

情感计算和儿童社会情绪能力评测

报告人_郑文明



报告人简介：郑文明，博士，东南大学学习科学研究中心教授、博士生导师，儿童发展与学习科学教育部重点实验室（东南大学）主任。

我的报告集中在情感计算和社会情绪能力的评测方面。情感活动是人类的基本心理活动，伴随在人的学习过程中，就像柏拉图所说，所有的学习都有情感基础。情感在人工智能方面也很重要。早在上世纪 80 年代，人工智能的创始人之一 Minsky 就指出了情感在计算机领域的重要性，2006 年他又出版了一本书《情绪机器 (Emotional Machine)》，再次强调了情感在人工智能中的重要性。在这里，我主要强调，情感具有普遍的存在性，社会情绪能力在儿童的发展过程中具有重要作用，儿童的社会情绪能力而不是他们的 IQ，是他们今后生活和工作能否成功的重要指标。

◎ 研究方法

首先介绍一下目前社会情绪能力评测常用的一些方法，可归纳为 3 种类型，第 1 种是基于生化分析的方法，包括基因检测、神经递质检测；第 2 种是生理分析的方法，比如心电、血压、体温、皮肤电、呼吸、脑电、近红外脑功能成像、功能核磁共振成像等；第 3 种是行

为观察法，包括心理学常用的量表、问卷、访谈、语音、表情、姿态、行为等。这些方法可以单独使用，也可以同时进行。我的研究主要集中在信号处理上，包括视频、语音的处理。我在这里介绍的是从工程技术的角度来讲，如何提高儿童社会情绪能力评测的效率，促进评测在教育领域方面的应用。例如，一段有关延迟满足实验的视频，这段视频如果用传统人工观察的方法需要一直盯着视频，关注小孩眼睛是否一直盯在糖果上，关注时间的长短，分析这样的视频通常需要几个小时甚至更长，耗费大量的人力、物力，而且同一个视频不同的评测人员还有个体差异。我们希望用计算机来进行自动检测、跟踪其行为，减少繁重的工作。在应用领域，有时需要远程监控，如在教室、课堂等，实时处理有着明显的优势。

另外一个例子，是在探究式科学教育的课堂中分析教师与学生的互动。课堂上有很多学生，要观察每一个学生的面部表情及其变化，采用人工方法会耗费很多时间，还缺乏准确性，儿童之间的个体差异也不能实时展现出来。运用面部表情自动识别的手段就能实时呈现。还有一个是有关孤独症儿童的例子，我们在测试他对面部表情的识别和反应，同时也采用了脑电技术，我们希望能用到多种测量手段，包括行为观察，所以也利用了摄像头。采用机器处理的方法可以和脑电进行同步对照评测的结果，这是我们的基本研究思路。

◎ 研究工作的实施

目前主要是结合“做中学”科学教育改革创新项目所开展的探究式科学教育课程进行研究。首先是建立了一个儿童的个性学习发展多媒体数据库，这个数据库不仅为我们的研究（技术研究、机器学习、计算机处理技术）提供技术来源，也为我们探究式科学教育项目的开展提供了

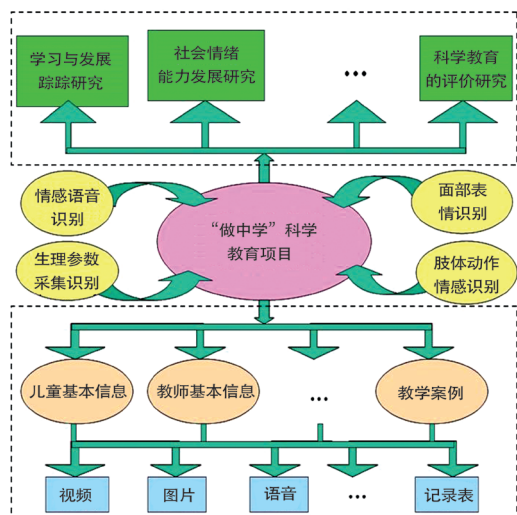


图1 儿童个性化学习与发展数据库结构图

研究方法和数据支持。随着10多年来“做中学”项目的开展，研究过程中我们积累了很多数据，特别是规模的扩大，我们的数据采集从全国4个城市扩展到20多个城市，包括2000多所小学和幼儿园的20万名儿童，积累了非常庞大的数据。基于这样的数据，我们建立了数据库，其基本构成主要有3个部分（图1）。

第1部分是原始数据，是有关教师信息、学生信息、学校信息，还有教学案例，有视频、图像、音频，还有学生在探究活动中的记录单等。教学案例记录教师针对某个主题的教学情况，可能是视频、图像或语音。

第2部分是应用层，帮助教师或家长了解项目的进展情况。其中一个重要的方面是，我们的数据库是个性化的，可以跟踪个体学生的学习和发展情况，家长和学校都可以了解学生的学习进展。这里还涉及到在技术方面如何开发更好的技术为数据处理提供服务的问题，我们采取了一些常用的技术，如视频检索或图像检索、表情识别、语音识别等，这都是技术上的支持。视频、音频和记录单记录了学生在学习过程中的情况。

第3部分是有关学生的个性化学习过程，主要记录学生在探究活动中学习的科学概念、探究技能和社会情绪能力的一些情况。以“小小水果篮”教学案例为例，需要了解的科学概念包括水果的名称、表面的光滑度、颜色、形状、重量、多样性等。探究技能包括语言表达能力、

交流能力、观察能力等。社会情绪能力体现在与人合作、理解别人、尊重别人、自我控制等方面。数据库可以随着时间记录学生在不同时期各方面技能的发展变化情况。这里涉及到很多关键技术，例如面部表情和语音识别等，应用层提供了这些技术手段。

关于面部表情，可以直接让计算机识别，可以实时监测儿童面部表情的变化情况。面部表情识别的具体技术不在此详述，其基本思想是：面部表情变化是由面部肌肉的变化引起的，Ekman把人脸部的肌肉划分为44块，每1块肌肉为1个运动单元，共44个运动单元，记录每个单元的实时变化。通过运动单元的识别，再根据对照表和编码表转化为情绪类别。围绕Ekman的工作，从解剖学角度把情绪分为6种基本情绪：高兴、悲伤、惊讶、生气、厌恶、恐惧；也可以从维度的角度划分为效价、唤醒度和强度。我们设计了很多的算法和模型来更好地识别脸部的情绪，可以计算儿童的情绪强度。

还有一种是微表情，它持续时间很短，在视频中可能只有几帧图像，在测谎中常用，因为为了掩饰真实情绪会有微表情的产生。对微表情的识别也是我们正在开展的工作。还有弱表情，其情绪强度非常弱。研究这2种表情的方法，就是把微表情或弱表情进行放大，以便更好地进行识别。另一方面是语音情感的识别，如何从语言信号中找到情绪的特点，之所以能找到是因为人在不同情绪状态下说的语句，会反映在声学特征和韵律特征方面。

对于生理参数特征例如心率的提取，传统做法通常需要佩戴传感器（心电电极），儿童戴着会有一些不舒服，所以在应用方面会有一些局限，我们希望从摄像头里把各种模态整合起来，语音、表情，包括生理信号整合在一起，这样就可以通过摄像头监测各种情感信号。从视频图像中获得心率的变化情况，其基本思想是，人的心脏在舒张或收缩时，血液在皮肤上会有微弱的变化，眼睛无法识别，但若把变化放大100倍，就会很明显。我们把这些工作和情感、面部表情整合在一起，就可以监测儿童在听一段音乐或观看一段视频时，从行为和生理参数的角度来观察其变化情况。■