



# 基于多特征多分类器的汉语手指字母流的识别系统<sup>1)</sup>

吴江琴<sup>1</sup> 高文<sup>1,2</sup> 陈熙霖<sup>1</sup>

<sup>1</sup>(哈尔滨工业大学计算机系 哈尔滨 150001)

<sup>2</sup>(中国科学院计算技术研究所 北京 100080)

(E-mail: wujiangqin2000@yahoo.com)

**摘要** 手指语是用手指指式进行交流,一个指式代表一个汉语拼音字母,按照汉语拼音方案拼成普通话.文中提出了一种基于多特征多分类器的汉语手指语识别方法,并利用该方法建造了手指字母流识别系统.实验表明,该方法的识别效率明显优于基于单分类器的识别方法.

**关键词** 手语识别,多特征,多分类器,手指语,神经网络.

## CHINESE FINGER ALPHABET FLOW RECOGNITION SYSTEM BASED ON MULTI-FEATURES AND MULTI-CLASSIFIER

WU Jiang-Qin<sup>1</sup> GAO Wen<sup>1,2</sup> CHEN Xi-Lin<sup>1</sup>

<sup>1</sup>(Department of Computer, Harbin Institute of Technology, Harbin 150001)

<sup>2</sup>(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

(E-mail: wujiangqin2000@yahoo.com)

**Abstract** Finger-spelling language is the language that uses finger patterns to communicate. One finger pattern stands for one Chinese spelling alphabet, which is spelled into mandarin according to Chinese spelling scheme. In this paper a learning and recognition algorithm of Chinese finger alphabet flow based on multi-features and multi-classifiers is proposed, and a Chinese finger alphabet flow recognition system is built on the algorithm. Experiment shows that the system apparently outgoes the recognition system based on a single classifier.

**Key words** Sign language recognition, multi-features, multi-classifiers, finger-spelling language, neural network.

1) 国家“八六三”高技术研究发展计划(863-306-ZT03-01-1)、国家自然科学基金(69789301)、国家教委跨世纪人才基金和中国科学院百人计划资助课题.

收稿日期 1999-06-07 收修改稿日期 2000-11-09

## 1 引言

手语是聋人交际的工具,是由手势辅之以表情姿态为符号构成的比较稳定的表达系统,是一种靠动作/视觉交际的特殊的语言.手语识别与手语合成相结合可实现聋人与周围环境的交流.

目前国际上从事手语手指语研究的主要有台湾大学的 Liang 与 Ming,利用 VPL 型号数据手套作为输入设备,采用窗口模板匹配策略实现了美国手语 26 个字母连续流的实时识别<sup>[1]</sup>.

中国手语手指语是由字母语言发展起来的,用手指指式进行交流,用一个指式代表一个汉语拼音字母,按照汉语拼音方案拼成普通话<sup>[2]</sup>.中国手语共有 30 个手指字母(参见文献[2]中的汉语手指字母图),汉语句子中的任意一个词语都可以根据汉语拼音由序列手指字母表示.一些专有名词、地名和人名只能通过手指语进行交流.所以,本文所讨论的手指语识别,即汉语手指字母流的识别,对于手语识别研究领域是至关重要的.

本文第 2 节简要地介绍了手语输入设备——具有 18 个传感器的 Cyberglove 数据手套及 Polhemus 位置跟踪器;第 3 节讨论了汉语手指字母流的分割方法;第 4,5 节提出了基于多特征多分类器的汉语手指字母流的学习识别算法,并在一种多层神经网络快速学习算法——单参数动态搜索算法<sup>[3]</sup>的基础上实现了基于多特征多分类器的汉语手指字母流的识别系统;第 6 节通过实验对基于多特征多分类器的汉语手指字母流识别方法与基于神经网络的单分类器的识别方法进行了比较.

## 2 输入设备

我们在进行汉语手指字母流的识别时采用具有 18 个传感器的 CyberGlove 数据手套作为输入设备,能测量出手指的弯曲与伸展、手掌曲率、手腕的旋转等手及手腕的角度(如图 1).在手腕上添加一个位置跟踪器,能够测量相对于发射器的空间位置以及与  $X, Y, Z$  三个轴的夹角.

