



[12] 发明专利申请公开说明书

[21]申请号 94115205.7

[51]Int.Cl⁶

G06F 3/00

[43]公开日 1996年3月27日

[22]申请日 94.9.21

[71]申请人 黄金富

地址 100101北京市安定门外安立路8号汇园
公寓D座1108室

[72]发明人 黄金富

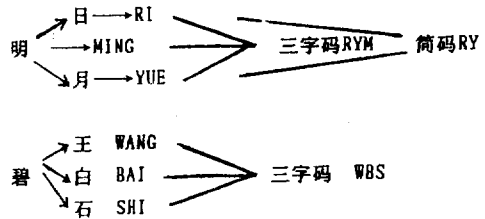
G06F 3/023

权利要求书 2 页 说明书 12 页 附图页数 2 页

[54]发明名称 汉字唯物拼音编码法及其中文电脑

[57]摘要

一唯物拼音中文电脑及其相应的唯拼赋码和输入方法，其特点是，将汉字按其结构自然地分解为两至三部分，分解为两部分时，取两部分的汉字的拼音的第一个字母即构成本发明唯拼简码，再取原汉字对应的汉语拼音的第一个字母排在简码后，就构成了唯拼三字码。当汉字被直接分解为三部分时，取每一部分的汉语拼音的第一个字母排列组合，即构成唯拼三字码，唯拼三字码由三个字母构成。本法由于把复杂的汉字分解成了简单的汉字或汉字及部首笔划，便利了输入和提高了输入速度。



权 利 要 求 书

1、一种中文电脑，包括键盘，CPU，中文字库，寄存器，显示器和打印器，其特征是，其中文字库所存贮的中文汉字是按唯物拼音赋码法所赋码的汉字和唯物中文码汉字，汉字输入方法是唯物拼音赋码的输入拼音字母和唯物中文码的输入方法，其汉字的唯物拼音赋码方法主要包括如下步骤：

1) 将汉字按其构成自然地分解成两或三部分，本方法限定分解不超过三部分；

2) 被分解的每一部分汉字或部首(包括笔划)赋予相应的汉语拼音或相应的指定汉字的汉语拼音；

3) 将被分解成两部分的汉字(包括部首)的相应汉语拼音的第一个字母取出按分解的先后顺序加以排列组合，即构成本发明下的汉字的简码，简码由两个字母组成；

4) 将上述汉字的简码在其尾部再加上原汉字的汉语拼音的第一个字母，即构成汉字的唯拼三字码，唯拼三字码由三个字母组成；

5) 将汉字按其构成自然分解为三部分时，按分解的左上、右上、左下、右下进行分解，将分解成的三部分汉字(包括部首)的汉语拼音的第一个字母取出，按分解时的顺序排列组合，即构成本发明的该汉字的唯拼三字码。

2、如权利要求1所述，汉字分解时，能分解成简单汉字时，必须分解为简单汉字。

3、如权利要求1所述，汉字分解时可分解成规定的部首和规定的笔划，其部首和笔划须赋予相应规定汉字的汉语拼音。

4、如权利要求1或3所述，规定的部首和笔划所相应的汉字和拼音可按表二、表三、表四的规定。

5、如权利要求1所述，对不能或不易分解的汉字，采用中文唯物数字码方法（本人发明并分布出版的唯物中文编码法——国际书号ISBN—111—01630—0/H·36）。

汉字唯物拼音编码法及其中文电脑

发明的技术领域：本发明涉及汉字的编码方法及其电脑输入方法

发明的技术背景：汉字是中国人民自远古以来创造的文字，汉字来源劳动和生活，它具有形象和函义相结合的特色。汉字是中华民族创造的最灿烂的文明之一，也是中华民族对世界所作出的最重要的贡献之一。

由于汉字源于生活，它是由象形字逐步演化而成的。因此，汉字的结构复杂，在现在电脑大量被使用之时，如何将汉字在电脑上输入，用什么样的编码法，正被许多人研究。

本人在1988年发明了汉字的唯物编码法及其中文键盘，并将其编印成字典公开出版，使世人使用，以促进中华文化的发展，（1988年，中国机械工业出版社，国际书号ISBN—111—01630—0/H·36）。

