

## 键位相关速度当量的研究

陈一凡 张鹿

(北京信息工程学院)

周志农

(北京电子技术服务中心)

**【摘要】** 本文从二百多万个实验数据中统计分析出通用小键盘连续击键键位相关速度当量矩阵。键位相关速度当量表是优化汉字键盘输入键位设计的人机工程基础据,也为自动测定汉字键盘输入方法速度素质提供了科学依据。本文还介绍了采集上述数据的实验设计原理和数据处理所采用的方法。

(关键词: 键位相关速度当量、最小值归一、时间相对值。)

### 一、键位相关速度当量

在设计汉字键盘输入系统的键位分布,为了提高击键速度,目前通常采用的原则是:

1. 尽量做到左右手交替击键;
2. 食指、中指、无名指、小指的击键频度依次降低;
3. 中间键位安排使用频度高的码元;
4. 尽量减少同一手指的越排操作。

这些原则与人们的直观感觉:双手交替击键快于单手操作;食指动作敏捷、小指迟钝等相一致。然而上述原则并没有给出快慢量的差别,从而无法确定键位设计量化的合理性指标。

某些文献曾发表过各个手指击键速度的定量比较,如以右手食指的击键时间为1.0,则右手小指的击键时间为1.4,左手小指的击键时间为1.5。然而这些孤立的数据只能作为设计的参考而不能作为依据。

键位相关速度当量为连续击键过程中两键之间击键时间的相对值,无量纲。只有测定连续击键过程中键位相关速度当量才能为评测汉字输入系统键位分布合理性和进行键位分布设计提供更为有效的数据。

为了得到通用小键盘键位相关速度当量表,选用了30名指法标准训练有素的中、英文专业操作员,使用专门设计的输入文本和数据采集软件,从二百多万个数据中得到表1键位相关速度当量。表中的数据为连续击键的时间相对值,采用向最小值归一,即击键速度最快的键位搭配的击键相对时间为1.0。表的纵坐标表示先行键,横坐标表示后续键。

由表 1 我们应对上面提到的设计键位分布的原则作较大的修正,并给出了量的比较。

1. 左右手交替击键速度最快,速度当量多数为 1.0~1.2 之间;

2. 同指连续击键多数为 1.3;

3. 食指、中指、无名指、小指、姆指在左右手交替击键和同指连续击键时它们之间无明显差异;

4. 同手连续击键的击键速度较慢,速度当量在 1.5~2.5 之间,反映了各个手指在同手操作时的灵活性不同和键位位置对击键速度的影响;

5. 空格键尽管总是处于左右手交替操作的位置,但由于它的惯性大于其它键位和姆指击键时的接触位置等原因,无论先击还是后击,速度当量均为 1.5 左右。

表 1 所示的人—机工程学基本实验数据,为汉字键盘输入方法键位设计提供科学依据。

下面介绍键位相关速度当量数据采集的实验设计和数据处理方法。

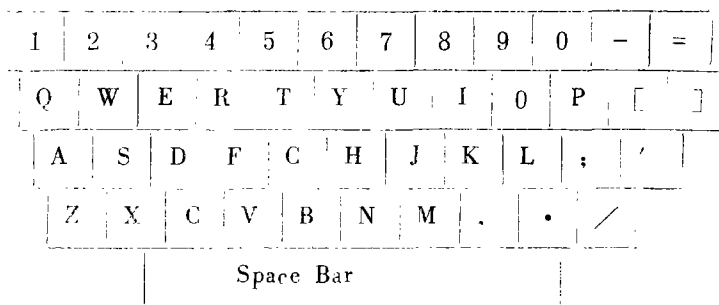
## 二、键位范围的确定

确定原则

1. 能复盖多数汉字键位输入系统;

2. 鉴于目前通用小键盘国际和国内均没有标准化,因此对某些位置不确定的键位(如  $\sim$ )不予采集。

根据以上原则,选定 26 个字母键、10 个数字键、9 个符号键、空格键共 46 个键位间的相关连击作为采集范围。除了功能选择键和控制键外,这 46 个键位的使用复盖了绝大多数汉字输入系统使用的键位。键位分布如图一。



图一 速度当量测定键位分布图

## 三、数据采集原理

在采集数据用的 PC 机系统中有一个由 8253 做计时器的时钟,基准时钟频率 1.19MHz,在正常情况下,通道 0 做计算机的系统时钟,每钟中断 18.2 次,为系统软件提供时钟计时。由于我们采集连续快速击键两键间的时间过程中需要中断频率更高的时钟计算,为此我们将时钟计算器初始化为每秒中断 400 次,这样键盘上的每个键位被按下时,时钟计数器单元记录下该键位被按下时间 t,采集软件同时记下被击键的位置,这样即可测出我们所要的测试

