

# 基于情感计算的 e-Learning 系统建模<sup>\*</sup>

马希荣<sup>1,2</sup> 刘琳<sup>1</sup> 桑婧<sup>1</sup>

(天津师范大学计算机与信息工程学院 天津 300074)<sup>1</sup> (北京科技大学信息工程学院 北京 100083)<sup>2</sup>

**摘要** e-Learning 也叫数字化学习,是通过因特网或其他数字化媒体进行学习与教学的活动。情感计算是指关于情感、情感产生以及影响情感的计算,试图创建一种能感知、识别和理解人的情感,并能针对人的情感做出智能、灵敏、友好反应的计算机系统。本文将 e-Learning 系统和情感计算结合在一起,提出了一个基于情感计算的 e-Learning 系统模型,旨在有效地解决 e-Learning 系统中情感交流匮乏的问题。

**关键词** e-Learning, 人脸表情识别, 语音情感识别, 情感计算

## The e-Learning System Model Based on Affective Computing

MA Xi-Rong<sup>1,2</sup> LIU Lin<sup>1</sup> SANG Jing<sup>1</sup>

(College of Computer and Information Engineering, Tianjin Normal University, Tianjin 300074)<sup>1</sup>

(Information Engineering School, Beijing University of Science and Technology, Beijing 100083)<sup>2</sup>

**Abstract** e-Learning is a kind of study system. In this system we can teach and learn through the Internet and other digitized media. Affective computing is related to, arises from and deliberately influences emotions. It tries to construct a kind of computer system which can recognize, synthesize and understand the human emotions to make intelligent, smart, friendly reaction. In this paper we combine the e-Learning system with Affective computing to propose the e-Learning system model based on Affective computing. This system can resolve the problem of communication deficiency between the computer and users effectively.

**Keywords** e-Learning, Facial expression recognition, Emotion Recognition of Speech, Affective computing

## 1 引言

进入 20 世纪 90 年代以来,多媒体技术和因特网的日益普及,信息技术正在改变人类社会的生产方式、工作方式、生活方式和学习方式,而且这种变化仍旧与日俱增。随着 e-Learning 的提出并迅速在全世界推广,现代信息技术对人类学习方式产生了深刻的影响并由此而产生了教育、教学领域的巨大变革。同时,由于情感计算的迅速发展,使得 e-Learning 与情感计算的融合成为了可能。在此基础上,我们提出了基于情感计算的 e-Learning 系统模型,将人脸表情识别技术、语音情感识别技术应用于 e-Learning 系统中,大大地弥补了 e-Learning 系统在情感交流和反馈方面的不足。下面我们主要从人脸表情识别和语音情感识别两方面论述情感计算在 e-Learning 系统中的应用。

## 2 e-Learning 概述

e-Learning 就是利用现代信息技术手段,通过信息技术与学科课程的有效整合来实现一种理想的学习环境和全新的、能充分体现学生主体作用的学习方式,从而彻底改革传统的教学结构和教育本质,达到培养大批高素质人才的目的<sup>[1]</sup>。即 e-Learning 主要包括三个内涵:首先,现代信息技术为 e-Learning 提供必要的硬件条件;其次,信息技术与课程整合为 e-Learning 提供先进的有利于建构理想学习环境和学习方式的教学设计方法;第三,教学深化改革与创新人才培养则是 e-Learning 的最终目标。其中“信息技术与学科课程整合”是起关键作用的核心要素<sup>[1]</sup>。

e-learning 系统特定的教学原则是,创设真实情境,开放学习、随机学习和发现式学习。e-Learning 概念影响着社会的方方面面,对现代教学的影响尤为深刻。

首先,是对教学结构的影响。e-Learning 改变了传统的以教师为中心的教学结构,创建既能发挥教师主导作用,又能充分体现学生主体作用的新型教学结构(即“学教并重”教学结构或“主导—主体”型教学结构)<sup>[1]</sup>。