由于汉字的复杂性，又有许多新的中文输入和赋码方法被发明出来。本人利用汉字的可拆性及结合中国政府规定的汉字拼音方法，发明了汉字唯物拼音编码法及其相应的在电脑上的输入方法。

发明目的：发明一种汉字的新的编码法，以更方便和更快捷地在中文电脑中被使用，尤其是，利用汉字的可拆性，大量的汉字可以被方便地拆成两三部分，并结合本人以前发明的唯物中文编码法和中国的汉语拼音方法，使汉字的赋码和电脑输入能减至在三码或三码以内。

发明的详细说明：参见图1，图1是本发明的唯物拼音中文电脑方框图。此部中文电脑主要由键盘，CPU，中文字库，显示屏，寄存器和打印器构成。中文字库要和键盘的输入方法即汉字的赋码方

法相对应地存贮着汉字。键盘输入汉字的码后，经CPU将汉字从字库中调出并显示在显示器上。打好的文章由CPU输入到寄存器中寄存，或由打印器打印出来。这是此部中文电脑的基本装置。除中文字库外，其余已都是现有技术了。但是，汉字的输入方法虽然很多，仍各有长处短处，仍有改进的余地。本发明的方法，有许多明显的不同和优点。

汉语拼音方法在中国已被广泛的采用。在中国与英美等外国交往时，汉字也是被利用汉语拼音拼写，中国的人名，地名等名称，都是采用中国规定的汉语拼音方法拼写的。它可以用英文二十六个字母的键盘直接输入，这是该方法重要优点之一。所以，在中文电脑中，它是一种重要的输入方法和汉字赋码方法。由于汉语拼音越来越普及和被人所熟悉，本发明亦采用了拼音赋码的方法。

但是，汉语拼音的赋码方法也有许多不足之处，例如①重码多。例如“西”的拼音是xi，吸也是xi，细也是xi，xi下至少有同意字67个之多。

再例如“指”的拼音是zhi，同音字有只、止、值、置等至少有79个同音字之多。

重码多，电脑输入时，挑选费时，影响汉字输入电脑的速度。

其缺点之②有些字的码很长。

例如：“双”字是shuang 有6个字母之多

“生”字是sheng 有5个字母之多

“成”字是cheng 有5个字母之多

“床”字是chuang 有6个字母之多

字母太多，也影响电脑输入的速度。

缺点之③不好拼。

有些字难拼，例如崔字，应当是cui。但根据韵母表，往往会拼成cui，使显示屏上找不到这个字。复合韵母的情况比较复杂。一个字内有两个韵母，甚至三个韵母的情况都有，不易拼对。要查字典。寻找如何拼得正确。影响了向电脑中输入的速度。

但是，汉字的拼音赋码方法中，声母发音简单而容易拼写，再加上不用特殊键盘，采用普通英文键盘就可使用拼音方法。所以，本发明吸收了以上优点。

本发明考虑到汉字复杂，汉字是方块字，汉字又是写意文字，是由象形文字发展而来。象形字必然适合于采用简单图案文字拼成复杂图案文字。发展到现在的方块字，仍然保留了这个特色。例如“碧”字，可拆成“王”、“白”和“石”字。“明”可拆成“日”和“月”字。复杂的字拆成了简单的字。这时，再利用汉语拼音的赋码方法赋码和进行中文电脑输入，就会方便和快捷得多。

因此，本发明汉字的赋码方法是，先将汉字拆开，拆成两部分或三部分，这两三部分分别赋予其汉语拼音的第一个字母（绝大部分情况是声母），并加以组合，即构成本发明唯物拼音汉字的简码及三字码，本说明书中称为唯物拼音赋码法。由被分解的两个字的第一个字母构成的码称为简码，由被分解成的三个字的第一个字母构成的码叫三字码。另外，考虑到减少重码问题。一个汉字拆成两个字后，由该两个字的第一个拼音字母顺序组合而成的简码，可再加上原汉字汉语拼音的第一个字母，也构成三字码。这些都属于本发明的唯物拼音赋码法之内容。