表1 键位相关

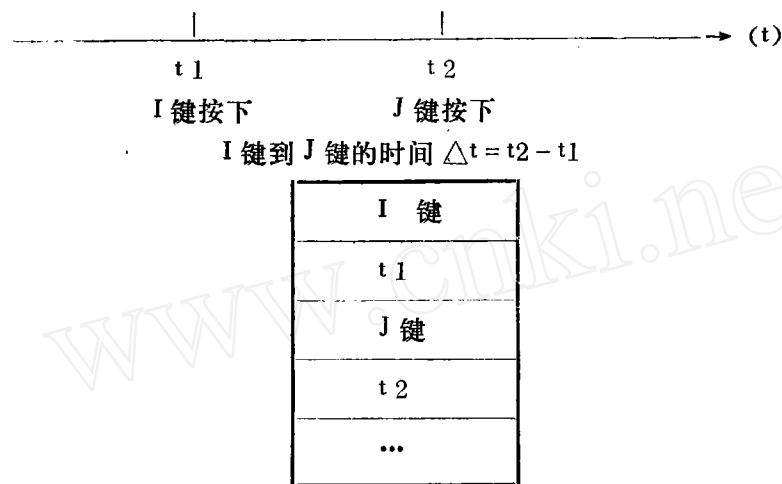
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
A,	1.3	1.8	1.7	1.5	1.5	1.5	1.5	1.1	1.1	1.1	1.0	1.0	1.1	1.1	1.1	1.2	1.9	1.7	1.6	1.5	1.2	1.5	1.8
B,	1.6	1.3	1.7	1.6	1.7	1.9	1.8	1.2	1.1	1.2	1.2	1.1	1.1	1.1	1.1	1.1	1.8	1.9	1.7	1.9	1.2	1.8	1.8
C,	1.6	1.6	1.3	1.7	1.8	1.7	1.7	1.1	1.1	1.1	1.0	1.0	1.1	1.0	1.1	1.1	1.8	1.9	1.6	1.9	1.1	1.5	1.8
D,	1.4	1.7	1.9	1.3	1.6	1.5	1.5	1.1	1.0	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.6	1.7	1.4	1.7	1.1	1.6	1.6
E,	1.5	1.8	1.9	1.8	1.3	1.5	1.6	1.1	1.1	1.0	1.0	1.1	1.1	1.1	1.0	1.2	1.4	1.3	1.7	1.6	1.1	1.7	1.4
F,	1.3	1.8	1.8	1.5	1.4	1.3	1.8	1.1	1.2	1.1	1.1	1.1	1.2	1.1	1.1	1.1	1.5	1.7	1.5	1.7	1.2	1.8	1.4
G,	1.4	1.8	1.8	1.5	1.4	1.6	1.3	1.2	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.4	1.8	1.5	1.7	1.1	1.8	1.5
H,	1.2	1.2	1.2	1.1	1.1	1.2	1.5	1.3	1.3	1.6	1.5	1.4	1.8	1.6	1.4	1.4	1.2	1.1	1.1	1.1	1.5	1.2	1.1
I,	1.2	1.2	1.2	1.1	1.1	1.1	1.2	1.7	1.3	1.5	1.6	1.6	1.7	1.5	1.4	1.5	1.2	1.1	1.1	1.1	1.4	1.2	1.1
J,	1.1	1.2	1.1	1.1	1.1	1.2	1.2	1.6	1.5	1.3	1.3	1.4	1.8	1.6	1.5	1.4	1.3	1.1	1.2	1.1	1.6	1.2	1.2
K,	1.2	1.1	1.2	1.1	1.1	1.2	1.2	1.6	1.6	1.5	1.3	1.3	1.6	1.6	1.6	1.8	1.2	1.1	1.1	1.1	1.6	1.2	1.1
L,	1.2	1.2	1.2	1.1	1.1	1.1	1.3	1.7	1.5	1.6	1.6	1.4	1.7	1.7	1.7	1.8	1.3	1.1	1.1	1.1	1.6	1.2	1.2
M,	1.2	1.1	1.2	1.1	1.1	1.2	1.1	1.8	1.6	1.6	1.6	1.5	1.3	1.6	1.7	1.7	1.3	1.1	1.2	1.2	1.8	1.2	1.2
N,	1.2	1.2	1.2	1.1	1.1	1.2	1.2	1.7	1.5	1.6	1.6	1.5	1.7	1.3	1.5	1.6	1.3	1.1	1.2	1.1	1.7	1.2	1.1
O,	1.1	1.1	1.2	1.1	1.1	1.1	1.1	1.4	1.5	1.5	1.8	1.7	1.7	1.5	1.4	1.5	1.3	1.2	1.1	1.1	1.3	1.3	1.1