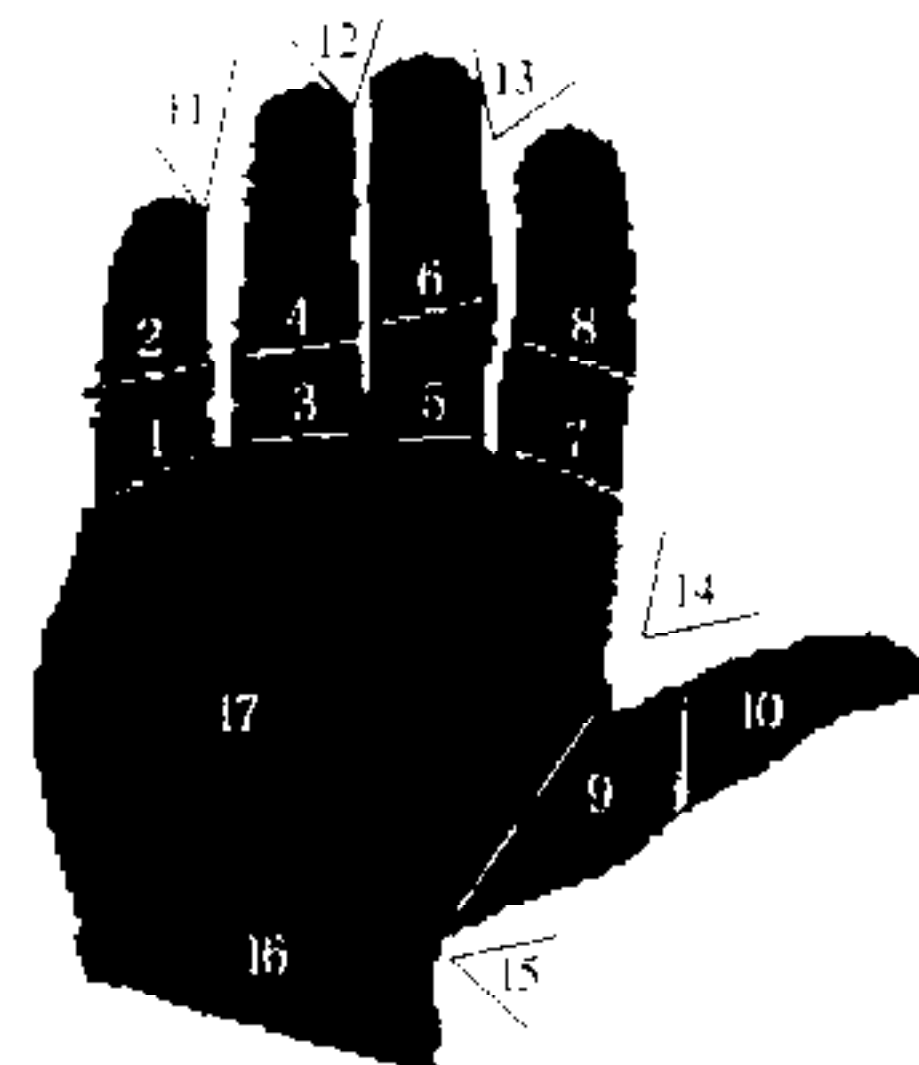


图 1 Cyberglove 测量角度示意图

## 3 原始数据的采集及汉语手指字母流的分割

进行手指字母流识别所需数据是由数据手套及跟踪器输出的 24 维序列.由于字母流中各字母之间的过渡能产生人为噪声,即运动插入,有必要对由数据手套及跟踪器输出的数据进行预处理,以提取表征手指字母的有效帧.

由于人在做每个手指字母动作的时候,在空间上都有一个较为明显的停顿,表现为各维数据都具有一个相对平稳段,进而各维都具有一个共同的平稳段.根据采样数据的平稳特点,采取如下方式进行表征汉语手指字母流的采样数据的分割:

- 1) 对每一维利用一维窗口滤波确定平稳段;
- 2) 确定各维的共同平稳段作为采样数据的平稳段;
- 3) 将采样数据平稳段的均值作为表征手指字母的有效帧数据.

#### 4 基于多特征多分类器的手指字母流的学习与识别算法

根据汉语手指字母自身的特点——单手静止手势,只需从单只手臂(譬如右手臂)的 24 个自由度中选取体现手形特征的 18 个自由度、体现方向特征的 6 个自由度做为汉语手指字母识别的特征. 本系统选取体现手形特征的由数据手套输出的 18 维传感数据直接构成系统的特征向量. 但是,特征向量的选择是否得当,对于一个汉语手指字母流识别系统来说是至关重要的. 因而本文根据汉语手指字母的特点,即某些字母的差距只是若干个自由度值不同引起的,提出了基于多特征多分类器的学习识别算法,即从整个特征集中,选取若干个使得汉语手指字母区分性更大的特征子集,构建用于融合的子分类器,以提高整个系统的识别效率.

