

摘要

随着计算机的发明，模式识别得以出现和发展。手语识别作为模式识别研究中的一个热点问题，近年来受到越来越多研究者的重视。手语识别，是借助计算机自动将手语信号转换为文本或语音的过程。手语识别具有重要的现实意义和巨大的理论研究价值。首先，手语识别能够在聋哑人与健听人之间架起桥梁，从而促进社会和谐发展。其次，手语是一种相对规范的手势集合，手语识别与其它手势分析问题相比相对简单，可以作为更为普遍的手势分析研究的一个前期平台。最后，手语识别涉及计算机视觉、模式识别、机器学习、智能人机接口等研究领域，其研究有助于促进这些领域内其它类似问题的研究。

手语识别研究经过多年的积累，已经在特定人识别领域取得了很好的结果。然而，当测试者与训练集中所有人手语打法差异较大时，系统性能下降明显。采集足够多的训练数据训练普适模型，能够部分解决该问题。然而，由于不同人做相同手语差异较大，因此，模型训练不易收敛。而且，普适模型参数分布比较平缓，能够对大部分测试者取得较好的识别结果；但是，对于特定用户，其性能与特定人模型差距明显。自适应手语识别利用新用户数据对普适模型的参数进行修正，使模型更适合于新用户。该方法与人类认知事物由一般到特殊的机理相契合。

本文围绕手语识别的自适应问题展开研究。根据自适应数据所属类别是否已知，自适应分为有监督自适应和无监督自适应。有监督自适应中需要已知自适应数据的所属类别，因此需要用户显式采集自适应数据。由于显式的数据采集过程需要用户参与，对系统的易用性造成损害。因此，有监督自适应的核心问题是：如何利用尽可能少的自适应数据对模型参数进行修正。无监督自适应中不需要自适应数据的所属类别，因此自适应数据可以在用户使用系统的同时自动采集，无需用户参与。然而，使用无标号数据之前必须对其进行标注，以确定其类别。因此，无监督自适应的核心问题是：如何有效利用大量的无标号数据对模型参数进行修正。

对于有监督自适应问题，提出了基于基本单元提取的手语识别自适应方法和基于模范均值选择和最大后验概率/循环矢量场平滑 (Maximum A Posteriori/Iterative Vector Field Smoothing, MAP/IVFS) 的手语识别自适应方法。由于基于词根的手语识别方法能够取得同基于词汇的手语识别方法相当的识别结果，因

此,本文提出基于词根的手语识别自适应方法。实验结果证明,同基于词汇的方法相比,基于词根的手语识别自适应方法能够在基本保持原有识别率的基础上,大大降低所需采集的自适应数据数量。进一步,分析中国手语的多数据流和词间片段数据相似的特点,可以对模型均值进行聚类以便得到更底层的手语词编码。根据此编码,可以通过部分手语词样本,生成词汇集中其它词汇的手语词样本,利用这些样本进行模型自适应,能够提高模型的识别率。实验结果证明,该方法能够进一步降低所需采集的自适应数据数量。为进一步减少自适应数据,提出了基于模范均值选择和MAP/IVFS的手语识别自适应方法。通过对手语词模型的均值向量进行聚类,可以提取出模范均值向量子集,进而得到模范手语词子集,该子集能够表征新用户的个性特征。仅采集该子集中的词汇的新用户数据,可以对相应的模型进行自适应。未得到自适应的模型,可以通过模型之间的相关性和得到自适应的模型参数估计得到。

尽管有监督自适应能够以较少的数据对模型参数进行修正,然而,显式的数据采集过程必不可少。无监督自适应可以通过隐式采集自适应数据对模型参数进行修正。对于无监督自适应问题,提出了结合简化多项式段模型(Simplified Polynomial Segment Model, SPSM)和隐马尔可夫模型(Hidden Markov Model, HMM)的手语识别无监督自适应方法和基于假设比较导引交叉验证的无监督自适应方法。HMM适合描述具有明显状态跳转的手语词,对于一些无明显状态跳转的渐变的手语词,其描述能力较弱,这源于HMM帧间数据独立同分布的假设。SPSM能够描述帧间数据的相关性,因此适合于描述另一类手语词。结合SPSM和HMM,对无标号数据进行标注,能够增加标注准确率,进而提升无监督自适应性能。传统的自学习自适应方法中,由于对无标号数据进行标注的模型和待适应模型为相同模型,因此出现错误累积和过适应问题。基于交叉验证的无监督自适应方法通过引入交叉验证思想,将对数据进行标注的模型和待适应模型分割开来,从而避免出现错误累积和过适应。通过引入假设比较,可提高标注的准确率,提升自适应的性能。

本文通过对自适应问题进行深入探讨和研究,为未来手语识别系统真正走向实用化提供了必要的准备;同时,也为其它领域自适应问题的解决提供了借鉴和参考。

关键词: 手语识别; 隐马尔可夫模型; 最大似然线性回归; 最大后验概率