P,	1.1	1.1	1.1	1.1	1.1	1.0	1.1	1.6	1.6	1.6	1.9	1.9	1.7	1.8	1.7	1.3	1.3	1.2	1.1	1.3	1.5	1.3	1.3
Q,	1.9	1.9	2.1	1.9	1.7	1.6	1.7	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.2	1.3	1.3	1.4	2.1	1.6	1.2	1.7	1.7
R,	1.5	1.9	2.0	1.7	1.4	1.7	1.7	1.1	1.1	1.1	1.2	1.1	1.2	1.1	1.2	1.2	1.5	1.3	1.7	1.7	1.1	1.9	1.5
S,	1.6	1.7	1.8	1.5	1.5	1.5	1.5	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.7	1.6	1.3	1.6	1.1	1.6	1.7
T,	1.5	1.9	2.0	1.7	1.4	1.7	1.7	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.4	1.7	1.5	1.3	1.2	2.0	1.5
U,	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.6	1.4	1.6	1.7	1.6	1.8	1.7	1.4	1.3	1.1	1.1	1.2	1.3	1.3	1.2	1.2
V,	1.5	1.8	1.6	1.6	1.6	1.7	1.8	1.1	1.1	1.1	1.1	1.2	1.1	1.2	1.2	1.3	1.7	1.9	1.5	1.9	1.4	1.3	1.5
W,	1.8	1.9	2.1	1.9	1.5	1.6	1.6	1.1	1.1	1.1	1.1	1.1	1.2	1.1	1.1	1.2	1.6	1.5	1.8	1.5	1.2	1.9	1.3
X,	1.7	1.5	1.5	1.5	1.8	1.7	1.5	1.1	1.1	1.0	1.1	1.0	1.1	1.1	1.1	1.1	2.0	1.7	1.7	1.8	1.1	1.6	1.9
Y,	1.2	1.1	1.1	1.1	1.1	1.1	1.1	1.6	1.4	1.8	1.7	1.7	1.9	1.7	1.5	1.4	1.2	1.2	1.1	1.3	1.5	1.3	1.3
Z,	1.8	1.5	1.6	1.7	1.8	1.7	1.7	1.0	1.1	1.0	1.1	1.1	1.0	1.1	1.2	1.1	2.1	1.8	1.9	1.7	1.2	1.5	2.1
0,	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.7	1.9	1.9	2.1	2.0	1.9	1.8	1.8	1.7	1.2	1.1	1.2	1.3	1.6	1.1	1.2
1,	2.2	2.2	2.3	2.2	2.1	2.0	1.9	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	2.1	1.8	2.4	1.6	1.3	2.2	2.1
2,	2.0	2.2	2.2	2.1	1.8	1.9	1.8	1.1	1.1	1.1	1.1	1.0	1.1	1.1	1.1	1.1	1.9	1.7	2.0	1.7	1.2	2.0	1.9
3,	1.7	2.0	2.1	1.9	1.8	1.9	1.8	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.2	1.6	1.7	1.8	1.6	1.2	1.9	1.7
