

基于索引结构的手语词库的设计

于美娟 许 力 刘岩恺 马希荣

(天津师范大学计算机与信息工程学院 天津 300387)

摘 要 随着人机交互应用的日益广泛,手语识别技术得到了很大的重视与发展。基于对当前手语识别技术的研究,针对手语模板库存在的缺点及中国手语的特点,对手语词库进行设计,并通过建立基于索引结构的手语词库,提高了手语识别的准确性和效率。

关键词 人机交互,中国手语,手语词库,索引结构

中图分类号 TP391.4 **文献标识码** A

Design of the Sign Language Template Library Based on Index Structure

YU Mei-juan XU Li LIU Yan-kai MA Xi-rong

(College of Computer and Information Engineering, Tianjin Normal University, Tianjin 300387, China)

Abstract With the increasingly widespread applications of HCI, the sign language recognition technology has gotten a lot of attention and development. On the basic of the research of the current sign language recognition techniques, for the disadvantages about sign language template library and the China sign language features, the sign language template library based index structure was designed. And with this method, the sign language recognition accuracy and efficiency are improved.

Keywords Human-computer interaction, China sign language, Sign language template library, Index structure

1 引言

随着计算机技术和网络技术的迅速发展,计算机在人们的生活和工作中无处不在。我们生活在一个普适计算的环境中,人们已经不满足于简单的计算机指令的交互,而希望计算机能够像人一样与用户进行情感交互,即人们对计算机的智能服务要求越来越高,因此对符合人际交流习惯的人机情感交互技术^[1]的研究有很重要的意义。当前比较常见的人机情感交互技术主要包括人脸识别、表情识别、手势识别、行为识别等,而手势是人们表达信息的重要手段之一,同时也是聋哑人的主要交际手段,因此手势识别技术成为了人们重点研究的课题。

手势识别根据其识别内容的不同可以分为静态手势识别和动态手势识别,静态手势识别即为人的手形识别,动态手势识别指对手势动作或多个手形序列的识别。相对来说,静态手势识别比较简单,只需获取手形的特征并对其进行分析识别即可。根据手势采集设备的不同,可以将手势识别分为基于数据手套的手势识别和基于视觉的手势识别。基于数据手套的手势识别可以利用数据手套和位置跟踪器获得精确的手势信息,识别率高;基于视觉的手势识别^[2]可以通过摄像头、

摄像机等设备采集手势信息,进而对其进行识别,图 1 所示为基于视觉的手势识别的步骤。

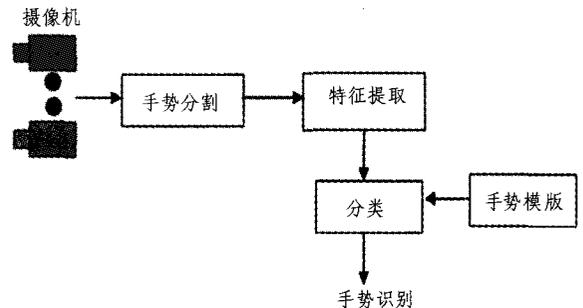


图 1 基于视觉的手势识别结构框图

首先通过摄像机采集手势信息,对手势进行分割并提取手势特征,最后通过与手势模板中手势样本的匹配对手势进行分类,得出手势识别的结果,所以手势识别的第一步工作就是通过采集手势图像建立手势图像库。由手势识别的步骤可以看出,手势模板库对手势识别的准确性和系统性能有重要的作用,而由于不同国家、不同应用领域的手势所代表的含义不同,中国手语具有复杂、多样性的特点,手势识别与人脸识别不同,它没有公开的现成的可供研究者使用的图像库;另

到稿日期:2012-02-21 返修日期:2012-07-15 本文受国家自然科学基金项目(60970060/F020508),天津市科技支撑重点项目(09ZCKFGX00500),天津师范大学校博士基金(52X09012),上海市高可信计算重点实验室开放课题(53H10058),天津师范大学青年基金(52XQ1102)资助。

于美娟(1985-),女,助理实验师,主要研究方向为模式识别、情感计算,E-mail:amy_1121@163.com;许 力(1959-),男,研究员,主要研究方向为计算机应用;刘岩恺(1984-),男,助理实验师,主要研究方向为人机交互;马希荣(1962-),女,博士后,教授,主要研究方向为和谐人机交互、人工智能。