其次,是对教学模式的影响。随着 e-Learning 的深入实践,一些新型的教学模式,诸如:探究型学习模式、案例研习学习模式、发现学习模式、资源型学习模式、协作型学习模式和虚拟教学模式等正在被广泛地应用和推广。

尽管 e-Learning 系统有很多优点,但是它忽略了学习者自身的能力。在 e-Learning 环境下,教师和学生空间上是分离的,他们之间缺少情感交流和反馈,不利于为学生选择适合他们自身条件的知识。为此,我们在 e-Learning 系统中引入情感计算技术,以加强师生的情感交流。

## 3 情感信息的获取和识别

情感计算是建立和谐人机环境的基础之一,美国 MIT 媒体实验室 Picard 教授领导研究小组最早开始情感计算研究工作,并提出“情感计算是指关于情感、情感产生以及影响情感的计算”,它试图创建一种能感知、识别和理解人的情感,并能针对人的情感做出智能、灵敏、友好反应的计算机系统<sup>[2]</sup>。目前情感计算的研究普遍受到学术界和企业界的关注,国际知名大学也纷纷建立情感计算研究小组,情感计算将有巨大的发展潜力和应用价值。但总的说来,情感计算的研究目前

<sup>\*</sup> 基金项目:天津市科技攻关重点课题(编号:04310731R)。马希荣 教授,博士后,主要研究领域为:多 Agent 系统,情感计算。刘琳,桑婧 硕士研究生,主要研究方向:多 Agent 系统,远程教育。

还处于起步阶段,主要侧重于情感信号的获取、情感计算的应用和可穿戴计算机等方面。关键技术包括:人类生理、心理及行为特征的情感状态分析,人类情感信息信号采集传感器的研究、人类情感的识别技术等。其中人类情感的识别又包括人脸表情识别、语音情感识别和身体姿态识别等技术。

人的情感是复杂多样的,所以情感分类的方法也很多,除了有主要情感和次要情感之分<sup>[3]</sup>之外,Plutchik 等人还提出了“情感轮”<sup>[3]</sup>的概念。而且作为人类情感的不同表达方式,人脸表情和语音情感也分别有各自的分类特点。

### 3.1 人脸表情的识别

人脸表情识别就是对人脸的表情信息进行特征提取分析,按照人的认识和思维方式加以归类和理解,利用人类所具有的情感信息方面的先验知识使计算机进行联想、思考及推理,进而从人脸信息中分析理解人的情绪。

人脸的表情是多种多样的,一般对表情的分析可以从维量分析和分类这两种角度来研究,人类表情学家 Paul Ekman 等人经过研究将人类的表情划分成六类,分别是高兴、惊奇、恐惧、悲伤、厌恶和愤怒。人类所有的情绪表情都是在以上六种表情的基础上,经过复杂的融合之后产生的<sup>[4]</sup>。

关于计算机面部表情识别的研究是在最近几年才逐渐发展起来的。人脸表情识别包括两大重要技术,即人脸的检测和定位技术以及人脸图像的提取和识别技术。其中人脸的检测和定位技术相对简单,而人脸图像的提取和识别比较复杂,因此得到了更为广泛和深入的研究。

目前,利用计算机进行人脸表情识别主要有三种方法:

(1)基于几何特征的人脸表情识别方法 该方法主要对人脸表情的显著特征,如眼睛、鼻子、眉毛、嘴等的位置变化进行定位、测量,确定其大小、距离、形状、相互比例等特征,然后进行表情识别。这种技术主要是通过对面部的几个特征点信息的采样来收集信息,但是,采用这种技术有可能造成一些重要的表情识别信息和分类信息丢失,因此,这种方法并不是特别理想的人脸表情识别方法<sup>[4]</sup>。

(2)基于模型的识别方法 该方法主要是建立精确的物理模型,根据解剖学知识确定关键特征并通过对关键特征采样等方法,并通过比较这些特征的变化来识别人脸表情,识别过程采用图像数据驱动。