4,	1.6	2.1	2.2	2.1	1.9	1.9	1.9	1.0	1.1	1.1	1.1	1.1	1.0	1.1	1.1	1.2	1.5	1.8	1.8	1.8	1.1	1.9	1.6
5,	1.7	2.0	2.1	1.9	1.8	1.9	1.8	1.1	1.0	1.1	1.1	1.0	1.1	1.1	1.1	1.1	1.5	1.8	1.7	1.7	1.2	2.1	1.6
6,	1.3	1.3	1.1	1.2	1.1	1.2	1.4	1.8	1.7	1.9	2.0	1.9	2.1	2.0	1.7	1.7	1.2	1.2	1.2	1.3	1.7	1.2	1.2
7,	1.2	1.2	1.1	1.1	1.1	1.1	1.2	1.8	1.7	1.8	1.9	1.9	2.0	1.9	1.7	1.7	1.1	1.1	1.2	1.1	1.6	1.2	1.2
8,	1.2	1.2	1.1	1.1	1.1	1.1	1.2	1.6	1.6	1.7	1.8	1.9	1.9	1.8	1.9	1.6	1.2	1.1	1.1	1.2	1.5	1.1	1.2
9,	1.2	1.1	1.1	1.1	1.1	1.2	1.2	1.8	1.8	1.7	2.0	1.9	2.0	1.9	1.8	1.7	1.2	1.1	1.2	1.2	1.6	1.1	1.2
,	1.2	1.2	1.2	1.1	1.2	1.1	1.2	1.9	1.9	1.8	1.6	1.7	1.6	1.7	1.9	2.0	1.2	1.1	1.2	1.2	1.9	1.2	1.2
.	1.1	1.1	1.2	1.1	1.1	1.1	1.1	2.0	1.9	1.8	1.7	1.8	1.6	1.8	2.0	1.9	1.2	1.2	1.1	1.2	1.9	1.1	1.2
/,	1.2	1.2	1.2	1.2	1.1	1.2	1.2	2.0	2.0	1.8	1.9	2.0	1.7	1.9	2.3	2.1	1.2	1.2	1.1	1.2	1.9	1.2	1.3
[,	1.2	1.2	1.2	1.1	1.1	1.1	1.2	1.8	1.9	1.8	2.2	2.2	2.0	1.9	2.1	1.8	1.2	1.2	1.2	1.2	1.6	1.2	1.3
],	1.2	1.2	1.2	1.3	1.2	1.3	1.2	2.1	2.1	2.1	2.3	2.3	2.0	2.1	2.2	2.0	1.2	1.2	1.1	1.2	1.9	1.2	1.2
{,	1.2	1.2	1.1	1.1	1.2	1.1	1.2	1.7	1.8	1.7	1.8	2.0	1.9	1.8	2.1	1.8	1.2	1.1	1.1	1.2	1.7	1.2	1.2
',	1.2	1.2	1.2	1.3	1.2	1.2	1.2	1.8	1.9	1.8	2.0	2.1	1.8	1.8	2.1	2.0	1.1	1.2	1.2	1.2	1.9	1.2	1.2
-,	1.2	1.2	1.3	1.2	1.2	1.2	1.2	1.9	2.1	1.9	2.3	2.4	2.1	2.0	2.0	1.8	1.2	1.2	1.2	1.2	1.7	1.2	1.3
=,	1.2	1.2	1.2	1.2	1.1	1.2	1.3	2.1	2.2	2.0	2.4	2.4	2.1	2.1	2.1	1.9	1.2	1.2	1.1	1.2	1.8	1.2	1.2
_,	1.4	1.4	1.5	1.4	1.4	1.4	1.3	1.3	1.4	1.4	1.3	1.3	1.4	1.3	1.4	1.4	1.5	1.4	1.5	1.5	1.5	1.5	1.5