由于汉语拼音有26个字母，如果只采用两字母组码，最多可以组成

$$26 \times 26 = 676$$

即676个字，这远远不够，因为常用的中文字有两、三千个。但如果是三位数，情况就有很大不同，按排列组合公式，26个字母构成三位数可构成

$$26 \times 26 \times 26 = 17576 \text{ 个}$$

但因I、U、V不能做第一个字母，所以应计算为

$$23 \times 23 \times 23 = 12167,$$

这也大大超过了常用汉字的数目，因此，可以比较有效地减少重码。

重码问题可按常规解决。在电脑输入时，重码字亦会在显示屏上显示，只要多按一个所选中的字的顺序号，该字就会被传送到显示屏中输入的文件里安排的位置上去。

现先举例说明本方法：参阅图2，图2举了两个字作例子。

“明”和“碧”字。汉字可以拆开，例如“明”字可以拆成“日”和“月”两个字，“日”字的汉语拼音第一字母是r，“月”的汉语拼音的第一个字母是y，“明”的汉语拼音第一个字母是M，因此，可以定义ry和rym为“明”字的中文唯物拼音码，ry是其简码，rym是其三字码，都是本发明的唯拼码。

下面再举例子：

					简码	三字码
zhong	zhong	xin				
忠	中 + 心		—————>		ZX	ZXZ
bi	wang	bai	shi			
碧	王 + 白 + 石		—————>		wb	wbs
kui	gui	dou				
魁	鬼 + 斗		—————>		gd	gdk
an	ge	an				
鞍	革 + 安		—————>		ga	gaa

zhang 章	li 立	+	zao 早	—————>	le	lee	
zhang 张	gong 弓	+	chang 长	—————>	gc	gcz	
ling 蛉	chong 虫	+	ling 令	—————>	cl	cll	
zao 皂	bai 白	+	qi 七	—————>	bq	bqz	
li 粒	mi 米	+	li 立	—————>	ml	mll	
gan 棍	mu 木	+	yue 日	+	bi 比	—————>	myb
xuan 轩	che 车	+	gan 干	—————>	cg	cgx	
lei 雷	yu 雨	+	tian 田	—————>	yt	yttl	

参阅表一，表一给出了更多的例子，全面说明了本发明的赋码方法。以上的例子指出，在采用本发明赋码方法时，先将一个汉字拆成两至三个汉字。如果拆成两个汉字，则采用这两个汉字的汉语拼音的第一个字母组合构成简码，再增加原汉字的第一拼音字母置于简码之后，构成三字码。如果汉字直接拆成三个汉字，则此三个汉字的第一个字母的组成构成为三字码。这里要强调的是，本赋码法中，最多三码，不须将汉字拆得更细，最多三笔就要输入。拆解汉字时，按照先左后右，行上后下，或左上右上左下右下顺序拆解，对应的字母也按此顺序构成简码和三字码。本说明书中大量例子已

给予说明。当用简码作电脑中文输入时，如已在显示屏上显示出所选汉字，则可即时打入文章中，而不用再输入第三码，节省了输入时间和提高了输入速度。

当用三字码输入时，如果字库中只有一个字与该码对应时，则该字自动跳入显示屏的文章中去，可减少击键的次数，便利输入。

以上赋码法解决了一大批汉字的赋码问题。

汉字太多，古老的中国字典里甚至收集了几万个中文汉字，但其中大多已不使用，据调查，如果掌握了三千个中文字，就已囊括了日常书报杂志上百分之九十五以上的汉字。

本发明人参阅了一册市场上较有名的中型语言工具书，是北京外国语学院英语系编写，商务印书馆1985年于北京出版的“汉英字典”内收有汉字六千余个。为了将汉字分解而使用本发明的唯物拼音编码法，按照分析分解中文字的习惯，分析汉字的构成，则从汉字的偏旁部首着手。每一字典的偏旁部首可能都互有相同互有不同，用此书进行说明。

根据其部首(本文中部首和偏旁函义基本相同)目录，该字典共采用部首226个。经分析得出，这226个部首的一大部分是单独的汉字，例如“寸”，“女”，“中”，“父”，“田”，“金”等。这些单独成字的部首所构成的汉字可直接使用上述赋码方法。