外,至今也没有一个高效的针对中国手语识别的手势词库,因此,建立一个针对中国手语的手语词库具有重要的研究价值和意义。

2 手势模板库的研究现状

手势识别是当前国内外研究的重要课题之一,研究者在对手势识别技术进行研究的同时,针对实验的特点及识别算法的不同,对手势模板库进行了相应的设计。例如西南大学的研究者在对基于切线距离的手势识别^[3]进行研究时,为验证算法的效率,对手指语图像库进行建立,主要实现了对中国手语中字母手势图像库的建立,并划定出测试集和训练集。其具体实现方法为:首先采用黑色背景,使用数码相机对字母手势图像进行采集;然后对照片做一些简单处理,包括图像大小归一化、对图像格式的转化等;最后对其进行存储建立字母手势图像库。还有研究者在对动态手势模板库进行建立时,利用光流计算及量化方法^[4],将一个手势视频序列转化成一个有序的方向直方图系列,其中包含整个视频中人手运动信息的统计特征,并在训练过程中,使用高斯金字塔光流算法来构造手势的直方图序列,对手势模板库进行更新。在基于手形特征的静态手势识别^[5]中,研究者在对手势模板库进行建立时,采用简单的黑色背景,并采用对手势增加附属物的方法去除手势上的冗余信息,如为被识别手势增加与背景颜色相近的附加物以去除手势中手腕部分的影响;为了提高手势的识别率,在大量实验的基础上,对每个手势提取一组特征,作为模板库的存储数据。在样本库数据的存储方面,为了缩小数据库中的存储量,有研究者将图像和反映图像特征向量分开存储^[6]。库中的图像包括数码相机拍摄的手势图片,样本的标准图像统一存在一个固定的目录下(…\标准),而每类图像的平均标准图像则统一存在另一个目录下(…\标准\类),图像和与其所属类别的关联可用数据库的两个表来表示。另外,考虑到各类图像的特征向量随着图像库的变化而变化,所以各类图像的特征向量则统一存放在一个库文件中(…\标准、类\ImageBase.db),其中样本图像及其所属类的关联用 Sample 表来表示,标准图像与类名称的关联用 Class 表来表示。国外也有一些对于手势模板库的研究,如针对美国手语的模板库的建立,针对数字手语的手势模板库的建立等。

根据当前国内外的有关手势模板库的研究现状,我们了解到当前手势模板库存在的问题有以下几点:首先,没有统一的针对中国手语的模板库,这给研究者在对手势识别技术进行研究时的实验带来很大麻烦;另外,目前针对静态手势的手势模板库比较简单,而针对动态手势的手势模板库还没有很好的设计与实现,尤其是很难保证其模板匹配的效率与准确性。针对这些问题,本文对动态手势模板库进行了详细设计与实现,主要设计思想为将静态手势模板库与动态手势模板库统一为一个,并设计为索引结构,以实现高效的模板匹配。

3 基于索引结构的动态手势词库的设计

基于对中国手语特点的分析,本节对手语词库进行设计,主要设计思想为针对手语的结构构成,建立针对静态手势识别与动态手势识别的综合手势模板库,将手语词库设计为索引表结构,与改进的手语识别算法相结合,提高了手语识别的

整体效率与准确率。

3.1 中国手语的特点

手语是由各手势词构成的,手势词又包括更小单位的词或词素。对于一个具体的手语来说,其可以表示为各手势词的组合,各手势词又可以表示为手形和手势轨迹(位置)的组合,此处运动轨迹是指手势的质心的位置变化。根据对中国手语的分析了解到,中国手语^[7]由简单到复杂可以分为以下几类:第一类为静态手语,即手形不变、运动轨迹(位置)不变,如爸爸、妈妈、数字一到十等;第二类为手形不变、运动轨迹发生变化,或手的位置不同,如吃、不、要是手形不变、运动轨迹发生变化,想和糖是手形相同,但手指所点位置不同等;第三类为手形变化、运动轨迹不变,如数字二十、三十、……、九十、谢谢等;第四类为手形和运动轨迹都发生变化。下面根据中国手语的特点对手语词库进行设计。