速度当量表

X	Y	Z	0	1	2	3	4	5	6	7	8	9	,	.	/	[	]	,	'	-	=	□
1.9	1.2	1.9	1.1	2.2	2.2	2.0	1.7	1.6	1.3	1.1	1.1	1.1	1.2	1.0	1.0	1.2	1.1	1.3	1.2	1.2	1.2	1.3
1.7	1.3	1.6	1.2	2.1	2.2	2.0	2.2	2.2	1.4	1.2	1.2	1.1	1.3	1.1	1.1	1.2	1.1	1.2	1.1	1.2	1.4	1.2
1.6	1.2	1.4	1.2	2.1	2.2	2.0	2.1	2.1	1.4	1.1	1.2	1.1	1.2	1.1	1.1	1.2	1.1	1.2	1.1	1.1	1.1	1.2
1.7	1.2	1.5	1.2	1.8	1.9	1.9	1.9	2.0	1.2	1.2	1.1	1.2	1.2	1.0	1.1	1.2	1.1	1.2	1.2	1.2	1.1	1.3
2.0	1.2	1.8	1.2	1.7	1.8	1.8	1.7	1.7	1.3	1.2	1.1	1.2	1.2	1.1	1.1	1.2	1.1	1.2	1.1	1.2	1.2	1.5
1.7	1.2	1.6	1.2	1.7	1.7	1.8	1.9	1.9	1.3	1.2	1.2	1.1	1.2	1.1	1.2	1.2	1.1	1.2	1.0	1.2	1.1	1.3
1.7	1.1	1.6	1.2	1.6	1.8	1.8	2.0	1.9	1.3	1.2	1.2	1.2	1.2	1.1	1.1	1.2	1.1	1.2	1.1	1.2	1.2	1.4
1.2	1.5	1.2	1.5	1.3	1.1	1.2	1.2	1.2	1.7	1.7	1.4	1.5	1.8	1.8	1.9	1.6	1.8	1.6	1.7	1.7	1.8	1.3
1.2	1.6	1.3	1.7	1.4	1.1	1.1	1.2	1.1	1.9	1.8	1.7	1.6	1.9	1.9	1.9	1.7	1.9	1.8	1.8	1.8	2.0	1.3
1.2	1.8	1.3	1.7	1.3	1.2	1.1	1.2	1.2	1.9	1.8	1.6	1.6	1.9	1.7	1.7	1.6	1.8	1.6	1.6	1.8	1.8	1.2
1.3	1.7	1.3	1.8	1.3	1.2	1.1	1.2	1.2	2.0	1.9	1.9	1.9	1.8	1.6	1.7	1.8	1.9	1.7	1.7	2.0	2.1	1.2
1.2	1.8	1.3	1.9	1.3	1.2	1.2	1.2	1.2	2.1	2.0	1.9	1.9	1.9	1.8	1.8	1.8	2.0	1.7	1.8	2.1	2.1	1.2
1.3	2.0	1.3	1.9	1.4	1.3	1.2	1.2	1.2	2.2	2.1	1.8	1.8	1.7	1.6	1.6	1.9	1.9	1.8	1.7	2.0	2.0	1.3
1.3	1.8	1.3	1.8	1.2	1.2	1.2	1.2	1.2	2.0	1.9	1.7	1.6	1.8	1.6	1.7	1.8	1.9	1.7	1.7	1.9	1.9	1.2
1.2	1.5	1.3	1.6	1.3	1.2	1.2	1.3	1.3	1.8	1.8	1.7	1.7	2.0	2.1	2.0	1.8	2.0	1.9	1.9	1.9	1.9	1.3
1.2	1.6	1.2	1.9	1.3	1.2	1.2	1.2	1.2	1.9	1.9	2.0	2.0	2.4	2.3	2.2	1.7	1.8	1.9	1.9	1.8	1.9	1.3
2.1	1.2	2.0	1.1	1.8	1.9	1.8	1.6	1.6	1.3	1.1	1.0	1.1	1.1	1.0	1.1	1.2	1.0	1.2	1.1	1.2	1.1	1.1
2.0	1.2	2.0	1.2	1.6	1.7	1.6	1.8	1.7	1.2	1.1	1.1	1.1	1.2	1.1	1.1	1.2	1.1	1.1	1.1	1.3	1.2	1.4
1.8	1.1	1.8	1.2	2.1	2.0	1.9	1.8	1.9	1.2	1.2	1.1	1.1	1.2	1.1	1.1	1.2	1.1	1.2	1.1	1.3	1.2	1.4
2.1	1.2	1.9	1.2	1.6	1.7	1.7	1.8	1.8	1.5	1.2	1.2	1.1	1.3	1.1	1.1	1.3	1.1	1.2	1.1	1.3	1.2	1.3
1.2	1.5	1.3	1.4	1.3	1.0	1.2	1.2	1.2	1.8	1.5	1.5	1.4	2.0	1.8	1.7	1.4	1.7	1.6	1.7	1.5	1.6	1.3
1.5	1.3	1.4	1.2	1.9	1.9	1.9	2.1	2.0	1.2	1.1	1.2	1.1	1.2	1.1	1.1	1.1	1.1	1.1	1.1	1.2	1.1	1.2
2.0	1.2	2.0	1.4	1.8	1.8	1.7	1.7	1.8	1.3	1.2	1.1	1.1	1.2	1.1	1.1	1.2	1.1	1.2	1.1	1.3	1.2	1.4