有如下几个部首，是简体之故，有繁体字对应，也可直接使用本发明的赋码法。

“讠”对应“言”字，涉及达137个字之多

“斗”对应“斗”字，(音pan，同音字“盘”)只涉及5个字，

“乞”对应“食”字，涉及达38个字，

有如下部首，可归为相应类似的汉字，而使用本发明的赋码法：

勹	→	爪
乙, 乚	→	乙
竹	→	竹
艮	→	艮
卜	→	卜
几	→	几
小	→	小
八	→	八
艹	→	艹
牛	→	牛
刀	→	刀

除了上述的以外，只有部首57个，不能直接作为单独汉字（无独立函义），和总数226个部首相比，只占25.2%，即只占约四分之一。这五十七个部首，又可分为三种情况，

第一种是简单笔划构成的部首，有五个。

第二种是有函义的部首，有二十余个，因为常用，所以人们习惯上给这些部首起了名字，例如“宀”叫做宝盖，“冫”叫做三点水，“亻”叫做单立人，等等，这二十余个部首对应着一千八百多个汉字，是最常用字中的重要部分，例如“亻”对应193个字，“冫”对应278个字，“艹”（草字头）对应286个字“扌”（提手）对应264个字，等等。

第三种是两划以上无常用函义的部首，例如“十”，“尢”等，计约三十余个，对应着约只有两百多字。

对这三种情况的汉字，本发明采取的方法是：

对第一种简单笔划构成的部首，取其对应的笔划的汉字，它们

如表二所示。笔划与指定汉字对应。这样笔划就有了拼音。

表二、简单笔划赋码表

笔划	名称	汉语拼音
丶	点	DIAN
丨	竖	SHU
丿	撇	PIE
フ	(横)勾	GOU
丁	丁(勾)	DING

这样，像“兵”字可马上分解为“丘”和“丿”简码为QP，三字码为QPP，像兵字可分解为“丘”和“丶”简码为QD，三字码为QDP，“凡”字可分解为“几”和“丶”简码为JD，三字码为JDF，等等。

对上述第二种非单字而有函义的部首，取其函义中的一个主要字的汉语拼音的第一个字母，作为该部首的组码字母，例如“亻”叫做单立人，形状像人字，取其对应者为人字，(REN)，则“亻”与“R”对应，例如“亿”字即可分解为“亻”和“乙”，简码为RY，三字码为RYY。再例如“借”字可即分解为“亻”和“昔”，简码为DX，三字码为DXJ，再举一例，“氵”通常被称为三点水，常被画为三个水滴，它可与水字对应，汉语拼音是SHUI，例如“海”字，即可分解为“氵”和“每”，简码为SM，三字码为SMH。再例如“江”字，可分解为“氵”和“工”字，简码为SG，三字码为SGJ。第二种有函义的部首意义重大。二十多个此种部首及赋名及分解字例见附表三。

第三种两划以上无常用函义的部首，本发明采用直接赋予一个对应字的方式，赋给其中的大部，其余的或归入表三，或采用本人发明的已用字典方式出版的唯物中文编码法编数字码。对于这第三

种两划以上的无常用函义的部首及分解其所对应的汉字举几例如下:

部首“宀”用零字,因为“宀”在本人唯物中文编码法中用0代表,“宀”涉及34个字之多。

例如:“宀”字分解为“宀”+“八”,简码LB,三字码LBL

“享”字分解为“宀”+“口”+“子”,三字码为LKZ

部首“隹”字用“佳”(Jia)字,因字形相近,“隹”涉及约十个字

例如:“隹”字分解为“隹”和“十”,简码JS,三字码JSS

“集”字分解为“隹”和“木”,简码JM,三字码JMJ

其余赋码情况,详见表三。

本赋码方法,即唯物拼音赋码方法中,能够采用上述自然分解的,都采用上述自然分解而予以赋码,凡用上述方法无法适当自然分解的少量汉字,仍采用本人以前发明的唯物中文赋码法,所以本发明称为唯物拼音赋码法。

采用本发明赋码方法的好处是:

- 1、只使用数量不多的简单汉字即可进行赋码和中文电脑输入;
- 2、由于分解了的字都是简单的字,其拼音便于记忆;
- 3、此赋码法大量使用声母,易于拼写;
- 4、码简单;
- 5、易于电脑输入;
- 6、输入速度可以提高,而借码和重码率低;
- 7、不识得的字,分解后即变得认识的而可输入(取→马+又)。