为了提高构造多特征多分类器的速度,在构造多特征多分类器的过程中,采用了一种快速多层神经网络学习算法——单参数动态搜索学习算法<sup>[3]</sup>. 它采用单变量轮换搜索方式,即每步搜索只让网络的一个参数变化,进行精确的一维搜索. 根据多层神经网络及误差函数的特点,该算法每步搜索只针对有变化的误差部分进行计算,尽管增加了一维搜索的次数,但大大减少了误差函数的计算量,进而加快了训练的收敛速度.

设特征集  $M$  的基数为  $|M|$ , 模式分类数为  $P$ , 根据各手指字母的具体手形特征,将用以分类的特征集  $M$  分成  $K$  个特征子集  $M_l (l=1, \dots, K)$ . 为了减少分类器个数,这里允许特征子集相交,显然  $|M_l| \leq |M|$ , 且  $\bigcup_{l=1}^K M_l = M$ . 对于每一特征子集  $M_l (l=1, 2, \dots, K)$ , 利用单参数动态搜索算法学习对应于该特征子集的分类器  $C_l$  相应的网络参数  $\{\omega_{ij}^{(l)}\} (i=1, 2, \dots, |M_l|; j=1, 2, \dots, H^{(l)})$ ,  $\{v_{jk}^{(l)}\} (j=1, 2, \dots, H^{(l)}; k=1, \dots, P)$ ,  $\{\theta_j^{(l)}\} (j=1, 2, \dots, P; l=1, \dots, K)$  和  $\{\gamma_k^{(l)}\} (k=1, \dots, P; l=1, \dots, K)$ , 其中  $\{\omega_{ij}^{(l)}\}$  为输入层至隐含层连接权,  $\{v_{jk}^{(l)}\}$  为隐含层至输出层连接权,  $\{\theta_j^{(l)}\}$  为隐含层各单元输出阈值,  $\{\gamma_k^{(l)}\}$  为输出层各单元输出阈值,  $|M_l| (l=1, \dots, K)$  为各分类器输入层单元数即特征数,  $H^{(l)}$  为隐含层单元数,  $P$  为输出层单元数即模式类数.

对于任一由数据采集模块得到的汉语手指字母流  $G$ , 利用数据采集及汉语手指字母流的分割模块得到有效帧序列  $x(1), x(2), \dots, x(T)$ , 对于每一帧数据  $x(t) = (x_1(t), x_2(t), \dots, x_{|M|}(t))$ ,  $t=1, \dots, T$ , 记其基于特征子集  $M_l, l=1, \dots, K$  的数据为  $x^{(l)}(t) = (x_1^{(l)}(t), \dots, x_{|M_l|}^{(l)}(t))$ , 计算其相对于分类器  $C_l$  的实际输出与第  $q$  个模式类 ( $q=1, \dots, P$ ) 理想输出的误差

$$E_q^{(l)}(t) = \sum_{j=1}^P \left( \left( f \left( \sum_{s=1}^{H^{(l)}} v_{sj}^{(l)} \left( f \left( \sum_{i=1}^{|M_l|} \omega_{is}^{(l)} x_i^{(l)}(t) \right) + \theta_s^{(l)} \right) \right) + \gamma_j^{(l)} \right) - t_q \right)^2,$$

其中  $t_q = (t_1^{(q)}, t_2^{(q)}, \dots, t_P^{(q)})$  为对应第  $q$  个模式类的理想输出,  $f$  为 Sigmoid 函数.

若选取误差精度  $\epsilon$  满足当  $E_q^{(l)}(t) \leq \epsilon$  时,  $x(t)$  隶属于第  $q$  个模式类, 则  $\{l: E_q^{(l)}(t) \leq \epsilon, l=1, \dots, K\}$  表示  $x(t) = (x_1(t), x_2(t), \dots, x_{|M|}(t)) (t=1, \dots, T)$  属于模式类  $q$  的分类器集合,  $|\{l: E_q^{(l)}(t) \leq \epsilon, l=1, \dots, K\}|$  为使得  $x(t) = (x_1(t), x_2(t), \dots, x_{|M|}(t)) (t=1, \dots, T)$  属于模式类  $q$  的分类器个数.  $A_q(x(t)) = \frac{|\{l: E_q^{(l)}(t) \leq \epsilon\}|}{K}$  即为  $x(t)$  隶属于第  $q$  个模式类的隶属度, 使得  $A_q(x(t))$  最大的模式类  $q_0(t) = \arg \max_q A_q(x(t))$ , 即为  $x(t)$  所属模式类,  $t=1, \dots, T$ .