1.3	1.2	1.7	1.1	2.3	2.2	2.1	2.1	2.2	1.3	1.2	1.1	1.1	1.2	1.1	1.1	1.2	1.2	1.2	1.1	1.2	1.2	1.3
1.5	1.3	1.3	1.5	1.3	1.2	1.1	1.2	1.2	1.6	1.5	1.4	1.3	1.9	1.9	1.8	1.6	1.9	1.7	1.7	1.6	1.8	1.3
1.5	1.4	1.3	1.2	2.2	2.4	2.1	2.0	2.0	1.3	1.2	1.1	1.0	1.2	1.0	1.1	1.2	1.0	1.2	1.1	1.2	1.1	1.3
1.3	1.7	1.4	1.4	1.4	1.4	1.2	1.2	1.2	1.6	1.6	1.6	1.7	2.4	2.2	2.2	1.8	2.0	2.0	2.1	1.7	1.7	1.4
2.3	1.3	2.4	1.2	1.3	1.5	1.7	1.5	1.5	1.2	1.1	1.1	1.1	1.2	1.0	1.1	1.2	1.1	1.2	1.1	1.2	1.1	1.2
2.2	1.3	2.3	1.2	1.6	1.4	1.6	1.6	1.5	1.2	1.1	1.0	1.1	1.3	1.1	1.0	1.2	1.1	1.2	1.1	1.3	1.1	1.3
2.1	1.2	2.1	1.2	1.5	1.6	1.3	1.4	1.5	1.3	1.1	1.1	1.1	1.2	1.0	1.1	1.2	1.1	1.2	1.1	1.3	1.1	1.3
2.1	1.1	2.1	1.1	1.4	1.4	1.5	1.3	1.8	1.2	1.2	1.1	1.1	1.2	1.1	1.1	1.2	1.1	1.2	1.1	1.3	1.1	1.4
2.0	1.2	2.0	1.2	1.3	1.4	1.4	1.7	1.3	1.3	1.2	1.1	1.1	1.3	1.2	1.1	1.2	1.1	1.2	1.1	1.2	1.1	1.4
1.2	1.6	1.4	1.4	1.2	1.2	1.2	1.3	1.5	1.3	1.4	1.4	1.4	2.2	2.1	2.2	1.8	2.1	2.0	1.9	1.5	1.8	1.4
1.2	1.6	1.2	1.4	1.2	1.2	1.1	1.1	1.3	1.7	1.3	1.3	1.4	2.1	2.2	2.1	1.9	1.9	2.0	2.0	1.6	1.8	1.3
1.2	1.5	1.2	1.5	1.1	1.1	1.1	1.2	1.2	1.6	1.5	1.3	1.4	2.1	2.0	2.1	1.8	2.1	2.0	2.0	1.7	2.0	1.3
1.2	1.7	1.3	1.4	1.2	1.1	1.2	1.3	1.3	1.7	1.5	1.5	1.3	2.2	2.2	2.1	2.0	2.1	2.2	2.2	1.8	2.0	1.4
1.3	2.1	1.3	2.0	1.3	1.2	1.2	1.2	1.3	2.3	2.2	2.1	2.2	1.3	1.6	1.8	2.1	2.1	2.0	2.0	2.3	2.2	1.3
1.2	2.0	1.3	2.1	1.3	1.1	1.2	1.2	1.3	2.3	2.2	2.2	2.1	1.6	1.4	1.7	2.0	2.1	1.9	1.9	2.2	2.2	1.3
1.2	2.0	1.3	2.3	1.2	1.2	1.2	1.1	1.4	2.3	2.3	2.2	2.4	1.9	1.9	1.4	2.0	2.0	1.9	1.8	2.3	2.2	1.3
1.2	1.8	1.3	2.0	1.2	1.2	1.2	1.2	1.3	2.1	2.0	2.1	2.2	2.5	2.2	2.0	1.4	1.6	1.9	1.8	1.8	1.7	1.3
1.2	2.1	1.3	2.3	1.3	1.2	1.2	1.2	1.4	2.4	2.2	2.3	2.5	2.5	2.4	2.1	1.9	1.4	2.0	1.9	2.0	1.8	1.4
1.2	1.9	1.3	2.0	1.2	1.2	1.1	1.1	1.4	2.1	2.1	2.1	2.3	2.0	2.0	1.8	1.8	2.0	1.4	1.7	2.0	2.0	1.4
1.2	2.0	1.4	2.2	1.2	1.2	1.2	1.2	1.3	2.3	2.1	2.2	2.4	2.1	2.0	1.8	1.8	1.7	1.9	1.4	2.0	1.9	1.3
1.2	1.8	1.2	1.8	1.2	1.2	1.2	1.1	1.3	1.9	1.8	2.1	2.2	2.5	2.5	2.2	1.8	2.0	2.0	2.2	1.4	1.6	1.3
1.2	1.9	1.3	1.9	1.3	1.2	1.1	1.4	2.1	2.0	2.2	2.2	2.6	2.5	2.3	1.7	1.8	2.0	2.0	1.8	1.3	1.5	1.5
1.4	1.6	1.5	1.4	1.7	1.6	1.4	1.4	1.6	1.6	1.6	1.5	1.5	1.4	1.5	1.5	1.5	1.4	1.5	1.5	1.6	1.6	1.3