本发明在于又提供了一条汉字电脑输入和汉字赋码的新途径和方法,能使更多的人更容易更方便的掌握中文赋码和输入,使中华文化在更大的范围内更加发扬光大。

例字	中文唯物拼音赋码法及实例(码)						
	分解		简码	分解			三字码
明	日 RI	月 YUE	RY	日 RI	月 YUE	明 MING	RYM
碧				王 WANG	白 BAI	石 SHI	WBS
茄	艹(草) CAO	加 JIA	CJ	① 艹 CAO	加 JIA	茄 QIE	① CJQ
				② 艹 CAO	力 LI	口 KOU	② CLK
丘	丘 QIU	丿(撇) PIE	QP	丘 QIU	丿 PIE	丘 PING	QPP
楸	木 MU	秋 QIU	MQ	① 木 MU	秋 QIU	楸 QIU	MQQ
				② 木 MU	禾 HE	火 HUO	MHH
欣	斤 JIN	欠 QIAN	JQ	斤 JIN	欠 QIAN	欣 XIN	JQX
偕	亻 REN	皆 JIE	RJ	① 亻 REN	皆 JIE	偕 XIE	RJX
				② 亻 REN	比 BI	白 BAI	RBB
刃	刀 BAO	丶(点) DIAN	DD	刀 BAO	丶 DIAN	刃 REN	DDR

表一，本发明中文唯物拼音编码法及实例

偏旁部首	习惯的中文称谓	本发定义字	定义字的汉语拼音	例字	唯物拼音中文赋码法实例	
					简 码	三 字 码
冫	两点水	二	ER	冰	冫+水, ES	冫+水+冰, ESB
刂	立刀	立	LI	剖		立+口+刂, LKL
冫	秃宝盖	宝	BAO	冠		冫+元+寸, BYC
亻	单立人	人	REN	仔	亻+子, RZ	亻+子+仔, RZZ
卩	单耳刀	刀	DAO	卫	卩+一, DY	卩+一+卫, DYW
阝	双耳刀	耳	ER	邓	又+阝, YE	又+阝+邓, YED
氵	三点水	水	SHUI	浒	氵+许, SX	氵+言+午, SMW
忄	竖心	心	XIN	恰	忄+合, XH	忄+合+恰, XHQ
宀	宝盖	盖	GAI	案		宀+女+木, GNM
辶	走之	之	ZHI	送	关+辶, GZ	关+辶+送, GZS
艹	草字头	草	CAO	葫	艹+胡, CH	艹+古+月, CGY
扌	提手	手	SHOU	挥	扌+军, SJ	扌+军+挥, SJH
彡	须字旁	须	XU	形	开+彡, KX	开+彡+形, KXX
攴, 攴	反文	文	WEN	赦	赤+攴, CW	赤+攴+赦, CWS
犭	犬字旁	犬	QUAN	猜	犭+言, QY	犭+言+猜, QYY
灬	下四点	点	DIAN	照	昭+灬, ZD	日+召+灬, RZD
衤	单衣补	补	BU	祛	衤+去, BQ	衤+去+祛, BQQ
衤	双衣补	衣	YI	衬	衤+寸, YC	衤+寸+衬, YCC
疒	病字旁	病	BING	病	疒+丙, BB	疒+丙+病, BBB
纟	乱绞丝	丝	SI	缙		纟+处+口, SCK