(3)基于整体面像特征的识别方法 该方法通过对整幅人脸或人脸图像中特别的区域进行数学建模,获取人脸各种表情的特征来进行识别。这种人脸表情识别技术需要坚实的数学基础,其中应用最多的是主元分析的统计方法。与基于几何特征的识别方法相比,此方法更注重尽可能地保留原始人脸表情图像的信息<sup>[4]</sup>。

### 3.2 语音情感的识别

语音情感处理主要是通过测量、分解、分析、合成等方法进行,其目的是识别、理解以及合成语音信号中的情感分量,进而使计算机具备一定的感情能力。主要研究内容包括:语音情感的识别、语音情感的合成、与其他情感信息的融合,以及在相关应用中的实现。其中,语音情感的识别是语音情感处理的基础部分,通过模式识别技术,可对语音信号中的特征参数和情感之间的复杂对应关系进行深入理解。

目前在语音情感识别的研究中常见的归类方法有“八大情感”或“四大情感”,我们采用的是愤怒、高兴、悲伤、害怕四大情感<sup>[5]</sup>,即,任一时刻的情感状态由四种基本情感混合组成,每种基本情感都是可以用“无、似乎有、有一些、有、确定有、很强烈”等状态进行刻画。与人脸表情一样,语音情感也是连续变化的,每种情感状态对应连续空间内的唯一坐标。

利用计算机进行语音情感识别主要包括:情感语音数据库建立、语音情感特征参数提取、语音情感模式识别、模式训练等。其基本流程如图1所示。

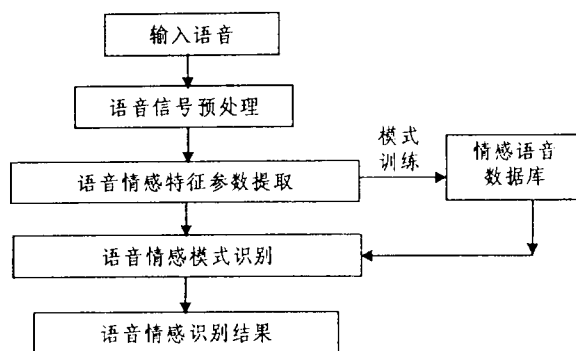


图1 语音情感识别基本流程图

(1)情感语音数据库的建立 不论是语音的情感识别还是合成,都需要一个真实、高效、数据丰富的情感语音数据库。语音素材可以通过多种途径获得,主要有两种,一种是录音法,通过邀请一些专业播音员或善于表达情感的人朗读一些指定的句子,采用录音的方法获取情感语言;另一种是剪辑法,主要从电影电视等多种媒体信息中截取包含情感的语音片断,相对而言,这种方法的情感真实度更高<sup>[5]</sup>。

(2)语音情感特征参数的选择 语音情感特征参数用来刻画语音情感状态。目前,国内外在语音情感特征参数选择方面已有许多研究成果。通过比较,我们尝试选择基音频率、共振峰、能量作为主要参数,同时衍生出平均值、最大值、平均变化率等其他参数。这样,根据实际的识别效果,确定了10个语音情感特征参数,包括:语句发音持续时间、平均振幅能量、振幅能量的动态范围、平均基音频率、最大基音频率、基音频率的平均变化率、共振峰频率的平均值、共振峰频率的平均变化率、共振峰峰值点回归直线的平均斜率以及共振峰峰值的平均值<sup>[6]</sup>。

(3)语音情感识别模型的建立 建立语音情感识别模型是语音情感识别的关键,其核心内容就是分类器的选择和使用。许多模式识别方法<sup>[5,6]</sup>都可应用于语音的情感识别,如:主元素分析法(PCA)、线性判别分类(LDC)、K-近邻法(KNN)、隐马尔可夫模型(HMM)和连续的隐马尔可夫模型(CHMM)、神经网络算法(BP)、支持向量机SVM、融合算法等。