24

数据。按照上面的思想，我们设计了数据采集软件。采集原理如图二革示。



图二 键位相关连击键序及时间采集原理

在数据采集过程中，当键盘中断和 DRAM 刷新或数据自动存盘 DMA 传送 发生 冲突时 将产生系统误差，这就要求尽量优化中断服务程序，因此采集程序用汇编语言编写。采集用

- 5

机器的石英时钟脉冲的稳定性为 10，足够满足采集精度要求。不同机器的时钟频率不一定完全一致，只要使同一操作员使用固定的机器，不同操作员的数据归一化后再合并，就不致因机器时钟频率不一致而带来误差。这样，系统误差将远小于本项目要求 5% 的偶然误差。

### 四、采集文本的设计

操作员在看稿键入过程中，其选择反应时间按海曼公式为：

$$T = a + bH(k) \tag{1}$$

式中  $H(k)$  为码元熵值， $a$  为操作员从看稿到击键的简单反应时， $b$  为与键位分布有关的常数。为尽量降低  $a$  值及缩小  $a$  值在不同操作员之间离散性，同时又要尽可能使采集过程手腕动作与输入汉字文本时的动作相接近，经过反复试验，采用 4 个码元为一组的字符串。英文录入操作员对诸如 of，ing 等高频使用搭配，汉字录入操作员对他们常用的码元组合 均具有条件反射式的反应，因此不能采用英文单词或中文单字（词）作为采集样本，而是采用由 46 个码元组成的离散文本。样本还要保证表 1 中每个数据的采样量的一致性和同一组键位组合前后键的多样性。

### 五、样本的大小确定

连续击键键位相关速度当量的母体为经过正规指法训练的通用键盘使用人员连续击键时两键间相隔时间的相对值。根据小子样抽样理论，我们选用的 30 名操作员作为人数样本，每个操作者在同一组相关键位提供 36 个数据为个人样本。这样，每个操作员提供的数据为

$46 \times 46 \times 36 = 76176$  个, 从而采集数据的总数为  $76176 \times 30 = 2285280$  个, 表 1 中每个数据的样本大小为  $36 \times 30 = 1080$  个。

## 六、采用数据的处理

为了使表 1 的数据能反映绝大多数经过正规指法训练的通用键盘使用人员的击键规律, 对悖于多数人规律的数据, 对采样过程中受干扰得出的数据, 误操作数据, 均作为非正常数据对待而采用肖维涅 (W. Chauvenet) 判据原则予以剔除。即将某一操作员对同一组速度当量的采集视作一系列  $N$  次等精度测量, 每一次测量的剩余误差的绝对值  $|\xi| = |X_i - \bar{X}|$  应小于  $K\sigma$ ,  $\sigma$  为  $N$  次测量值的方差,  $K$  为与测量次数有关的常数。当  $|X_i - \bar{X}| \geq K\sigma$  时,  $X_i$  被认为是非正常数据。

设在  $N$  次等精度数据采集,  $|X_i - \bar{X}| \geq K\sigma$  的次数为  $M$  次, 当  $N$  足够大时, 概率  $P\{| \xi | \geq K\sigma\} \approx \frac{M}{N}$ , 由于要求  $M$  为零, 因此

$$\frac{M}{N} \approx P\{| \xi | \geq K\sigma\} \rightarrow 0$$

或

$$M = NP\{| \xi | \geq K\sigma\} \rightarrow 0$$

式中  $a = K\sigma$ 。

设  $K$  值取定后最低限度应有

$$NP\{| \xi | \geq K\sigma\} \leq \frac{1}{2}$$

则

$$\begin{aligned} P\{| \xi | \geq a\} &= 2 \int_a^{\infty} P(\xi) d\xi \\ &= 1 - \operatorname{erf}(k) = \frac{1}{2N} \end{aligned} \quad (2)$$

或

$$P\{| \xi | \leq a\} = \operatorname{erf}(k) = 1 - \frac{1}{2N} = \frac{2N-1}{2N} \quad (3)$$

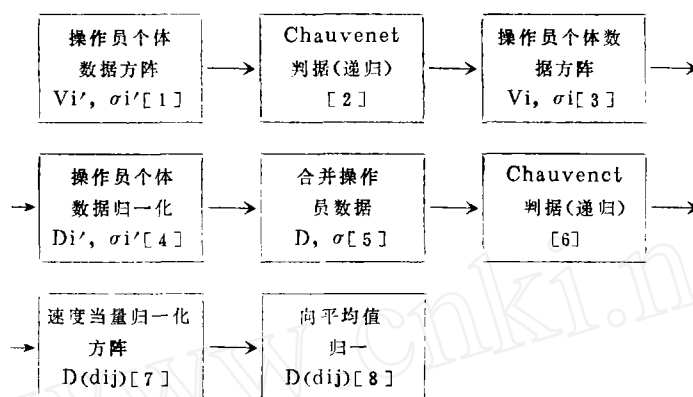
式中  $\operatorname{erf}$  为  $K$  值的补余误差函数。由式 (2) 或式 (3) 可得到表-2 所列的对应于不同采集次数的  $K$  值。