## 5 手指字母流识别系统的实现

系统实现的具体过程如下: 1) 由数据手套输出的汉语手指字母流的原始传感数据  $G$ , 通过分割模块(参见第 3 节), 得到有效帧序列  $x(t), t=1, \dots, T$ ; 2) 构造  $K$  个特征子集  $M_l, l=1, \dots, K$ ; 3) 对应于  $M_l$  的有效帧序列记为  $x^{(l)}(t), l=1, \dots, K, t=1, \dots, T$ , 对于每一分类器  $C_l$  利用单参数动态搜索学习算法学习其所对应的参数; 4) 利用分类器  $C_l$  对应的参数计算  $x(t)$  相对于模式类  $q$  的  $E_q^{(l)}(t), q=1, \dots, P$ ; 5) 计算  $A_q(x(t))$  及  $q_0(t)$ ; 6) 参照字母表输出模式类序列  $q_0(t) (t=1, \dots, T)$  所对应的手指字母流.

## 6 实验结果

本实验未考虑手的方向特征, 因而将手形相同, 但手的方向特征不同的手指字母并为一类进行识别. 选取 27 个手指字母(A, B, C, D, E, F(V), G(I), H, J, K, L(R), M, N, O, P, Q, S, T, U, W, X, Y, Z, CH, SH, ZH, NG)的 270 个样本作为训练样本. 分别利用多特征多分类器(这里采用了三个分类器)及基于神经网络的单分类器对汉语手指字母按照从 A 到 NG 的顺序进行 100 次在线测试, 统计结果如表 1 所示. 表中数字表示每个手指字母的统计识别率.

表 1 单分类器与多分类器识别结果对比

字母 \ 算法	A	B	C	D	E	F	G	H	J	K	L	M	N	O
多分类器	100	100	98	100	100	98	100	97	100	100	100	97	100	96
单分类器	100	100	97	100	100	90	100	96	100	100	100	95	98	90
字母 \ 算法	P	Q	S	T	U	W	X	Y	Z	CH	SH	ZH	NG	
多分类器	100	100	98	100	100	100	98	100	100	98	100	100	100	
单分类器	100	100	90	100	100	92	90	100	100	90	100	100	100	

由统计结果可见, 多特征多分类器的识别效果明显优于基于神经网络的单分类器识别结果. 例如, 对于字母 H, X(如图 2), 用区分两个字母的主要特征: 拇、食指间外展, 食、中指间外展以及中、无名指间外展构成的特征子集构造的子分类器的识别效果明显优于用由数据手套输出的原始数据构成的特征向量构造的分类器的识别效果.

