



低成本手语手套设计与中等词汇量连续手语识别

华东师范大学第二附属中学 马浩宇
指导教师 王振堂

一、研究背景

手语识别现今主要有两种较为成熟的技术：基于佩戴式输入设备的手语识别与基于机器视觉的手语识别。尽管目前已经对基于数据手套的手势识别有了较为详尽的研究，但是市面上仍然没有一款产品广泛地被聋哑人所使用。原因之一是数据手套价格较高，普通人难以承受。

二、研究过程

（一）手语手套硬件设计

为了实现手语识别的目标，手语手套需要采集手指与手势两方面的数据。手指数据用来判断手姿，使用的是弯曲传感器，手势数据用来判断手的运动轨迹，使用的是三轴陀螺仪传感器，为了能够对上述两组数据进行接收与预处理，需要使用单片机；为了将数据以无线的方式传送至智能终端，需要使用蓝牙模块。



（二）数据采集与识别系统设计

系统需要达成的设计目标主要有三方面：①为了能实时监测蓝牙方面的连接与数据情况，需要蓝牙数据接收与存储模块；②为了能确定数据的有效性，并保证在手套出现异常（如出现断线情况）时能及时发现，需要数据处理与分析模块；③为了方便地录制样本与了解自己所录制的样本内容，需要样本录制与语料标注显示模块。

（三）算法的应用与分析

现今手语识别的算法主要有动态时间规整（下简称 DTW），深度神经网络（下简称 DNN）和隐马尔可夫算法（下简称 HMM）。本次研究旨在通过比较三种算法之间的复杂度与准确率，寻找最适合中词汇量连续手语识别的算法。

数据维度上，每只手的弯曲传感器数据均为 5 维，陀螺仪数据中加速度 3 维，角速度 3 维，故每只手的数据维度为 11 维，双手的数据总维度为 22 维。

在对手语手势进行了分析后，发现在手语中，左手主要作为辅助手使用，其手指的弯曲度在识别过程中可能影响不大。如果能去除左手的弯曲传感器而不影响最终的识别结果，不仅能减少算法复杂度，更能大大降低系统的成本，契合低成本的研究目的。为了验证这个设想，在分析数据时将提供包含左手弯曲传感器（22 维）的结果与不包含左手弯曲传感器（17 维）的结果，并做出比较。

（四）实时译码

由于本次研究目的为完成实时连续手语识别，在将训练好的模型参数反灌回智能终端后，需要在智能终端上进行实时的译码以达到识别的目的。在实时译码的过程中，本项目将着重研究译码算法的复杂度，并通过测试得出实际处理需要的时间作为结果。

本次研究主要使用了两种译码方式，一种为针对 DTW 的译码，另一种为针对 HMM 的 Viterbi 译码，DTW 算法的优势在于需要的训练样本很少且识别准确率较高，缺点在于算法复杂度高，无法实时处理数据，延展性极差。使用 DTW 进行识别的先决条件在于被识别的数据与模板在手势上必须完全一致，否则识别效果便会大打折扣。GMM-HMM 算法的优势在于算法复杂度低，延展性较好，缺点在于训练样本大，在无语法的情况下识别准确率极低。DNN-HMM 的特点是延展





性好,在有语法与无语法的情况下都能做到较好的识别准确率。未来,手语识别的趋势将是更大的词汇量与更少的语法,也正因为这个原因,DNN-HMM 相较于 DTW 与 GMM-HMM 有着更良好的发展前景。

三、研究结论

本研究实现了设计并制作低成本低功耗的手语手套的目的,利用常见的低成本元器件,如蓝牙模块、陀螺仪、弯曲传感器完成了数据的采集工作,总成本仅 782 元,远低于市场上销售的同功能数据手套。制作出的手语手套功率较低,保守估计连续工作时间为 10 小时。

在本研究中所涉及的模型中,可以使用去除左手弯曲传感器的 17 维数据代替包含左手弯曲传感器的 22 维数据进行模型的训练与识别,不会影响最终的准确率。这样不仅可以降低算法复杂度,还可以降低成本。

使用 DNN-HMM 混合模型代替 GMM-HMM 模型可以有效地增加识别准确率,且在无语法的情况下提高倍数更为显著,符合手语识别的发展趋势,是未来更有发展前景的手语识别方法。

本项目实现了在智能终端上完成实时识别的课题目的,实时译码系统可以在规定时间内完成数据处理,可以保证数据的实时性。

专家评语

本项目立足于手语手套研究,成本低、适用面广,有一定的实用价值,具有很好的社会推广性。