此外,随着计算机图形技术和多媒体技术的飞速发展,作为人类情感表达的一种方式,人体运动姿态识别<sup>[7]</sup>技术也得到迅猛发展,在e-Learning系统的情感识别和交流中必将发挥重要作用。

## 4 基于情感计算的e-Learning系统

在教与学的过程中,人类的表情和语言是丰富的,当教师讲授的内容我们能够理解和接受的时候,会出现情绪高涨的状态,学生会用微笑等表情或欢快的语言来表示;反之,情绪会低落,学生可能会用眉头紧锁、目光呆滞、低沉丧气等方式来表示他们无法理解讲授的知识。

学生的这些情感表现是很重要的反馈信号,为了在e-learning学习环境中,有效地解决情感交流匮乏的问题,我们开发了基于人脸表情识别与语音情感识别技术的e-Learning系统。该系统将对学生的感情信号进行捕捉和识别,并以此

作为重要依据,判定学生对所学知识的接受情况。

#### 4.1 系统模型

根据以上讨论,我们给出了基于情感计算的 e-Learning 系统模型(如图 2 所示)。

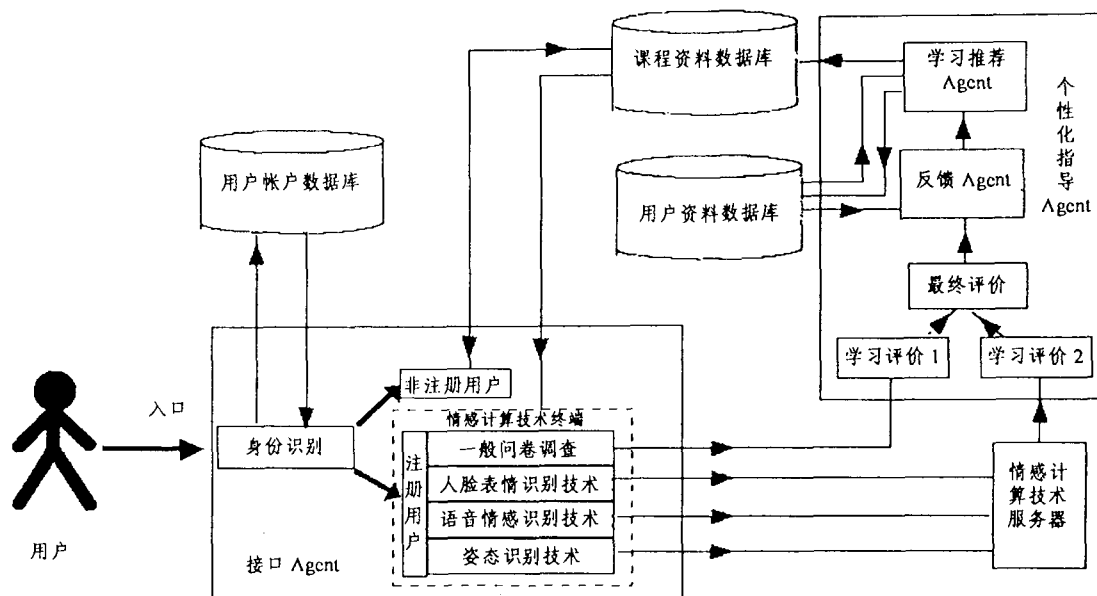


图 2 基于情感计算的 e-Learning 系统

基于情感计算的 e-Learning 系统模型可分为两个大的部分:用户信息导入部分和系统指导学习部分。用户信息导入部分包括接口代理、情感计算服务器和用户帐户数据库,系统指导学习部分包括个性化指导代理、课程资料数据库以及用户资料数据库。