3.2 手语词库的设计

根据中国手语的特点,对手语各手势词进行分类存储。根据手势的手形和运动轨迹对其进行分类:首先按静态手形的不同对手势进行分类,可以直接满足第一类静态手语的识别匹配问题;然后按运动轨迹的不同进行分类,通过静态手形与动态轨迹的组合可组成不同的动态手势,满足第二类手语的识别;针对第三类手势可以通过静态手势之间的组合进行匹配识别;而第四类手势的识别是最复杂的,在具体识别的实现上,可以对整个手势按照一定的规则进行分段,进行逐段识别再得出最后结果。

在具体实现上,首先对静态手形进行存储,如存储 N 个不同的静态手形类,然后再对不同的手势轨迹进行存储,如定义 M 个不同的运动轨迹,通过手形与运动轨迹的组合,可以定义 $N \times M$ 个动态手势。通过这样简单的组合,可以实现部分简易手语的模板设计。另外对于比较复杂的动态手语,其可以由多个手势的组合构成,组合手语的具体含义可以在单独的一个数据表中进行设计,通过多个手势的组合搜索匹配的手语,并得出其含义。本文通过手语的多层次分割对手势模板库进行设计与实现,建立基于索引结构的模板库。该种结构的手势模板库的主要优点是扩展性强,节省存储空间,并且能够与前台手势识别算法相结合提高手势匹配速度,实现高效的手势识别。

在手语模板库的具体存储上分为 3 级存储,分别为手形的存储、手势轨迹的存储及手语语义的存储。首先在手形存储上本系统中定义了 5 种基本手型,分别表示好、一、二、五(手掌)、八,并分别标记为 a、b、c、d、e,如图 2 所示。

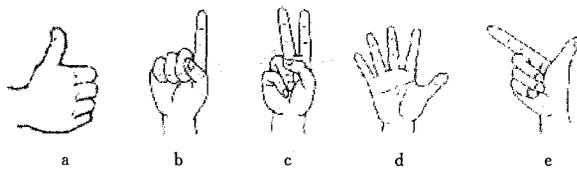


图 2 基本静态手势词

根据手势轨迹的类别,定义 6 种基本轨迹,分别表示上、下、左、右、前、后,然后与手形进行组合,可以组合成 30 种基本动态手势,进而几种动态手势又可组合成不同的复杂手语,并在不同的特定的环境下代表不同的含义,具体实例如表 1、表 2 所列,其中表 1 为由静态手形与动态轨迹组合成的基本

动态手势,表2为由各种基本手势组合成的各种类型的手语的具体含义。表1存储了5种静态手形,6种基本轨迹,它们共组合成30种基本动态手势。表2为具体的各种手势及其含义的存储,首先有静态手语,如a、b、c分别代表“好”、“一”、“二”,简单手语d1,即为手掌向前,表示的含义为“前进”,进而由几种简单手势组合成各种复杂手语,如d5&d6,即为手掌向左、手掌向右,表示的含义为“不”,手语d5&d6&d2即为手掌向左、手掌向右、手掌向后,表示的含义为“不要”。以此类推,通过这样的不断组合,可以构成很多种复杂的手语,并存储其含义。

表1 基本动态手势

轨迹	手形				
	好(a)	一(b)	二(c)	五(d)	八(e)
向前(1)	好向前	一向前	二向前	五向前	八向前
向后(2)	好向后	一向后	二向后	五向后	八向后
向上(3)	好向上	一向上	二向上	五向上	八向上
向下(4)	好向下	一向下	二向下	五向下	八向下
向左(5)	好向左	一向左	二向左	五向左	八向左
向右(6)	好向右	一向右	二向右	五向右	八向右

表2 手语含义表

手语	含义
a	好
b	一
c	二
...	...
d1	前进
...	...
d5&d6	不
...	...
d5&d6&d2	不要
...	...