表三，有习惯中文称谓的偏旁部首的赋码法及实例

偏旁部首	指定字	指定字之拼音	例字	唯物拼音中文赋码法实例	
				简码	三字码
亠	零	LING	享		亠+口+子, LKZ
ナ, 又	麦	MAI	友	ナ+又, MY	ナ+又+友, MYY
匚	考	KAO	𠄎	匚+也, KY	匚+也+𠄎, KY Y
冂	肆	SI	周	冂+吉, SJ	冂+士+口, SSK
凵	黑	HEI	画		一+田+凵, YTH
乚, 丩	贰	ER	午	乚+十, ES	乚+十+午, ESW
勹	捌	BA	勺	勹+丶, BD	勹+丶+勺, BDS
マ	妈	MA	勇		マ+用+力, MYL
厶	木	MU	矣	厶+矢, MS	厶+矢+矣, MSY
尢	尤	YOU	无	一+尢, YY	一+尢+无, YYW
車	车	CHE	幹		車+人+斗, CRD
隹	佳	JIA	隼	隹+十, JS	隹+十+隼, JSS
ヨ, ㄩ, ㄩ	要	YAO	寻	ヨ+寸, YC	ヨ+寸+寻, YCX
彭	发	FA	髻	彭+吉, FJ	彭+士+口, FCK
夂	老	LAO	孝	夂+子, LZ	夂+子+孝, LZ X
虍	虎	HU	虔	虍+文, HW	虍+文+虔, HWQ
耒	来	LAI	耕	耒+井, LJ	耒+井+耕, LJG
夫	秦	QIN	春	夫+日, QR	夫+日+春, QRC
苜	蒙	MENG	莹	苜+玉, MY	苜+玉+莹, MYY
尚	尝	CHANG	赏	尚+页, CY	尚+口+贝, CKB
美	券	QUAN	誉	美+言, QY	美+言+誉, QYT
戔	裁	CAI	戴		戔+田+共, CTG

表四，无习惯中文称谓的偏旁部首的赋码法及实例

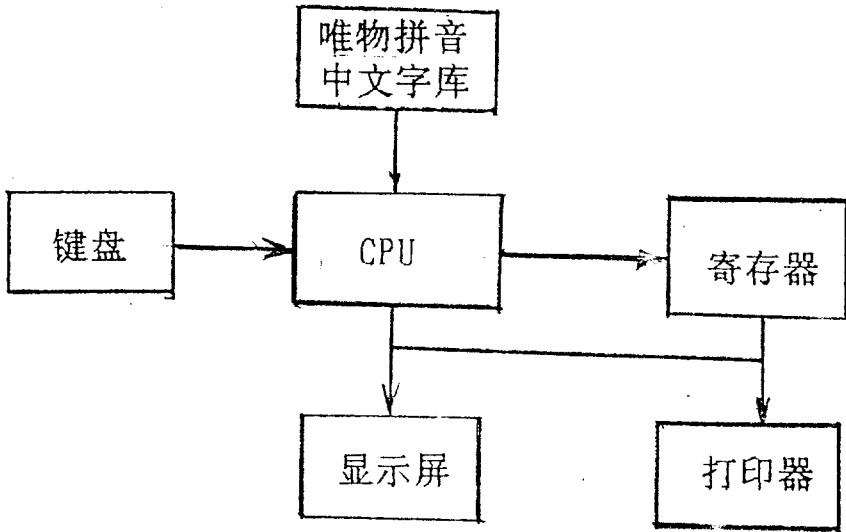


图1, 唯物拼音中文电脑方框图

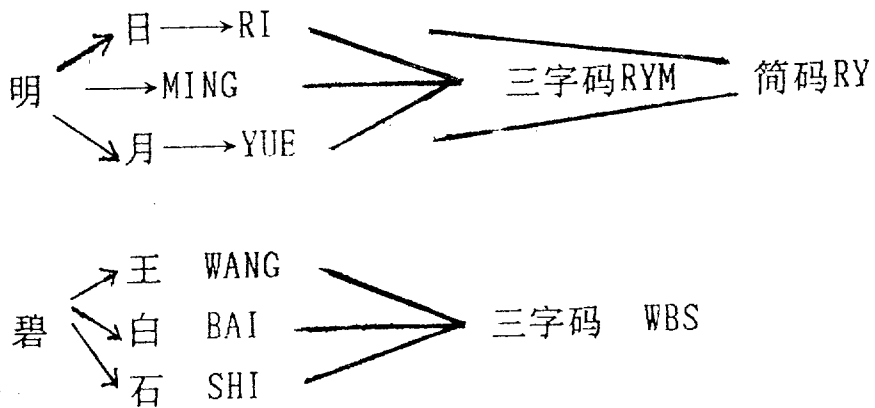


图2, 唯物拼音赋码法说明图