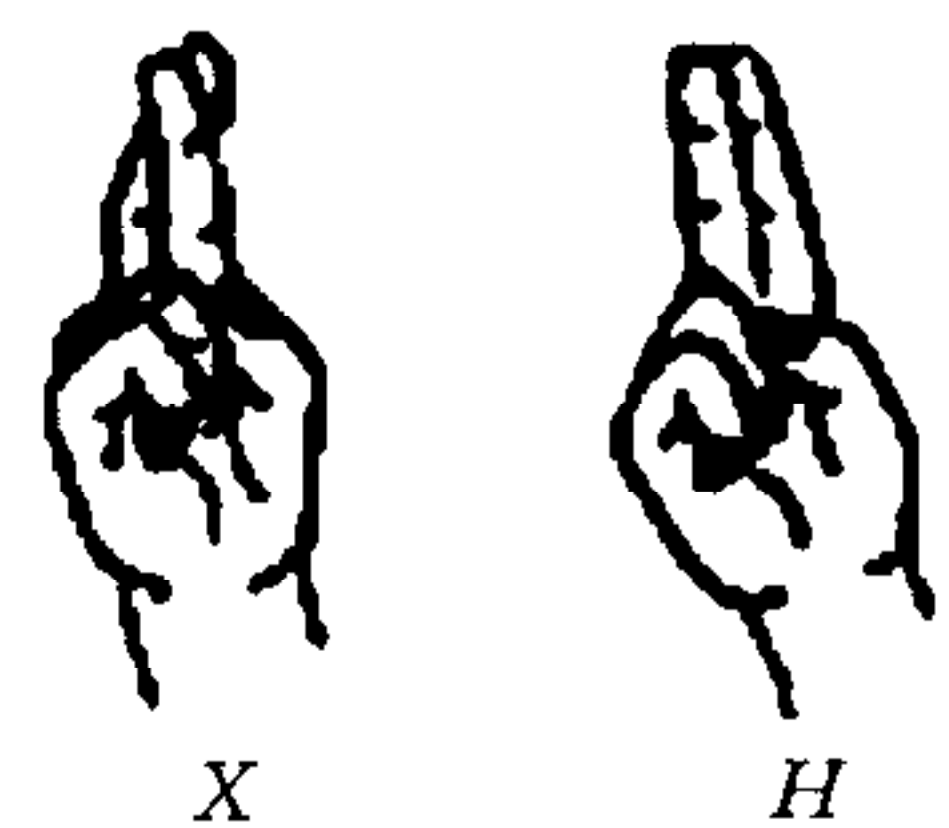


图 2 字母图

实验表明,本文所提出的基于多特征多分类器识别算法,使得汉语手指字母流识别系统的识别效率大大提高.

### 参 考 文 献

- 1 Liang R-H, Ouhyoung M. A real-time continuous alphabetic sign language to speech conversion VR system. *Computer Graphics Forum*, 1995, 14(3):C67~C77
- 2 中国聋人协会. 中国手语. 北京:华夏出版社, 1991, i~xi
- 3 王雪峰,冯英俊. 多层神经网络的一种新的学习算法. 哈尔滨工业大学学报, 1997, 29(2):23~25

**吴江琴** 1965年生,副教授,1991年获哈尔滨工业大学应用数学专业硕士学位,2000年获哈尔滨工业大学计算机应用技术专业博士学位. 研究领域为模式识别与多媒体技术.

**高文** 见本期.

(上接第 805 页)

周彤	林行刚	林作铨	林健	林元烈	林学閤	易继锴	宗光华	金以慧	金芝
金聪	金士尧	武玉强	武际可	范玉顺	郁文生	郑大钟	郑应平	郑南宁	庞国仲
庞哈利	罗跃虎	苗夺谦	姚天顺	姜旭升	姜启源	姜长生	姜平	封兴富	段志生
段广仁	俞立	施颂椒	闻新	钟宜生	胥布工	洪炳镕	洪奕光	胡包钢	胡卫明
胡占义	胡寿松	胡泽新	胡剑波	胡跃明	胡保生	胡昌华	胡恒章	胡德文	赵凯
赵军	赵千川	赵明旺	赵克友	赵杰	赵英凯	赵家璧	赵沁平	荆海英	战守义
费树岷	郝飞	贺国光	贺星钊	郭雷	郭树理	郭晨	徐世杰	徐道义	徐科军
徐德民	徐晓飞	徐乃平	徐波	徐光祐	徐南荣	徐心和	徐文立	徐宁寿	徐立鸿
殷爱茹	诸静	贾沛璋	贾培发	贾英民	贾利民	贾云得	秦世引	耿志勇	袁波
袁震东	袁著祉	袁保宗	钱积新	钱敏平	谈大龙	高黛陵	高龙	高东杰	高存臣
夏文杰	侯格贤	顾冠群	顾仲权	原魁	唐志敏	唐荻	唐云	唐泽圣	唐明
席裕庚	席在荣	柴跃庭	柴天佑	涂健	涂序彦	涂莘生	秦化淑	桂卫华	晏雄伟
陶卿	龚至豪	龚幼民	龚光鲁	龚诚	曹晋华	盖如栋	盛昭翰	梁学斌	梁斌
靳蕃	章毅	章毓晋	章燕申	常文森	常东来	康立山	康景利	梅生伟	葛成辉
阎平凡	程兆林	程侃	程鹏	程代展	程乾生	舒迪前	蒋平	蒋昌俊	蒋慰孙
蒋伯峰	谢贤亚	谢胜利	谢广明	谢惠民	董士海	黄继武	黄亚楼	黄永宣	黄家英
黄泰翼	黄海军	黄心汉	黄正良	黄琳	喻文焕	喻学刚	焦李成	焦文品	彭思龙
彭士戈	曾黄麟	曾建平	曾三友	温香彩	韩战钢	韩正之	韩志刚	韩明	韩兵
韩建达	韩京清	韩崇昭	韩曾晋	解学军	解学书	楚天广	虞忠伟	裘聿皇	裘丽华
褚健	熊有伦	熊光楞	谭跃进	谭民	谭铁牛	管晓宏	蔡开元	蔡安妮	蔡自兴
廖晓昕	慕小武	慕春棣	潘士先	潘启敬	潘永泉	颜文俊	樊晓平	薛安克	薛劲松
薛景瑄	霍伟	魏晨	魏建峰	魏权龄	瞿寿德	戴先中	戴冠中	戴建设	