建立基于索引结构的手势模板库,一方面可以节省存储空间,如在中国手语中,有很多不同含义的手语都会用到“手掌”这个基本手形,此时只需对手掌进行一次存储,然后对复杂手语进行手形标记的链接即可,如本数据库中手掌的标记为d;另一方面,使用基于索引结构的手势模板库,可以进行高效的手势匹配与识别。如表1所列,在该手势模板库中共定义了5种静态手形,6种动态轨迹(向上、向下、向左、向右、向前、向后),本手势模板库中共存储了表2中所列6种具体手语及其含义。对于静态手语可以直接搜索静态手势词库得出手势匹配结果,如静态手势“好”,直接在手语含义库中搜索a,即可得到识别结果。对于手语“前进”的搜索,首先可以对手形进行搜索,得出静态手形为“手掌”,然后在对手势轨迹进行搜索匹配时,可以根据存储的具体手语缩小搜索范围,即在手语含义库中可知手掌向前、手掌向左及手掌向右是有具体

意义的,在轨迹搜索时只需在“向前”、“向左”、“向右”中进行搜索,并得出手势轨迹结果,进而得出该手势为“手掌向前”,最后在手语含义库中进行搜索得出该手语含义为“前进”。

结束语 从手势搜索匹配过程中可以看出,基于索引结构的手势模板库的设计可以在很大程度上缩小搜索范围,提高搜索效率,实现高效的数据搜索,保证识别系统的准确性与实时性。在对手势模板库进行设计时,分别按静态手形、运动轨迹及手语含义对基于手势词进行分类存储,可以简单地将其应用于静态手势的识别中,而且可以针对手势模板库的使用背景对各分类表进行相应的修改,如手势“八”在普通的手语环境中可定义为数字:“八”,而在作战手语中可将其含义修改为“枪”,在这种情况下只需修改手势语义表即可,修改简单,扩展性强。另外对于比较复杂的动态手语(连续动态手势),可以认为是由多个基本动态手势组成的,同样也可以使用该手势模板库。

本文也存在着一一定的不足,如在对手语进行存储的时候,中国手语在表达含义上不仅仅使用手,有些手语的表达还结合了其他人体部位的表达,如“悲伤”的手语表达是一手虚握于胸部,并转动几下,脸露愁容,即结合了人脸的表情,有些手语的表达要结合手臂或双腿及整个身体的动作进行辅助,这样在对手语模板进行存储时,就需要考虑多方面的因素,在对手语进行识别时要综合考虑并得出其最后结果。另外,该手势模板库只是实现了单手的手语存储,对于双手的手语表达及存储没有充分的研究,其研究重点在于双手的定位与配合,可以通过对双手作标记的方法来实现,相信在未来的研究中可以取得进一步的成果。

参考文献

- [1] Alejandro J, Sebe N. Multimodal human-computer interaction [J]. Computer Vision and Image Understanding, 2007, 108(1/2): 116-134
- [2] Pavlovic V I, Sharma R, Huang T S. Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7)
- [3] 黄国范. 基于切线距离的手势识别[D]. 重庆:西南大学, 2009
- [4] 曹陶科. 基于光流计算和DTW算法的动态手势识别研究与实现[D]. 天津:天津师范大学, 2009
- [5] 于洋. 基于手形特征的静态手势识别[D]. 天津:河北工业大学, 2007
- [6] 郭兴伟. 基于视觉的手势识别算法研究[D]. 上海:上海海联学院, 2003
- [7] 中国聋人协会. 中国手语[M]. 北京:华夏出版社, 2003
- [8] Wang Chi, Wang Hao, Xie Zhao, et al. Exact String Matching With Variable Length of Don't Cares Based on Suffix Tree[A]// International Conference on Information Science and Technology, 2011[C]. Nanjing: IEEE Computer Society, 2011: 109-113
- [9] Kucherov G, Rusinowitch M. Matching a Set of Strings with variable length don't cares [J]. Theoretical Computer Science, 1997, 178(1/2): 129-154
- [10] Chen Gong, Wu Xin-dong, Zhu Xing-quan. Efficient string matching with wildcards and length constraints[J]. Knowledge and Information Systems, 2006, 10(4): 399-419
- [11] 武优西, 吴信东, 江贺, 等. 一种求解MPMG00C问题的启发式算法[J]. 计算机学报, 2011, 34(8): 1452-1462
- [12] NCBI National center for biotechnology information website [OL]. <http://www.ncbi.nlm.nih.gov/>

(上接第180页)