124 枢：12135

桧：12344 校：12414 根：12514 板：12334 杖：12134

由此可见，用左右数码时，其数字编码都不相同，原来重码的以上各字都不重码了。所付出的代价仅仅是一条“左右字”的编码规则。

二、由于把上下、上中下、包围、半包围等结构的汉字与独体字归为一类，并采用统一的取码规则（前四加末一）编码，会产生较高的“重码率”，这至少对计算机用户来说，输入效率太低（速度慢），因此，用拼音的第一个字母作为码元，放在第一码的位置上。它不仅大大降低了重码率，也为那些对“ch, c”、“sh, s”、“zh, z”无法正确发音的人解除了烦恼。

三、采用两组使用频度高的组字部件“口、日”和“艹、木”，作为码元，放在键位6、7的位置上，目的是为了降低重码率。

四、尽管只取拼音的第一个字母，但对有些人说，对一些字的声母，如：“h、f”、“l、r”、“s、z”，也会混淆，因而导致无法正确输入。解决这一问题的办法是，用某个字符，如“？”或“1”作为“万能键”，来替代第一码的拼音。

以国标一级字库3755个字为例，去掉拼音后，其“键选率”也只有9.43%，也能满足手机、电话机上汉字输入效率的要求。

总之，左右数码汉字输入法较好地做到了“规范、易学、高效”，它不仅可当作计算机的“普及码”，也可作为上手即用的手机、电话机的输入法。理由是：

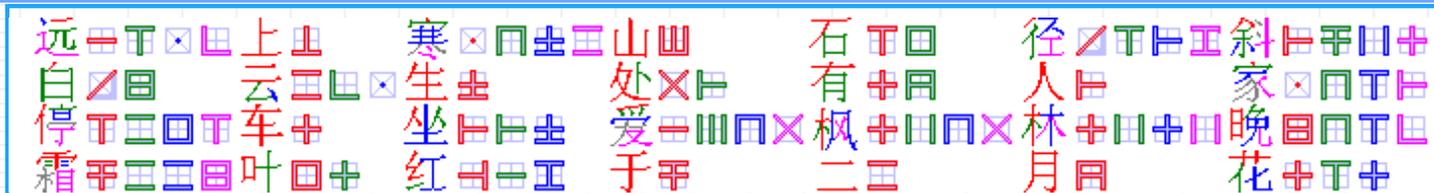
一、简单易学。只要记住“左右字”、“整体字”的两条取码规则、两组部件与6、7的对应关系即可。至于“横、竖、撇、捺、折”与“1、2、3、4、5”的对应和汉字拼音的第一个字母，人们已经习惯

了，几乎不要去刻意记忆。

二、“键选率”大大低于国家8%的标准。以国标一级字库3755个字为例，

简码后，键选率为0.18%。（单字的使用频度取自李公宜、刘如水主编的《汉字信息字典》）

三、可利用软件的功能来提高输入速度。由于“0”、“8”、“9”等数字键不可能出现在编成的等2—6位上，因此，可用作单字的断码键。“1”也不可能出在编码的第一位上，所以，“1”作为词组输入的引导，使词组与单字不重码；另外，“0”、“00—09”、“88”、“89”、“98”、“99”也不可能出现在单字编码的末位或首位上，所以都可作为功能键。



井田汉字，独一无二的汉字结体构形理论，能够科学地解决数码时代汉字所面临的问题！

推荐：[井田汉字](#)、[汉字书同文研究](#)、[中文虚拟学校](#)、[WPL 语言文学网](#)、[汉字编码设计学](#)、[《现代语文》](#)、[《中文》](#)、[百度](#)、[谷歌](#)

湘ICP备05008125号 [语言文字网](#) YYZW.COM©版权所有