#### 4.2 用户信息导入部分

作为用户与计算机的接口,这部分主要负责对用户的身分进行验证,收集用户反馈信息、呈现课程资料和分析处理用户反馈信息等。

用户帐户数据库主要用来存储注册用户的资料,例如用户名、密码、邮箱地址和用户的其他基本信息等。

接口 Agent 模块由身份识别、一般问卷调查、情感计算终端(包括人脸表情识别技术、语音情感识别技术、姿态识别技术)等部分组成,主要负责验证用户身份,收集用户反馈信息和呈现课程资料。

情感计算服务器对收集到的情感信息分别进行情感识别,再通过统计方法计算出在该段学习中表情、语音的种类和出现的比例,识别出用户的整体情感反应,并作出适合的学习评价。

情感计算终端和情感计算服务器部分是本 e-Learning 系统模型的特色部分。用户在学习过程中要被情感计算终端实时监测,并将收集到的关于人脸表情、语音情感等信息传递给情感计算服务器,在一段学习结束后,自动统计出出现的表情的种类、语音的种类以及所占的比例。用户可以主动填写问卷调查以形成学习评价 1,以调整学习进度;系统在认知心理学的支持下,形成学习评价结果 2;一般情况下,计算机会自动监测并分析用户的学习情况,自动调整学习进度,用户无需自己去填写问卷。因此,在本系统中,学习评价 1 的优先级要高于学习评价 2。

#### 4.3 系统指导学习模块

系统指导学习模块主要是根据“用户信息导入”传来的反馈信息,为学生寻找合适的课程资料。该部分由个性化指导代理、用户资料数据库和课程资料数据库三部分组成。课程

资料数据库包括各种优秀的课程资料和信息;用户资料数据库则保存了注册用户所有的学习评价、学习进度和对学习资料的访问历史。

个性化指导代理又由反馈代理和学习推荐代理组成,反馈代理主要是负责收集从用户信息导入部分传过来的诸如学习评价之类的结果,然后转化成相应的评价参数,同时从用户资料数据库中调出该用户前一次学习的记录,一起传递给课程推荐代理;学习推荐代理接受参数和该用户前一次学习的记录,并据此进行调整,从课程资料数据库中搜索出适合用户水平的学习资料传递给接口代理,这样,学生就可以根据自身水平及时调整学习进度,得到适合自身水平的学习资料。其中与该注册用户相关的学习评价、学习进度和各种各样的对学习资料的访问历史等都会被记录到用户资料数据库中,以此作为下一次学习评价和学习资料选择的依据。

**结束语** 情感计算作为刚刚兴起的技术得到了广泛关注,将其融入 e-Learning 系统中必将使学习方式产生巨大变革。基于情感计算的 e-Learning 系统以人脸表情、语音以及姿态等多种情感表现作为反馈信息,应用于教学策略调整,可为用户提供个性化的学习系统环境,从根本上改变了我们的学习,在网络时代的今天,基于情感计算的 e-Learning 系统必定会有更大的发展。

#### 参考文献

- 1 何克抗. e-Learning 的本质——信息技术与学科课程的整合. 电化教育研究, 2002(2)
- 2 Picard R W. Affective Computing[M]. MIT Press, London; England, 1997
- 3 罗森林, 潘丽敏. 情感计算理论与技术. 系统工程与电子技术, 2003, 25(7)
- 4 何鹏, 陶建华, 谭铁牛. 看懂人的“脸色”——人脸表情处理技术. 计算机世界报, 2002, (05)
- 5 王青, 谢波, 陈根才. 基于神经网络的汉语语音情感识别. 浙江大学, 2003
- 6 赵力, 王治平, 卢韦, 邹采荣, 吴镇扬. 全局和时序结构特征并用的语音信号情感特征识别方法. 自动化学报, 2004, 30(3)
- 7 季白桦, 袁修干, 温文彪. 三维人体运动数据提取的人机交互方法及实验. 北京航空航天大学学报, 2000, 26(1)