表-2 非正常数据剔除标准

$N$	$K = a/\sigma$	$N$	$K = a/\sigma$
10	1.96	22	2.78
11	2.00	24	2.82
12	2.04	26	2.84
13	2.07	28	2.87
14	2.10	30	2.89
15	2.13	40	2.90
16	2.16	50	2.92
17	2.18	60	2.94
18	2.20	80	2.96
19	2.22	100	2.98
20	2.24	185	3.00

数据处理流程如图三所示。

图中 [1] 为从第  $i$  个操作员采集到的数据, 此方阵由  $46 \times 46 = 2116$  组组成, 每组为两个相关码元键入平均时间和方差组成。对于同一组相关码元, 同一个操作员在采集过程不同阶段的 36 次输入过程中, 由于精力集中、疲劳、字符串结构、外界干扰等众多因素, 难免存在少数非正常数据, 经肖维涅判据处理将这些数据剔除后得到每个操作员的数据方阵 [3]。因为要对每一组数据反复作期望值和方差值运算后才能进行判决, 因此这是一项计算量较大的工作。每个操作员的数据在 Sun386i 工作站上处理约需一小时以上, 因此从 [1] 到 [3] Sun-



图三 键位相关速度当量数据处理流程

386i 需运行 40 小时以上。

每个操作员的击键速度不同，每台采样用计算机的时钟频率也不可完全一致，因此必须对每个操作员的数据进行归一化处理使每一个数据成为相对时间值才能进行数据合并，第二次肖维涅判据用于剔除远离群体特性的个别数据，从而得到反映群体共性的 D 矩阵。

在 D 矩阵（表）中，定义两个最快搭配键位的连击速度为 1.0，给出 46 个键位各种组合的连击速度当量。表体数据为先击纵列键位后连击横列键位的速度当量。

表 1 的相对或然误差  $\sigma/V$  或  $\sigma/D$  均小于 0.05。

## 参 考 文 献

- [1] 李慧勤：计算机汉语输入技术的研究 博士论文 1989
- [2] 陈文熙：刺激-效应的统一模型 心理学报 1984 年第三期
- [3] 林少宫：基础概率与数理统计 人民教育出版社 1978
- [4] 冯师颜：误差理论与实验处理 科学出版社 1964
- [5] 张世箕：测量误差及数据处理 科学出版社 1979
- [6] Hyman e.Stimnlus; Information as a Determinant of Reaction time Journal of Experimental Psychology Vol.45, No.3 1953
- [7] Harald Cramer; The Elements of Probability Theory and Some of its Application John Wiley Sons Inc. 1955

（下转第 11 页）

## **AUTOMATIC PROCESSING CHINESS WORD**

Wang Yongcheng, Su Haiju, Mo Yan

*(Computer Application Technology Institute of Shanghai Jiao Tong University)*

### **ABSTRACT**

In this paper, a summary on the advances of Chinese word processing and several main methods of Chinese automatic word segmentation is made, and a new better method identifying Chinese word is also proposed. Dissecting Chinese sentence to several segments, extracting key words by proper matching methods and AI, and correctly handling remaining strings in proper way, these all are the key points of the new method.

---

(上接第18页)

## **THE INVESTIGATION OF THE RELATIVE SPEED EQUIVALENCE OF THE CHINESE CHARACTER KEYBOARD ENTRY**

by

Chen Yi-Fan Zhang Lu

*(Beijing Information Technology Institute)*

Zhou Zhi-Nong

*(Beijing Electronic Technology Service Center)*

### **ABSTRACT**

In this article, It is introduced how a global ASCII Keyboard succession character entry speed equivalence matrix, which is relative to the code location, is figured out from more than two million experiment data. Keyboard code location relative speed equivalence is a basic data for optimizing Chinese-Character keyboard entry and can be used to compare various Chinese-character entry methods scientifically. This article also explains the design principle for collecting above experiment data and the method for data treatment.

**KEYWORDS:** Keyboard code location relative speed equivalence, Returning to one for minimum value, Time relative value.